

Measuring Brain Activation Patterns from Raw Single-Channel EEG during Exergaming: A Pilot Study

*Original*

Measuring Brain Activation Patterns from Raw Single-Channel EEG during Exergaming: A Pilot Study / Amprimo, Gianluca; Rechichi, Irene; Ferraris, Claudia; Olmo, Gabriella. - In: ELECTRONICS. - ISSN 2079-9292. - ELETTRONICO. - 12:3(2023), p. 623. [10.3390/electronics12030623]

*Availability:*

This version is available at: 11583/2977770 since: 2023-04-05T12:45:27Z

*Publisher:*

MDPI

*Published*

DOI:10.3390/electronics12030623

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

## Article

# Measuring Brain Activation Patterns from Raw Single-Channel EEG during Exergaming: A Pilot Study

Gianluca Amprimo <sup>1,2,\*</sup>, Irene Rechichi <sup>2,\*</sup>, Claudia Ferraris <sup>1</sup> and Gabriella Olmo <sup>2</sup><sup>1</sup> Italian National Research Council, CNR-IEIT, 10129 Turin, Italy<sup>2</sup> Department of Control and Computer Engineering, Politecnico di Torino, 10129 Turin, Italy

\* Correspondence: gianluca.amprimo@polito.it (G.A.); irene.rechichi@polito.it (I.R.)

† These authors contributed equally to this work.

**Abstract:** Physical and cognitive rehabilitation is deemed crucial to attenuate symptoms and to improve the quality of life in people with neurodegenerative disorders, such as Parkinson's Disease. Among rehabilitation strategies, a novel and popular approach relies on exergaming: the patient performs a motor or cognitive task within an interactive videogame in a virtual environment. These strategies may widely benefit from being tailored to the patient's needs and engagement patterns. In this pilot study, we investigated the ability of a low-cost BCI based on single-channel EEG to measure the user's engagement during an exergame. As a first step, healthy subjects were recruited to assess the system's capability to distinguish between (1) rest and gaming conditions and (2) gaming at different complexity levels, through Machine Learning supervised models. Both EEG and eye-blink features were employed. The results indicate the ability of the exergame to stimulate engagement and the capability of the supervised classification models to distinguish resting stage from game-play (accuracy > 95%). Finally, different clusters of subject responses throughout the game were identified, which could help define models of engagement trends. This result is a starting point in developing an effectively subject-tailored exergaming system.

**Keywords:** rehabilitation; RGB-depth cameras; EEG; user engagement; BCI; exergames; Machine Learning; telemedicine; single-channel EEG; low-cost BCI



**Citation:** Amprimo, G.; Rechichi, I.; Ferraris, C.; Olmo, G. Measuring Brain Activation Patterns from Raw Single-Channel EEG during Exergaming: A Pilot Study. *Electronics* **2023**, *12*, 623. <https://doi.org/10.3390/electronics12030623>

Academic Editor: Gwanggil Jeon

Received: 27 December 2022

Revised: 12 January 2023

Accepted: 19 January 2023

Published: 26 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the development of advanced wireless technology and in the wake of the COVID-19 pandemic, telemedicine applications have rapidly grown by virtue of their versatility, portability and accessibility. Indeed, telemedicine allows patients living in remote areas to benefit from continuity of care, while significantly reducing healthcare costs, compared to conventional inpatient and outpatient care [1].

This becomes significantly relevant in the context of neurodegenerative disorders, such as Parkinson's Disease (PD) and Alzheimer's dementia, which are becoming a challenge worldwide. Notably, PD features a prevalence of 1% in the population over 65 [2], expected to rapidly increase as the global population ages [3]. This may become a critical issue, due to the lack of a curative treatment and the complexity of the daily disease management. Patients, indeed, experience a series of motor alterations (e.g., bradykinesia, akinesia, muscle stiffness, tremors, balance impairments) and cognitive decline (e.g., memory loss, reduced communication skills, reduced planning and reasoning, mood swings) that worsen as the disease progresses [4–6]. This translates into reduced independence and need for continuous care and monitoring; hence, high costs for patients and national health systems [7,8], and complexity in the implementation of continuous follow up strategies [9,10] arise.

As the pathophysiology of PD is still unclear, at present symptomatic treatment is the only viable option to improve patient's Quality of Life (QoL) [11,12]. Physical and cognitive rehabilitation is often considered a crucial adjuvant to pharmacological therapy

for the amelioration of symptoms [13–15]. According to some studies, they are also able to slow down the disease progression [16,17]. Physical exercise can improve both motor and non-motor symptoms [18–21] and may even reduce the risk of developing the disease itself in healthy subjects [22,23]. The origin of this positive effect is largely investigated through studies in animal models of PD, which suggest that exercise and learning induce a dynamic interplay between degenerative and regenerative mechanisms [20]. Moreover, increasing evidence suggests that physical exercise reduces chronic oxidative stress and stimulates the synthesis of neurotransmitters and trophic factors [24].

Motor rehabilitation is usually carried out in the context of hospital or gym facilities, under clinical and technical supervision. As an example, motor rehabilitation protocols for PD can include stretching, muscle strengthening, balance and postural exercises, occupational therapy, cueing, treadmill training. Dance (e.g., tango), music-coordinated training and martial arts have proven to be effective alternatives to traditional muscular rehabilitation [25,26].

Cognitive rehabilitation is often coupled with motor rehabilitation to achieve better results [27], since this latter may stimulate neuro-plasticity [28,29]. Cognitive training aims at stimulating specific aspects of cognitive functioning (e.g., memory, complex reasoning) through structured and guided tasks that can be carried out alone or in a group [30]. As previously introduced, a general trend in the last years is the increasing use of technological devices such as tablets, smartphones or computers [31]. This provides several benefits: flexibility (i.e., the same task can be adapted to different applications and needs); high interactivity and more immersive tasks; deployment of the training also through telemedicine and Internet-based solutions; automatic qualitative and quantitative feedback, available as soon as the cognitive task is completed [32].

A similar approach is being sought also for physical rehabilitation, in the form of telerehabilitation, through the employment of so-called *Exergames* or *Serious Games*. Exergames are broadly investigated for motor-cognitive rehabilitation in healthcare [33,34], exploiting new generation devices such as RGB-Depth cameras (e.g., Kinect) [35,36], balance boards (e.g., Wii balance board) [37], and virtual-reality headsets [38,39]. The users can control the game through their own body and carry out goal-oriented tasks, which aim at stimulating specific motor and cognitive skills.

Several systems have been proposed in the last years to combine exergames with traditional rehabilitation protocols and remote assessment tools [35,40,41]. Preliminary results suggest that this alternative form of training could produce a positive effect similar to physical, in-person activity; however, the results are largely subjective and general trends have not yet been identified in the population of patients affected by neurodegenerative diseases [42,43]. Indeed, especially for these latter, rehabilitation protocols must be tailored on the single specific subject to optimise the expected results [44]. Moreover, when carrying out the tasks at home—often with little to none supervision—the patient should be emotionally and mentally committed, or, more generally speaking, engaged.

Engagement is regarded as a facilitator to the fruitful adoption of telemedicine, as it likely translates into continuity of use [45], hence improving the probability of success of the rehabilitation protocol. Therefore, the development of next-generation exergames, able to automatically tune themselves according to the level of engagement and/or mental effort required to the subject, is crucial. For example, the complexity of the required task could be reduced or increased depending on the subject's conditions. To implement this concept, it is necessary to gather information on the physical and mental state of the patient, in a simple and unobtrusive, yet reliable way.

However, to date there is a lack of a stable and reliable framework to assess user engagement based on physiological parameters, both in healthy and pathological subjects. Given the complex clinical picture of neuro-affected subjects, preliminary explorations should focus on the healthy population, with the final goal of developing a solid framework before translating the obtained solution into the medical scenario. Furthermore, in view of

integrating this framework in a telemedicine approach, low-cost, limited computational burden and minimally-invasive-oriented solutions should be investigated.

Among the various bio-signals suitable for retrieving a metric of engagement, electroencephalography (EEG) is considered one of the most significant. It reflects the electrical activity of the brain and is commonly recorded through electrodes placed at the scalp. The electrical waves are further divided into different rhythms (or bandwidths, detailed in Section 2.7.1), each with its own clinical significance. EEG signals are commonly employed in clinical practice to diagnose and monitor brain conditions (e.g., epileptic syndromes) and widely explored in neuroscience to characterise the response to specific events or triggers [46], as well as in Brain-Computer Interface (BCI) technology. In particular, this latter exploits the different activation patterns of brain areas in order to steer devices or collect the subjects' response to specific stimuli.

Various EEG-based indices have been proposed in literature [47,48] to describe engagement. They mainly rely on the power in the  $\beta$ - (13–30 Hz),  $\alpha$ - and  $\theta$ -ranges (8–13 Hz and 4–8 Hz, respectively). The first is usually associated to visual attention and sustained alertness [49], whereas  $\alpha$ -waves are commonly related to relaxed states. As for  $\theta$ -waves, various studies highlighted their role in cognitive processing, task execution and sensory-motor coupling [50,51]. This consists in a task-specific motor output relying on the integration of multi-source sensory information (notably, goal-oriented hand-movements such as grasping [52,53]). These task-related indices have been employed in adaptive frameworks in order to monitor the participants' attention levels [54] or their engagement during different tasks, such as cognitive activities [55,56] or interaction with video-games [57].

In addition to EEG signals, eye-blinking is also considered a clear indicator of attention and concentration [58–60]. Therefore, features related to eye blinking could be significant for inferring the level of attention and engagement in subjects performing the exergaming task.

This method article describes a framework for computing the level of engagement during an exergame, specifically designed to stimulate engagement with a combined physical and cognitive task. The game exploits an Azure Kinect to track the movement of the hand in front of the camera. A commercial EEG headset (Dreem Headband v2.0) is used to evaluate the mental activity of the subject during an initial relaxation phase, to obtain a baseline reference signal, and then during the exergame play. In addition, eye blinks are identified by analysing the video recordings collected by the Azure Kinect, using the Mediapipe open source library.

As a preliminary investigation of this method, before translating it to a clinical scenario, data from 50 healthy volunteers were collected and analysed with the following goals: (i) to compute a set of features from single-channel EEG and eye blinking data to characterize the users' engagement while playing the exergame; (ii) to automatically distinguish the relaxation phase from the game play using shallow classification methods; (iii) to investigate the possibility of finely distinguishing the specific levels of the game, as each of them was designed to elicit a different level of engagement; (iv) to explore the subjects' behaviours and identify trends in the response to the exergame. This could allow to apply the investigated features for automatically tuning the exergame according to the data models inferred from the collected signals.

The rest of the paper is organized as follows. Section 2 describes in detail the observational study designed for this work, including the information about the recruited subjects, the exergame design process, the sensors employed, the data acquisition protocol, and the methods applied to analyse the collected data. The achieved results are presented and discussed in Sections 3 and 4 respectively. Finally, in Section 5 conclusions are drawn, taking into account the limitations of the current study and the future research directions.

## 2. Materials and Methods

An observational study was specifically designed for the purpose of this study. The experimental data included both surveys and physiological signals collected during a motor-cognitive task. The involved subjects, the Data Acquisition Protocol (DAP), the experimental data processing and the motor-cognitive task are detailed in the following subsections.



### 2.1. Subjects

A private database (NeAdEx Dataset) was employed in this study. Data were collected at our R&D Laboratory at the Polytechnic University of Turin (Turin, Italy). It included 50 healthy volunteers (37 males), aged  $26 \pm 4.5$  years. All participants were recruited through the University Study Portal. Table 1 reports the demographic data of the subjects. Inclusion criteria required the ability of reading and/or understanding Italian at the A1 Common European Framework of Reference (CEFR). Exclusion criteria included narcolepsy, insomnia, history or ongoing psychiatric conditions preventing the correct execution of the test and a diagnosis of neurodegenerative disorders.

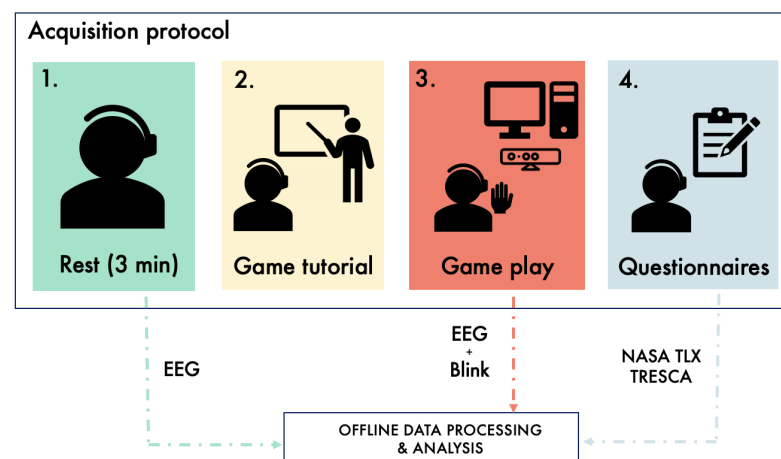
All procedures have been conducted in accordance with the Declaration of Helsinki, supervised by a clinician, and approved by the Ethics Committee of the Hospital A.O.U. Città della Salute e della Scienza di Torino (Approval No. 00384/2020). The participants received detailed information on the study purpose and execution, as well as on the employed instrumentation; informed consent for observational study was obtained.

**Table 1.** Involved subjects and related demographic information.

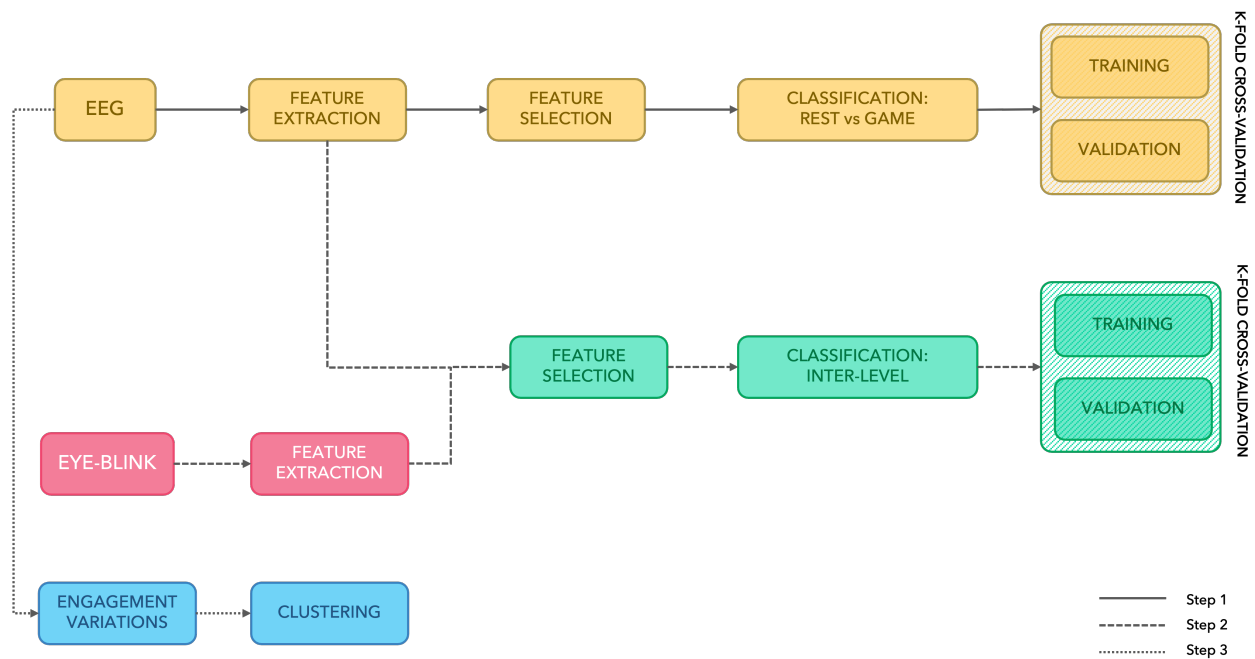
Age	Sex	Education Level
$26 \pm 4.5$ years	37 males (74%) 13 females (26%)	Bachelor's Degree: 13 (26%) Master's Degree: 34 (68%) Ph.D.: 3 (6%)

### 2.2. Data Acquisition Protocol

The experimental data were collected following a specifically-designed DAP. The protocol was organised in a four-stage session, with a total duration of 30 min per subject, including the time required for presenting the study, signing the informed consent and the instrumentation-setup procedure. Figure 1 summarizes the stages of the DAP. First, the participants were asked to rest for three minutes with their eyes closed, to obtain a baseline signal for EEG data. In the second stage, they were instructed about the basic interactions with the game and their goal while playing; however, no preview about the specific challenges of the game levels was provided. In the third stage, the subjects played the game while EEG signals and RGB video were recorded. Once the game ended, they were asked to fill in the questionnaires included in the study—i.e., the NASA Task Load Index (NASA-TLX) [61] and the study-specific Task-Related questionnaire (TRESKA), detailed in Section 2.4. The offline data processing and analysis are described in detail in Sections 2.6.1–2.10; a flow-chart of the different steps carried out in the analysis is displayed in Figure 2.



**Figure 1.** Experimental data acquisition protocol: an EEG baseline is computed from a 3-min resting phase, then the subject is instructed about the game, plays the game and finally answers the NASA and TRESKA questionnaires.



**Figure 2.** Offline data processing and analysis flowchart.

### 2.3. The Grab-Drag-Drop Exergame

For this study, an ad hoc exergame was developed in Unity®(Unity Technologies, San Francisco, CA, USA), named Grab-Drag-Drop (GDD) exergame. It was designed following well-established guidelines for video games targeting the elderly population [62]. The overall goal consists in repeatedly selecting the correct object among four alternatives shown on-screen. The player must grab the object, drag it from its starting position on top of a collecting box, then drop it inside the box before running out of time. Each object is characterized by a shape (i.e., a cube, a sphere, a cone, a cylinder) and a colour (e.g., red, blue, green). A textual message appears in the middle of the screen, instructing the user about the object to select; the time left is also displayed. In addition, the colour of the background progressively turns from green to red as time runs out. Additional stimulation is provided by a pressing background music, which is considered a significant source of stress and engagement in game playing [63]. The human computer interaction (HCI) is based on the GMH-D algorithm [64] to track the motion of the dominant hand of the player, which is visualised as an hand-shaped cursor onscreen. An example of a game scenario during the play is reported in Figure 3.



**Figure 3.** Game scenario during Level 4 of the GDD exergame. The user has to select the red sphere, but the instruction is written in green. The box is moving and the time is expiring (appreciable from the red background and the 1 second left to complete the task).

The fundamental gesture is hand open–closing. When the hand is opened, the subject can explore the game objects without interacting with them; on the contrary, when the hand closes in the proximity of an object, this translates into a *grab* command. The *drag* gesture consists in keeping the hand closed after a grasp, whereas the *drop* corresponds to a transition from hand closed to hand opened. In case the player grasps the wrong object, drops it outside the box or the time runs out, an error is assigned; otherwise, one point is scored. Errors and scored points are associated to a negative and a positive acoustic feedback, respectively.

The game is structured in four levels of complexity, here described in the detail.

- **Level 1 (L<sub>1</sub>):** it is designed as a tutorial level, to become acquainted with the task and the game environment. Each GDD execution can last up to 10 seconds, and background music is played at a normal pace. All objects are cubes; the discriminating factor in the selection of the correct one is only the colour, which is not repeated among the displayed objects. Six selections must be performed in sequence.
- **Level 2 (L<sub>2</sub>):** it introduces new possible shapes besides the cube. Colours and shapes can be repeated among the objects, so the correct colour-shape combination has to be identified. Ten objects have to be selected, while the background music plays with a 1.2x speed factor.
- **Level 3 (L<sub>3</sub>):** this level introduces a Stroop-Test-like [65] challenge in the recognition of the correct object. In fact, the written message onscreen can appear with a non-matching colour, e.g., the message could ask the user to select the red object while being written in yellow. The time to select each of the 10 correct objects is reduced to 6 seconds, while the music is sped up by a 1.5x factor.
- **Level 4 (L<sub>4</sub>):** this level is the most difficult one. Besides the challenges of the previous levels, the collecting box is now moving. Hence, the player has to select not only the correct object, but also the correct timing for dropping it. The time for each GDD executions is kept at 6 s as in L<sub>3</sub>, whereas the background music is played at 2x.

The transition between levels is not evident to the player, who perceives the overall game play as a whole. This design choice was made to prevent explicit transitions between levels from altering the concentration/engagement of the player. This also mimics what would happen in a real adaptive-exergaming scenario, where the exergame should re-tune itself in real-time, without explicit notifications to the player. All the information about the start and the end of the each level, along with other significant events (e.g., start/end of each GDD movement, errors and points) are stored during the game play in JSON format, making all this information available for offline processing and analysis.

The GDD exergame merges motor and cognitive rehabilitation aspects. The motor component is designed to stimulate the functionalities of the hand through the reaching, grasping and dragging gestures, which could be challenging for pathological subjects with hand dexterity impairment (e.g., PD patients or post-stroke subjects with hemiplegia). Moreover, the level L4 requires a good eye–hand coordination to follow the box movements, which diminishes in the elderly population and could be severely impaired in neurological diseases [66,67]. The cognitive aspect is related to the recognition of the correct object to select, given a set of stressors/sources of distraction that might make the decision complex (e.g., the repetitions of shapes and colours, the Stroop-Test, the countdown timer). It is worth noticing that the goal of this preliminary work was to demonstrate the possibility of measuring mental engagement and burden using simple instrumentation. Hence, the game was designed to solicit the cognitive burden as much as possible, including several simultaneous stressors in order to achieve a measurable response in the collected signals. However, in its current configuration the exergame may be too complex for pathological subjects with severe motor or cognitive impairment, and has been tested on healthy young adults only. Proper modifications and tests on the elderly and pathological subjects is left to future developments.

## 2.4. Questionnaires

After performing the GDD exergame, the participants were administered two questionnaires, in order to evaluate the task effectiveness and their self-perceived involvement. We employed the NASA-TLX [68] and the TRESCA, a novel questionnaire specifically designed to rate the users' engagement and appraise the GDD elements and features.

### 2.4.1. NASA Task Load Index

The NASA-TLX is a scale useful to retrieve information about a task, in terms of perceived workload and user's performance, and can be administered immediately after the task. It evaluates the responses on six categories, designed to explore different facets of the task, rated numerically on a 100-points scale (0: very low, 100: very high). They are [68]:

- **Mental Demand:** Extent of mental activity required by the task.
- **Physical Demand:** Extent of physical activity required by the task.
- **Temporal Demand:** Extent of time pressure felt by the subject, due to the task pace or the pace at which the task elements occurred.
- **Overall Performance:** Self-perceived success and satisfaction with the performance.
- **Effort:** Extent of mental and physical workload required to accomplish the level of performance.
- **Frustration Level:** How irritated or frustrated the subject felt during the task.

Additionally, the questionnaire includes a set of pairwise comparisons between categories. For each category pair, users are required to select the item they deem more impacting on the workload. A weighted score is then worked out for each subject. However, some studies demonstrated the efficacy of a shortened version that discards this latter part (Raw-TLX), claiming that a briefer test is beneficial for improving experimental validity [69,70]. Indeed, a shorter test is simpler and more time efficient [71], and the scores provided by the Raw-TLX highly correlated to the weighted scores of the original NASA-TLX. Based on this assumption, in this work we have implemented the Raw-TLX. The final score is the mean of the scores in all subscales, ranges between 0 and 100 and is then interpreted on a 5-point scale (Low to Very High perceived workload) [72], as reported in Table 2.

**Table 2.** Interpretation of the Raw TLX Score.

Range	Perceived Workload
0–9	Low
10–29	Medium
30–49	Somewhat High
50–79	High
80–100	Very High

### 2.4.2. TRESCA: Task-Related Scale

As previously stated, the TRESCA survey was specifically designed for this observational study. It aims at assessing the proposed task in terms of self-perceived engagement and cognitive workload, as well as evaluating the impact of the different features of the GDD exergame and the environmental stressors on the subjects' performance. It consists of seven questions, arranged over three different items (Table 3), each covering different aspects and perspectives that contribute to an effective task completion.

The first item, *Environmental and Game Features*, aims at exploring the perceived effect of stressors. The participants are required to sort, in increasing order, the exergame features and environmental factors according to their distracting effect—i.e., starting from the element that made it most difficult to carry out the task without interference. The considered elements are: ( $E_1$ ): *Repeated colour*, ( $E_2$ ): *Repeated shape*, ( $E_3$ ): *Time*, ( $E_4$ ): *Inconsistent stimuli (unmatched colour/word)*, ( $E_5$ ): *Music*, ( $E_6$ ): *Moving objects*. The subjects were also required to report whether they noticed any change in difficulty during the exergame.

The second item, *Mental and Cognitive Workload*, includes two questions that investigate the self-perceived engagement and attention levels, respectively. These two elements contribute to the validation of the exergame effectiveness in compelling and maintaining the users engaged, as described in Section 3.1. Indeed, they converge into the TRESCA Engagement Score (TENS), defined on a 10-point scale as the average of the two questions.

Finally, in the third item, *Perceived Shifts in Performance*, the participants are asked to rate on a 10-point scale their level of effort and dedication throughout the task, fatigue and concentration. These are combined into a single score, named the TRESCA Effort Score (TEFS) and defined as the average of the three self-perceived scores.

**Table 3.** TRESCA Questionnaire: items description and score range.

Item	Score
<b>I. Environmental and Game Features</b>	
Exergame Features and Environmental Factors	1–6 (1: most distracting)
Change in difficulty during the game	10-point scale (10: maximum)
<b>II. Mental and Cognitive Workload</b>	
<i>The questions refer to the parameters throughout the game</i>	
Engagement Level	10-point scale (10: maximum)
Attention level	10-point scale (10: maximum)
<b>III. Perceived Shifts in Performance</b>	
<i>The questions refer to the parameters throughout the game</i>	
Effort and Dedication	10-point scale (10: maximum)
Fatigue Level	10-point scale (10: maximum)
Active concentration throughout the game	10-point scale (10: maximum)

## 2.5. Instrumentation

This Section describes the instrumentation employed. As previously introduced, the use of low-cost, lightweight equipment is a key factor in this feasibility study. Hence, only wearable technology and 3D cameras of limited cost were employed in the experimental data collection: these are described in the following paragraphs.

### 2.5.1. Azure Kinect

The Kinect devices, initially conceived as RGB-Depth cameras for commercial gaming, have a long history of applications in the field of exergaming [73–75] and as a research tool for clinical applications. The new Azure Kinect camera presents with an improved depth sensor with respect to its predecessors [76], but still lacks a proper system for tracking hand joints in its native body tracking algorithm. Therefore, the GMH-D solution proposed in [64] was employed in this work. GMH-D merges the depth estimation, provided by the Azure Kinect depth map, and the virtual joint tracking by Google Mediapipe Hands (GMH). It provides high accuracy hand tracking, also in a scenario where GMH alone fails due to complex hand gestures [64].

The tracking works in real-time (30 frames per seconds) on a ZOTAC ©(Zotac, Fo Tan, New Territories, Hong Kong, China) ZBOX EN52060-V model. It is equipped with a 9th generation Intel ®Core™ processor (2.4 GHz quad-core), 16 GB RAM, NVIDIA GeForce RTX 2060 6GB GDDR6, HDMI and USB3 ports, Windows 10 Operating System. This makes GMH-D a suitable tracking solution for managing the real-time HCI inside the GDD game. In addition, the colour stream of the Azure Kinect camera during the game play is saved as an AVI file using the standard functions from the OpenCV library. This file is processed offline using the Google Mediapipe Face Mesh (GMFM) solution [77] to extract the eye blink features described in Section 2.7.2. Figure 4 shows the setup of the Azure Kinect and the Zotac miniPC employed in these experiments.

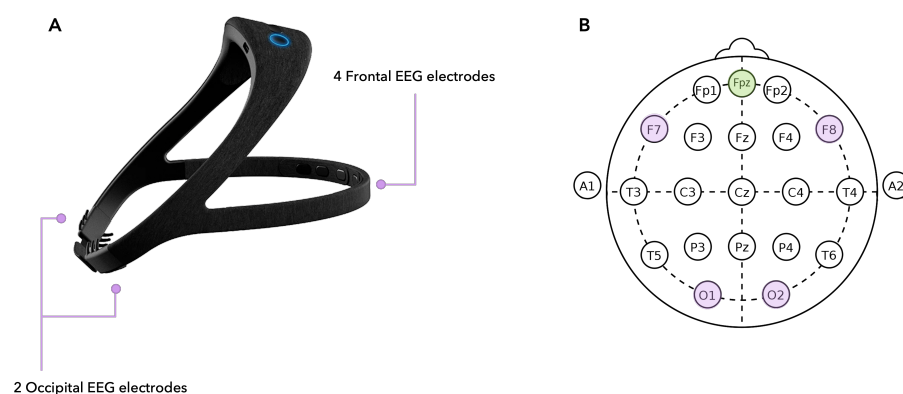


**Figure 4.** Acquisition system setup for the experimental sessions.

### 2.5.2. EEG Headset

A wireless headset (Dreem 2 Headband) was employed for the EEG signal acquisition; the subjects wore the headband throughout the execution of the experiment, from the rest stage to the end of the game play. The Dreem 2 Headband (Figure 5A) is a non-invasive, wearable device (soft fabric and TPU) specifically designed for the collection of physiological data. Generally employed for sleep tracking purposes [78], it contains several sensors, among which 6 EEG dry electrodes. Dry electrodes do not require skin preparation and the use of conductive gel. On the other hand, they necessitate close contact with the skin to ensure low contact-impedance, and therefore a good quality signal. The headband is one-size; three size adjusters (small, medium, large) are also provided to allow for proper fitting.

The EEG signal was recorded at the frontal and pre-frontal sites, following the International 10–20 System, through channels Fpz, F7, F8 with occipital channels O1 and O2 set as reference (Figure 5B). All EEG signals were acquired at a sampling frequency of 250 Hz. Data were exported as EDF files for further processing. In the perspective of developing a low-computational cost BCI based on a single-EEG channel, only the Fpz channel (Figure 5B) was employed in the analysis. This was chosen out of the other available recording sites, not only for its frontal location, but also because it featured the highest recording quality.



**Figure 5.** (A) Dreem 2 EEG Headset employed for the data collection and sensors location. (B) EEG electrodes available on the headset, 10–20 standard (green: channel selected for the study).



## 2.6. Data Pre-Processing

The collected data—i.e., raw EEG signals and video recordings for the extraction of blinking events during game play—were properly pre-processed in order to improve the quality of the extracted features. This procedure is described in detail in the following subsections.

### 2.6.1. EEG Data

The EEG signals were resampled at 256 Hz, to allow for better implementation of the frequency-domain processing. A preliminary check on the overall quality of the recordings was performed using the *channel quality metrics* provided by the Dreem acquisition system. These metrics are based on electrode and skin impedance and headband placing, and are provided on a scale 0–100, with 50 being the threshold to identify low quality recordings. Hence, subjects featuring a channel quality below 50 were discarded from subsequent analysis, leading to a final dataset including 37 subjects.

Waking EEG is a low-amplitude process, frequently affected by physiological and non-physiological artefacts of higher amplitude. In this work, artefact rejection only encompassed pre-processing filters and very simple operations. In fact, we aimed at devising a framework for real-world applications, hence we chose to avoid heavy computation. The EEG signals were bandpass-filtered through a zero-phase Chebyshev Type 1 filter (bandwidth: 0.5–40 Hz), to attenuate high-frequency noise related to EMG and the slow drifts. Artefacts related to poor headband placing, motion or electrode detachment result in very high-amplitude samples. Hence, thresholding was applied to the whole EEG recording, and samples with absolute amplitude exceeding 250  $\mu$ V were discarded. Finally, Independent Component Analysis (ICA) was employed to remove ocular artefacts (i.e., eye-blinks), using the Infomax algorithm [79].

### 2.6.2. Video Recordings for Offline Blink Detection

The subjects were recorded through the Azure Kinect camera while playing the game. The camera was placed approximately 1 m in front of the subjects for HCI and game requirements. This positioning allowed the camera to record the subjects' face during the whole game play. The video sequences collected during the test were analysed using the GMFM to extract 472 facial landmarks for each frame. To identify eye blinking events and their features, only the landmarks referring to the left and the right eyes (6 for each eye) were identified in each frame and stored in a JSON format file. The eye blinking event analysis, based on this file, is described in Section 2.7.2.

## 2.7. Feature Extraction

The features employed in the Machine Learning (ML) models were extracted from the EEG features and the eye landmarks.

The EEG feature extraction process followed three configurations, defined as follows:

- **Rest:** Features are extracted from the whole rest period preceding the game;
- **Game:** Features are extracted from the whole exergame, from start to end (with no level distinction);
- **Levels:** Features are extracted from each single level.

The features related to eye blinking were evaluated only in the Game and Levels configurations, since the subjects were asked to keep their eyes closed during Rest.

### 2.7.1. EEG Features

EEG Features were extracted from the Fpz channel in an epoch-wise fashion, and in the Time (TD) and Frequency (FD) domains respectively. The EEG records were divided into 10 s segments and TD features were extracted and concatenated as an array. Five statistics—i.e., mean, median, mode, 25th and 75th percentiles, standard deviation—were computed for each feature and employed as variables. TD features were used to describe the amplitude, morphology and statistical properties of the EEG segments. Among these, the Hjorth Parameters (Activity, Mobility, Complexity) are normalised slope descriptors

and account for the waveform variability [80]. They are computed on the signal and its 1st and 2nd order derivatives, as stated in Equations (1)–(3), where  $x$  is the EEG epoch.

$$Activity = var(x) \quad (1)$$

$$Mobility = \sqrt{\frac{\frac{dx}{dt}}{var(x)}} \quad (2)$$

$$Complexity = \frac{Mobility(\frac{dx}{dt})}{Mobility(x)} \quad (3)$$

Impulsive metrics (Form, Crest and Impact Factors) were also extracted from the EEG records, to describe the peak amplitude and the waveform properties through a measure of RMS, as proposed in [81].

On the other hand, features in the FD were computed from the Power Spectral Density (PSD), estimated using the Welch modified periodogram with sliding Hanning windows of length 1.5 s and 25% overlap. Absolute and Relative Power (AP, RP, respectively) were computed from the PSD on each waking-EEG clinically relevant band: Theta ( $\theta$ , range: 4–8 Hz), Alpha ( $\alpha$ , range: 8–13 Hz), Beta ( $\beta$ , range: 13–30 Hz) and Gamma ( $\gamma$ , range: 30–40 Hz). Additionally, AP and RP were computed on the Mu ( $\mu$ , range: 7–11 Hz) and Sensorimotor (SMR, range: 13–15 Hz) rhythms, as they have been widely employed in BCI technology [82,83].

The power in the different frequency bands was further exploited to compute parameters accounting for concentration, attention, motor planning and execution. Among these, we mention the Concentration Index (CI) [48], which is proportional to the band-power (BP) in the  $\beta$  and SMR range (Equation (4)), and the event-related desynchronisation/synchronisation (ERD/ERS) described in [48] and reported in Equation (5).

$$CI = \frac{BP_{SMR} + BP_{\beta}}{BP_{\theta}} \quad (4)$$

$$ERD/ERS_{\%} = 100 * \frac{BP_{baseline} - BP_{task}}{BP_{baseline}} \quad (5)$$

Finally, the Engagement Index (EI) was computed. It is an adimensional parameter employed in [84] to describe the engagement level and detect mental changes during the execution of a task. The index is proportional to the  $\beta$ -BP and is computed as:

$$EI = \frac{BP_{\beta}}{BP_{\theta} + BP_{\alpha}} \quad (6)$$

As an increase in the  $\beta$ -BP is commonly linked to an increased level of concentration [47,48], higher EI values may indicate higher mental engagement. A complete list of the extracted EEG features is displayed in Table 4.

**Table 4.** Employed EEG features, along with the domain and proper reference.  $\diamond$ : adapted from the cited study.

Category	Feature	Reference
Time and Morphological	Amplitude metrics: root mean square, kurtosis, maximum and minimum value	various
	Hjorth Parameters (Activity, Complexity, Mobility)	[80]
	Form, Crest and Impact Factors	various
	Approximate Entropy	[85]

Table 4. Cont.

Category	Feature	Reference
Frequency	Relative Power for each relevant frequency band ( $\theta, \alpha, \mu, SMR, \beta, \gamma$ )	various
	Absolute Power for each relevant frequency band ( $\theta, \alpha, \mu, SMR, \beta, \gamma$ )	various
	Peak Frequency in the $\alpha$ -band	[86]
	Frontal $\theta$ to $\alpha$ Ratio	[87]
	Engagement Index	◇ [47,48]
	Concentration Index	◇ [48]
	Event-Related Desynchronisation and Event-Related Synchronisation	◇ [48,53]

### 2.7.2. Eye Blinking

Eye blinking is considered a significant factor in the identification of attention and engagement [58–60]; it tends to decrease during a challenging task that requires the user to keep a steady gaze on the screen. In this work, blinking features were added to EEG ones with the aim of improving the recognition of the levels of engagement during the game. They could not be evaluated during the initial rest phase, as the subjects were asked to keep their eyes closed to better relax and avoid distractions due to the testing environment.

Several algorithms have been proposed in literature to automatically extract eye blinking from video, once the facial landmarks related to the eyes have been identified [88–90]. In this work, the algorithm based on the Eye Aspect Ratio (EAR) proposed in [91] was employed to capture blinking from the facial landmarks tracked by GMFM. The features listed in Table 5 were extracted from the identified blinking events. This list includes also trends of the EAR, since changes in the shape of the eyes such as squeezing or widening could be associated with concentration, engagement or loss of attention [58,92].

**Table 5.** Employed blink and EAR features, along with proper reference. \*: first presented/employed in this work.

Category	Feature	Reference
Blink	Blink Absolute Frequency (BAF)	*
	Blink Relative Frequency (BRF), with respect to level duration	*
	Mean Blink Duration (MBD)	[58]
	STD Blink Duration	*
	Blink Rate	various
EAR	Mean EAR	[58]
	Standard Deviation (STD) EAR	*

### 2.8. Statistical Analysis and Feature Selection

The statistical analysis was conducted using the open-source statistical tool Jamovi [93]. After data inspection through box- and violin plots, the normality of the employed EEG and eye blinking features was investigated by means of the Shapiro-Wilk test. Then, the parametric independent sample Student's t-test (for normal features) and the non-parametric Mann–Whitney independent sample U-Test (for non-normal features) were used to identify characteristics differently distributed between the rest stage (REST) and the game play stage (GAME) of the protocol. All statistical tests were performed at the 95% confidence interval.

Feature Selection was performed by means of the the ReliefF algorithm [94] to select the top- $K$  relevant features for the classification task. The parameter  $K$  was chosen by identifying the elbow on the feature importance scores computed by the algorithm. For classification, only the features above threshold were considered when applying the ML models.

### 2.9. Automatic Classification of Mental Activation

To prove the feasibility of an EEG-based assessment of mental activation during exergames, supervised ML methods were employed to automatically discriminate among different mental states. In particular, as introduced in Section 2.7, the following two configurations were explored:

- **(1) Rest vs. Game:** binary classification between the rest stage and the task (whole exergame, levels 1 to 4);
- **(2) Inter-Level Classification:** 4-stage classification intended to detect the four levels designed in the exergame.

In the first classification task, only EEG features were included, as during the REST stage the subjects kept their eyes closed; hence, it was not possible to retrieve the blinking pattern. The eye-blinking features were introduced in the second classification task (discrimination among levels).

Four supervised models were explored, namely:

- **Support Vector Machine (SVM):** the model aims at finding the hyperplane that best separates the samples in the dataset according to their class.
- **K-Nearest Neighbour (KNN):** it is a non-parametric method that employs similarity measures to classify the elements in the dataset.
- **Discriminant Analysis (DA):** it is a multi-variate data analysis technique that aims at finding the linear combination of features that best characterises the classes in the dataset.
- **AdaBoost:** it is an adaptive ensemble learning method based on Random Forest classifiers. In the binary classification configuration, it trains learners sequentially and, at each iteration, it updates the observation weights by increasing the weights of the misclassified ones and decreasing those of correctly classified.

A  $k$ -fold Cross Validation (CV) ( $k = 10$ ) was employed in the training phase to allow for better generalisation capability and mitigate the risk of overfitting. This technique partitions the available data into  $k$  different subsets. Then, the models are trained on  $k - 1$  subsets and tested on the remaining one. This procedure is repeated until all folds have been explored. The obtained performance metrics are the average of those yielded in the different iterations. Hyperparameters were optimised through a Grid Search approach or Bayesian optimisation (maximum number of iterations: 50) to further ensure the robustness of the trained classifiers. A complete summary of the employed models and optimised parameters is provided in Table 6.

**Table 6.** Summary of the explored classifiers, along with the hyperparameters search range and the optimised parameters. For each classifier, the chosen normalisation and cross-validation technique are also displayed.

	SVM	KNN	DA	AdaBoost
<b>Normalisation</b>	z-score	z-score	z-score	z-score
<b>Optimisation Method (Iterations)</b>	Bayesian (50)	Grid Search (50)	Grid Search (50)	Bayesian (50)
<b>Hyperparameters and Search Range</b>	Kernel function: linear, quadratic, cubic, gaussian Penalty (C): 0.001–1000	Distance metric: Euclidean, Chebyshev, Hamming, Manhattan, Mahalanobis, Spearman, Minkowski K: 1–12	Discriminant: linear, quadratic	Splits: range 1–50 Learners: 1–100 Learning rate: 0.1–1
<b>Optimised Parameters</b>	Linear kernel C = 1	Euclidean Distance K = 1	Linear Discriminant	Splits: 20 Learners: 30 Learning rate: 0.1
<b>Cross-Validation</b>	10-fold	10-fold	10-fold	10-fold

### 2.10. Subjects' Response to Levels Variation

Several works proposed in the literature exploit the EI as an objective metric to evaluate engagement. Moreover, in this work, the EI was found relevant for the classification REST vs. GAME (cf. Section 3). As discussed in Section 2.7.1, the EI is adimensional and its range may vary across different subjects. Consequently, if one wants to explore how different subjects cluster with respect to this parameter, this process cannot be based on the actual EI values, and an alternative metric, accounting for its increasing/decreasing trend, must be employed. We performed a preliminary analysis of the variation of the EI among the four game levels and for each subject. In more detail, for each subject  $s$ , a vector  $I_s \in R^4$  was computed such that each element  $i_L$ , with  $L = 1, \dots, 4$  and  $L - 1 = 0$  representing the REST stage, is defined as:

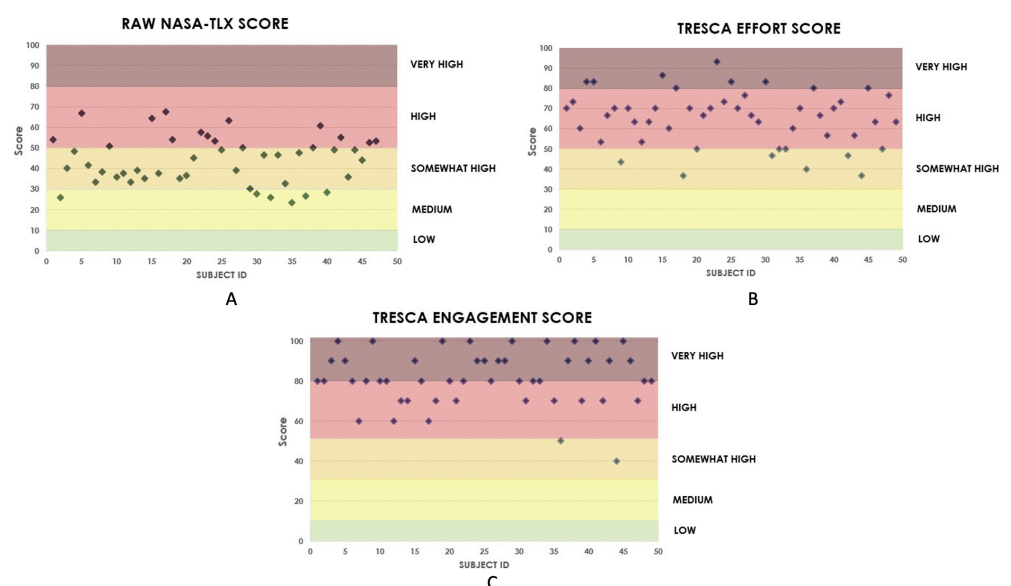
$$i_L = \begin{cases} -1 & EI_L - EI_{L-1} < 0 \\ 1 & EI_L - EI_{L-1} \geq 0 \end{cases} \quad (7)$$

This binary vector describes an EI increase (+1) or decrease (−1) between levels  $L - 1$  and  $L$ , allowing to cluster the subjects' responses according to the overall trend in the four levels. The correspondence between the dominant clusters and the answers provided in the TRESCA questionnaire was explored to identify relevant correlation.

## 3. Results

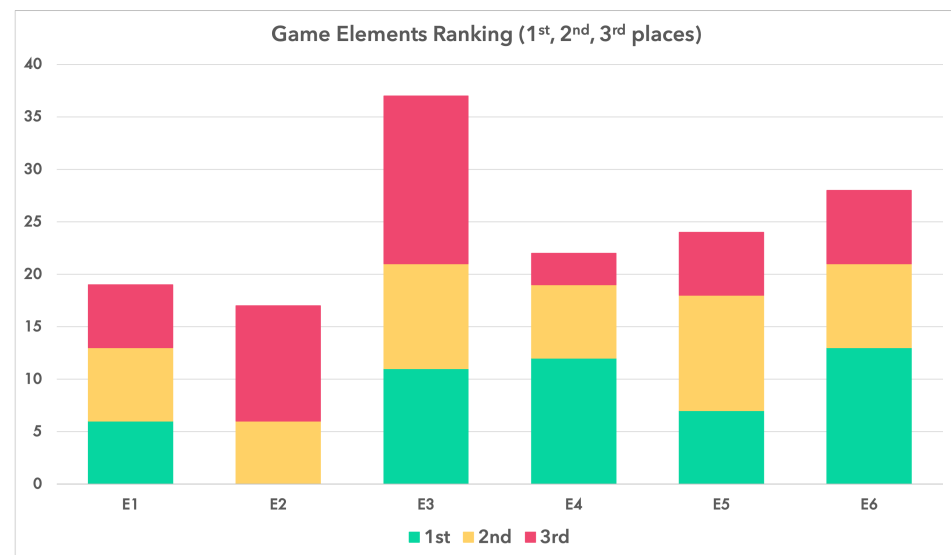
### 3.1. Exergame Validation

Figure 6A displays the Raw NASA TLX scores for all the tested subjects (also the subjects excluded from the data analysis are included here, since they actually performed the game, and hence could provide a feedback through self-assessment). The score ranges highlighted in the plot are as in [72]. As it can be appreciated, all scores fall in the MEDIUM to HIGH range. The TEFS and the TENS scores (Figure 6B,C) both report values in the SOMEWHAT-HIGH to VERY-HIGH range. This shift upwards is likely related to the fact that these scores are more specific to the goals of this study (effort and engagement) than the NASA TLX. This latter, even if largely validated, still remains a general-purpose questionnaire. Overall, the evaluation provided by the users support the engaging capability of the designed exergame.



**Figure 6.** Exergame validation through administered questionnaires. (A) NASA-TLX raw score per each subject; (B) TRESCA Effort score per each subject; (C) TRESCA Engagement score per each subject.

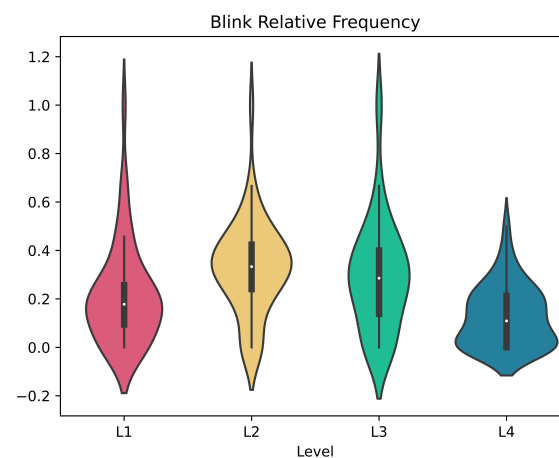
In Figure 7, for each Environmental and Game Feature considered in Item 1 of the TRESCA questionnaire, the number of subjects that have ranked that element in the first three positions is reported. For the sake of clarity, we recall the element definitions: ( $E_1$ ): Repeated colour; ( $E_2$ ): Repeated shape; ( $E_3$ ): Time; ( $E_4$ ): Inconsistent stimuli (unmatched colour/word); ( $E_5$ ): Music; ( $E_6$ ): Moving objects. From this Figure, it can be appreciated that  $E_3$  was the most frequently ranked element among the first three positions, followed by  $E_6$  and  $E_5$ . As for the first ranking position,  $E_6$ ,  $E_4$  and  $E_3$  were considered as the most challenging elements by more than 72% of the players.



**Figure 7.** Distribution of the first three positions in the ranking of the most distracting game elements.

### 3.2. Feature Analysis: Statistical Analysis and Feature Selection Results

Regarding the characteristics of blinking features, Figure 8 reports the violin plots of the Blink Relative Frequency (BRF), defined as the number of blinks divided by the duration of each level. This parameter exhibits a bell-shaped behaviour across the levels. The median values are lower in  $L_1$  and  $L_4$ , and they raise in  $L_2$  and  $L_3$ . This seem to suggest similar visual attention patterns in the player at the beginning and the end of the game, when the maximum stimulation is enforced to the user.

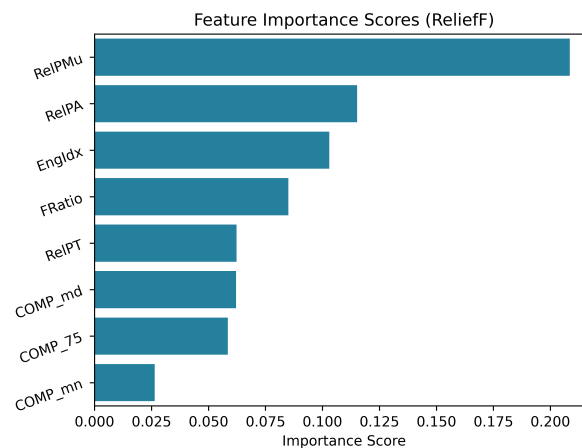


**Figure 8.** Violin plot of the Blink Relative Frequency across the four levels.

Regarding feature selection, after inspecting the importance scores provided by the ReliefF algorithm,  $K = 8$  (out of 79) was chosen as the number of EEG features to be kept in the REST vs. GAME classification task. Figure 9 shows the importance scores, whereas



Table 7 reports their statistical description. In the second classification task, the eye-blinking features were included. In this case, a total of 15 features (out of 85) were selected through the ReliefF scores. The statistics of the selected eye-blink features are reported in Table 7. It can be noticed that most EEG features are normally distributed according to the Shapiro-Wilk test results, and all the EEG features are statistically significant. On the other hand, the eye-blinking features are not normally distributed in the levels.



**Figure 9.** Importance scores of the selected features, computed through the ReliefF algorithm.

**Table 7.** Descriptive and Independent Samples Statistics of the features employed in the classification. The  $\diamond$  mark represents normal features, the \* represents significant features (95% confidence interval).

Features (EEG)	Shapiro-Wilk	Independent Sample Test
Relative Power ( $\mu$ )	0.057 $\diamond$	<0.001 *
Relative Power ( $\alpha$ )	0.142 $\diamond$	<0.001 *
Engagement Index	<0.001	<0.001 *
Frontal Ratio	0.088 $\diamond$	<0.001 *
Relative Power ( $\theta$ )	0.841 $\diamond$	<0.001 *
Complexity <sub>mode</sub>	0.245 $\diamond$	<0.001 *
Complexity <sub>75pctl</sub>	0.036 $\diamond$	<0.001 *
Complexity <sub>mean</sub>	0.100 $\diamond$	<0.001 *
Features (Eye-Blink)	Shapiro-Wilk	Independent Sample Test
Blink Relative Frequency	<0.001	<0.005
Eye Aspect Ratio <sub>STD</sub>	<0.001	<0.005

### 3.3. Automatic Classification Results

As discussed in Section 2.9, a binary classification between the REST and the GAME configurations was performed. The results are reported in Table 8, with the positive class being GAME. The performance of each classifier is reported through the Accuracy, Sensitivity, Specificity, Precision (PPV) and F-1 score—the latter being the harmonic mean of Sensitivity and Precision. The metrics are computed according to Equations (8)–(12)—where TP, TN, FP, FN stand for True Positives, True Negatives, False Positives, False Negatives, respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$Specificity = \frac{TN}{TN + FP} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$F_1 = 2 * \frac{Precision * Sensitivity}{Precision + Sensitivity} \quad (12)$$

As can be appreciated, all the explored models scored a high overall accuracy (in excess of 90%), suggesting good classification performance and high predictive power of the features selected through the ReliefF algorithm. According to the computed metrics, the Linear DA performed the best, yielding a 98.61% accuracy, with precision and F-1 values of 97.3% and 98.6%, respectively.

**Table 8.** Classification performance of the employed classifiers, REST vs. GAME.

	SVM	KNN	DA	AdaBoost
Accuracy	95.83%	93.05%	98.61%	90.27%
Sensitivity	97.23%	94.62%	97.21%	91.67%
Specificity	94.45%	91.41%	98.35%	88.89%
Precision	94.59%	91.89%	97.29%	89.19%
F-1	95.89%	93.15%	98.63%	90.41%

The models employed in binary classification REST vs. GAME were then exploited to attempt a 4-stage classification, employing the features extracted from each level of the exergame as well as the eye-blinking features (cf. Table 5). However, even after a thorough feature selection process, it was not possible to attain satisfactory as those presented in Table 8. Indeed, though being considerably unaffected by false positives (with the best configuration featuring a macro-averaged specificity over 98%), the implemented models could not effectively discriminate the four game levels.

In order to understand the reason behind this performance impairment, the features reported in literature as those most linked to engagement and attention were further investigated. In particular, as previously discussed, BRF is a key element in evaluating mental states, since lower values of the feature are generally linked to a higher level of concentration. As appreciable from Figure 8, both  $L_1$  and  $L_4$  present with small values of BRF, suggesting a similar pattern in the initial and final stages of the exergame, as well as a high level of engagement. On the other hand, the violin plots display an increase of BRF in  $L_2$  and  $L_3$ ; the two levels also present with an alike data distribution. This may indicate a reduction in engagement with respect to the previous level, probably due to an adaptive mechanism. Based on this assumption, it is possible to infer that not only  $L_2$  and  $L_3$  follow a similar trend, but also they may represent the steady state of the game. In  $L_1$ , the low BRF may be due to the fact that the subject has to become used to the game, whereas the low BRF values observed in  $L_4$  may denote a significant change in difficulty when shifting from  $L_3$ .

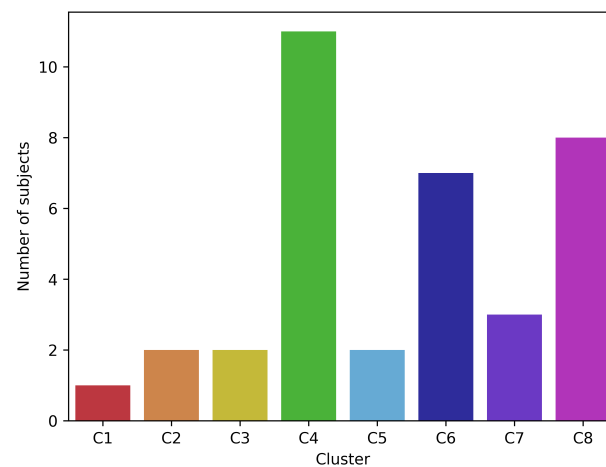
For this reason, a third classification task was performed to discriminate the Low-Engagement (LE) and High-Engagement (HE) stages of the game, which, based on the feature analysis, were identified in  $L_2$  and  $L_4$ , respectively. Various supervised models were trained and tested for this purpose. For the sake of brevity, Table 9B reports the classification performance of the best model (AdaBoost). A 10-fold CV was applied to allow for better generalisation capability. The model hyperparameters were optimised through Bayesian Optimisation (50 iterations); their search range and the optimised model parameters are displayed in Table 9A. The model yielded an overall accuracy of 72.2%, with Sensitivity and F-1 score of 75% and 72.3%, respectively. These results denote a reasonable capability to discriminate between a steady, low-engagement phase of the game and a high-engagement one.

**Table 9.** AdaBoost classifier (Low-Engagement vs. High-Engagement): model hyperparameters and classification performance.

A. Model Hyperparameters					
	Splits	Learners	Learning rate		
Search Range	1–50	1–100	0.1–1		
Optimised Parameters	20	30	0.1		
B. Classification Performance					
	Accuracy	Sensitivity	Specificity	Precision	F-1
Optimised AdaBoost	72.2%	75%	69.44%	71.05%	72.3%

### 3.4. Subjects Response to Levels

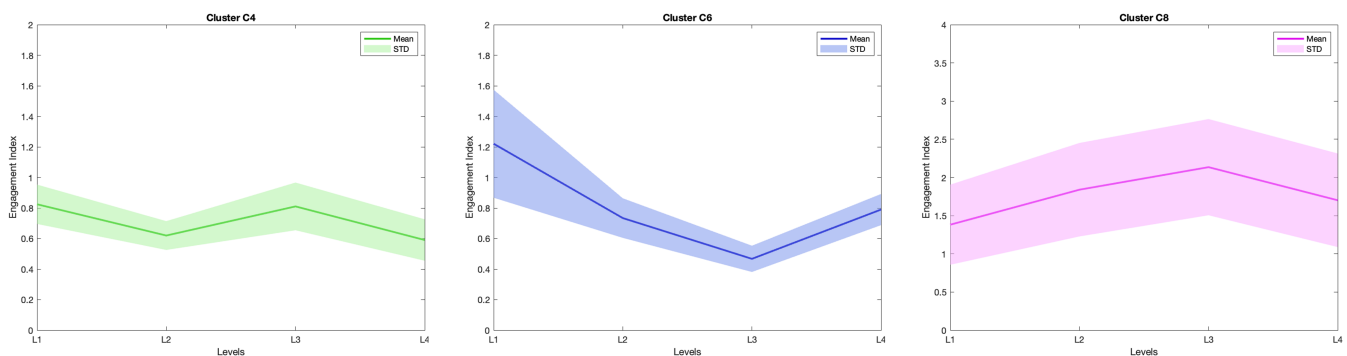
As introduced in Section 2.10, the EI values across the four exergame levels  $L_i$ ,  $i = 1, \dots, 4$ , can evaluate the variation of the subjects' engagement. The inter-level variations  $i_L$ , defined in Equation (7), were employed to identify different clusters reflecting the response of the subjects to the game levels. Figure 10 displays the results. Out of the possible  $2^{L-1}$  clusters, C4, C6, and C8 emerged as the most populated, with 11, 8 and 7 subjects, respectively, while the other clusters only encompassed a limited number of subjects (up to 3). Due to the limited population size, it was not possible to infer whether these latter represent actual trend clusters or outliers. Therefore, they were excluded from the subsequent analysis.

**Figure 10.** Inter-level variation clusters, along with the number of included subjects

A deeper analysis of the most populated clusters is reported below; for the sake of clarity, an inter-level EI decrease (increase) is represented as D (I), respectively.

- **Cluster C4.** This resulted to be the most populated cluster, including 11 subjects. The trend follows a  $D - I - D$  pattern, meaning that the subjects showed an increase in the value of EI while shifting from  $L_2$  to  $L_3$ .
- **Cluster C6.** This cluster includes 7 subjects and follows a  $D - D - I$  pattern. This reflects an increase of the EI value in  $L_4$ .
- **Cluster C8.** This cluster encompasses 8 subjects. The inter-level variation pattern is  $I - I - D$ . This denotes an increase in the EI value when shifting from  $L_1$  to  $L_2$  and from  $L_2$  to  $L_3$ , and then a decrease when moving to  $L_4$ .

Figure 11 displays the EI values (mean  $\pm$  STD) across the four levels, in the three most populated clusters C4, C6, and C8. As discussed in Section 2.7.1, the EI is adimensional and its range may vary across different subjects. Consequently, the clustering process is based on its increasing/decreasing trend.



**Figure 11.** Engagement Index (EI) across the four levels, in the three most populated clusters.

After evaluating the inter-level EI variations in the most populated clusters, the subjects' response to the TRESCA questionnaire (Item 1, Question 1) was analysed, in order to identify pertinent connections between the designed stressors and the change in engagement. As regards C6, which displays an EI increase in L<sub>4</sub>, the totality of the subjects indicated the *Moving Objects* and *Fast-Paced Music* as being the most distracting features—i.e., both stressors that characterise L<sub>4</sub>. On the other hand, C8 shows a linear increase of EI from L<sub>2</sub> to L<sub>3</sub>, and a drop in L<sub>4</sub>. From the TRESCA Questionnaire, 63% of the subjects in this cluster marked the *Repeated Colour/Shape* and *Incongruent Stimuli* (Stroop-Test) in the top-3 distracting stressors. Finally, cluster C4 shows a varying trend, with a decrease of EI in L<sub>2</sub>, a rise in L<sub>3</sub> and a subsequent drop. For this cluster, it was not possible to find a direct correspondence with the subjects' responses; however, 55% of the subjects indicated the *Incongruent Stimuli* among the top-3 most distracting elements.

#### 4. Discussion

In the scenario of neurodegenerative disorders, motor/cognitive rehabilitation plays a fundamental role in the slowdown of the degeneration process and in symptoms mitigation, with demonstrated improvements in quality of life. However, in-person and in-hospital rehabilitation is difficult to carry out on a daily basis. The solution could rely on the implementation of new generation rehabilitation strategies such as exergames, in which the subjects are involved in the accomplishment of a video game whose goals are coupled with specific cognitive and motor rehabilitation strategies. Exergames could apply also to telerehabilitation scenarios, allowing for the implementation of continuous, at-home protocols.

However, two main issues may arise in this approach. On the one hand, rehabilitation protocols should be properly tailored on the need of the single subject; on the other hand, the effectiveness of the solution is inevitably linked to continuity of use, which can be achieved only if a good level of engagement is elicited in the player. EEG signals and eye-blinking related patterns are a promising approach for the measurement of engagement in game play. A low-cost BCI based on coupling video analysis and a non-invasive EEG headsets could be used to detect engagement while playing and automatically adjust the game so as to increase or decrease the solicitation provided to the user.

In this work, the above-described BCI system, based on Dreem Headband v2 and Azure Kinect for HCI interaction and eye-blinking detection, has been preliminary investigated. Fifty healthy subjects were recruited to test the GDD exergame, specifically designed to stimulate motor and cognitive performance through a sequence of four levels of increasing difficulty. Subjects were equipped with the EEG headband and played the game while their EEG signals from the frontal and pre-frontal cortices were recorded. The user interaction with the game was made possible through the Azure Kinect device, which recorded also the face of the user during the game play. From these recordings, the blinking events, that could convey additional information on the engagement of the player, were extracted and included in the analysis.

The effectiveness of the GDD exergame in stimulating the subjects was proven by the results of the questionnaires administered after the trials. From the Raw NASA-TLX, 88% of subjects found the game as “SOME-WHAT HIGH to HIGH”, in terms of workload required. From the game-specific TRESCA questionnaire, effort and engagement scores were found to be in the HIGH and VERY-HIGH ranges for 80% and 96% of the players, respectively. Moreover, among the stressors included in the game to solicit the player, Time, Inconsistent Stimuli (i.e., the Stroop-Test), Moving Objects and Fast-paced Music were all considered as significant stressors. All these elements appear or increase their magnitude during the transition between levels; therefore, the answers from the players seem to suggest that an increase in the required mental workload was correctly achieved following the initial design choices. As a limitation, the absence of an explicit indication of the game levels did not allow to inquire more specifically about which of the four sequences was actually the most stimulating one. Future work should take into account whether to convey this information, to extrapolate from the user more specific feedback at the end of the task.

From the EEG signals and facial video recordings a series of features related to mental activation, engagement and blinking events were extracted offline, after proper data pre-processing. The statistical analysis of such features showed a mix of normally and non-normally distributed features.

As for the EEG features, the employed Independent Sample Statistical Tests (i.e., Student's T or Mann–Whitney U) highlighted how the features *Relative Powers* ( $\alpha, \theta, \mu$ ), *Complexity*, *Engagement Index*, and *Frontal Ratio* exhibit different statistical distributions during the rest and the game play stages. This is reflected in the high performance achieved through the classification methods employed in the REST vs. GAME task, in terms of all the computed performance metrics. This result supports the idea that a discrimination between these two states may be properly measured by means of a low-cost BCI such as that proposed in this work. Indeed, all the tested classifiers achieved an overall accuracy over 90%, and the best model attained an accuracy, sensitivity and F-1 score of 98.61%, 97.2% and 98.63%, respectively.

An extensive comparison with previous works is not trivial, for several reasons. First, to the best of our knowledge, this is the first study that aims at evaluating engagement during a motor-cognitive task presented by means of an exergame, providing multi-source stimulation. Second, the framework proposed in this work relies on ML algorithms, whereas the vast majority of works investigating engagement only performed statistical analyses on the collected data [57,95]. Last, previous works mainly commit to multi-channel EEG recorded with traditional EEG caps—therefore being invasive and requiring higher setup times [56].

In [96], a configuration similar to the one proposed in this work is presented; the Authors employ ML models in a Human-Computer Interaction scenario. EEG signals are collected from the participants both during rest and while playing a commercial videogame. Our results outperform those in [96], where the Rest vs. Game configuration attained a 92.7% accuracy through a Bayesian Network classifier. This is the only reported parameter in the mentioned study, therefore a thorough comparison with the other performance metrics was not possible, also due to the different nature of the employed classification model. Furthermore, in [96] the Authors exploit two EEG channels for the analysis, and, though being reasonably low-cost instrumentation, the Ag/Cl electrodes employed still required full-skin preparation to reduce skin impedance and the use of conductive paste to allow for good electrical connection, in contrast to the dry-electrodes employed in our study. In [56], the Authors collect numerous EEG channels and employ a SVM classifier to detect cognitive-only tasks based on EEG parameters, among which the EI. The only performance metric available is the accuracy. The best reported model yielded a 93.33% accuracy, averaged across the subjects; our results for the SVM classifier are in line with these findings. However, though the Authors claim to have successfully performed a multi-task recognition, it is not clear how the reported accuracy for each subject was

computed. Finally, the study only involved six subjects, and the EEG was recorded through a traditional electrode-cap.

As regards the blinking features, BRF exhibits a particular trend (Figure 8): blinking is reduced in the first and the last levels, whereas it tends to increase in levels two and three. This behaviour may find its explanation in the following: at the beginning of the task, the player does not know what to expect, and blinking is suppressed to keep the visual focus steady on the game scene. Once the player becomes used to the game procedures in L<sub>2</sub> and L<sub>3</sub>, the blinking frequency increases again. The sudden change in the game in L<sub>4</sub> (i.e., the box starting to move), which involves a new significant visual stimulation, produces again a reduction in the occurrence of blinking. This result seems to support the importance of eye blinking in the evaluation of attention/engagement and will be further explored in future works.

Nevertheless, including the eye-blinking features did not prove sufficient to improve the Inter-Level classification, that yielded overall poor results, even though in line with [97]. This could be explained by several factors. On the one hand, the different levels combine several stimuli, which could reflect in different types of responses in the EEG signal. This superimposition of effects could be challenging to disentangle and therefore to associate to a specific level response. In addition to this, only the EEG recordings from a single channel were used to extract the features, inevitably reducing the overall quality of the computed metrics. On the other hand, the answers to the questionnaires show how the different subjects had different responses to the provided stimulation, which, for instance, indicates that not for all participants L<sub>4</sub> was implicitly more challenging and engaging than L<sub>3</sub>. Finally, each level had a different duration, spanning from a minimum of 10 s to a maximum of 40 s on average, according to the player's performance. It is possible that the short duration of some levels did not allow for the recognition of level-specific characteristics in the EEG signals, whereas in the classification REST vs. GAME, the parameters were obtained over longer recordings (around 2–3 min for each phase). For these reasons, a classification between L<sub>2</sub> and L<sub>4</sub> was performed considering the former as a low-engagement level after the initial training stage L<sub>1</sub> (as proven by the average increase in the BRF parameter), and the latter as the most stimulating one, due to all the stressors that it simultaneously encompasses. The best model achieved an overall accuracy of 72.2%, with sensitivity and F-1 score of 75% and 72.3%, respectively. This provides support for the fact that a finer classification might be feasible, and its investigation is left to future work.

Finally, it may be of interest to reduce the stimulation or adapt it according to the tested subject, in order to better identify clear variations in the EEG patterns. To this end, the results of the TRESCA questionnaire and the results of the clustering could provide insight on the response pattern of the subjects. Indeed, the exploration of the engagement trends in the subjects aimed not only at validating the proposed stressors, but also at assessing the feasibility of an adaptive system that may provide the user with the proper stimuli for a successful rehabilitation strategy. For example, cluster C6 showed a strong connection between (1) the elements marked as most distracting and (2) an increase in engagement in the corresponding level. However, as the transition between levels was designed to be as smooth as possible—i.e., without informing the user of any level shift or change in difficulty—a clear distinction between levels is challenging, also due to the fact that the stressors are subsequently added as exergame features, or increased in intensity throughout the game. This may explain why such a definite correspondence *Level–Stressor* was difficult to attain in clusters C4 and C8, which, however, showed encouraging connections.

This work is not without limitations. First, being this a pilot study, only healthy subjects were recruited; therefore engagement patterns and neuro-physiological characteristics were explored only in that sample. When including subjects with neurodegenerative diseases, proper calibrations and adjustments may be required in the offline data processing stage. For instance, one should take into account the difference in EEG morphology that could occur in that population, as well as disease-specific or abnormal eye-blinking patterns. Second, dry-EEG electrodes allow for a faster and easier setup procedure but, at



the same time, the lack of skin preparation and conductive paste may result in higher skin impedance and lower recording quality. In fact, 13 subjects had to be discarded from the study as the recording quality did not allow for a reliable analysis. Last, in this work only the Fpz channel was explored. Given that the proposed task is mainly a visual-stimulated one, it could be beneficial to evaluate the EEG activity at the occipital channels or the central channels—the latter, however, were not available in the employed EEG headset. A single-channel-based analysis may limit performance when trying to detect finer variations in engagement. However, in view of a telemedicine application, employing a single electrode would undeniably allow for a less costly system, and remove the need for cumbersome and invasive instrumentation.

## 5. Conclusions

This pilot study investigated the feasibility of a low-cost BCI system based on single-channel EEG for the assessment of mental workload and engagement. In the future, this would allow for the development of adaptive telerehabilitation strategies without the need for expensive and invasive equipment. The results, from a clinical point of view, are promising; indeed, it was possible to classify an *active* mental state—with respect to *rest*—and to identify different engagement paradigms during game play. Future work will focus on the discrimination of finer stimuli from EEG and facial features; to this end, each game level could be associated to a specific predominant stressor, therefore avoiding the influence of other factors and superimposition effects. Finally, the investigation of engagement patterns, mental and emotional workload will be replicated and validated on pathological subjects.

**Author Contributions:** Conceptualization, G.A. and I.R.; methodology, G.A., I.R., C.F. and G.O.; software, G.A. and I.R.; validation, G.A., I.R., C.F. and G.O.; formal analysis, G.A., I.R. and G.O.; resources, C.F.; data curation, G.A. and I.R.; writing—original draft preparation, G.A. and I.R.; writing—review and editing, C.F. and G.O.; supervision, C.F. and G.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of A.O.U. Città della Salute e della Scienza di Torino (approval No. 00384/2020).

**Informed Consent Statement:** Written informed consent was obtained from all study participants.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

**Acknowledgments:** The Authors would like to thank the volunteers who took part in the study. G.A. and I.R. would like to thank T. Bergling for his positive influence, words of wisdom and inspiration.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript, and are presented in alphabetical order:

AP	Absolute Power
AVI	Audio Video Interleave
BCI	Brain-Computer Interface
BP	Bandpower
BRF	Blink Relative Frequency
CEFR	Common European Framework of Reference
CI	Concentration Index
CV	Cross-Validation
DA	Discriminant Analysis

DAP	Data Acquisition Protocol
EAR	Eye Aspect Ratio
EEG	Electroencephalography
EI	Engagement Index
ERD	Event-Related Desynchronisation
ERS	Event-Related Synchronisation
FD	Frequency Domain
FN	False Negatives
FP	False Positives
GDD	Grab-Drag-Drop
GMFM	Google Mediapipe Face Mesh
HCI	Human Computer Interaction
HE	High-Engagement
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbour
LE	Low-Engagement
NASA-TLX	NASA Task Load Index
NeAdEx	Neuroadaptive Exergame
PD	Parkinson's Disease
PPV	Precision
QoL	Quality of Life
R&D	Research and Development
RP	Relative Power
SMR	Sensorimotor Rhythm
SVM	Support Vector Machine
TD	Time Domain
TEFS	TRESCA Effort Score
TENS	TRESCA Engagement Score
TN	True Negatives
TP	True Positives
TRESCA	Task-Related Scale

## References

1. Peretti, A.; Amenta, F.; Tayebati, S.K.; Nittari, G.; Mahdi, S.S. Telerehabilitation: Review of the state-of-the-art and areas of application. *JMIR Rehabil. Assist. Technol.* **2017**, *4*, e7511. [[CrossRef](#)] [[PubMed](#)]
2. Simon, D.K.; Tanner, C.M.; Brundin, P. Parkinson disease epidemiology, pathology, genetics, and pathophysiology. *Clin. Geriatr. Med.* **2020**, *36*, 1–12. [[CrossRef](#)]
3. Alzheimer's Association. Alzheimer's disease facts and figures. *Alzheimer's Dement.* **2018**, *14*, 367–429.
4. Sveinbjornsdottir, S. The clinical symptoms of Parkinson's disease. *J. Neurochem.* **2016**, *139*, 318–324. [[CrossRef](#)]
5. Arvanitakis, Z.; Shah, R.C.; Bennett, D.A. Diagnosis and management of dementia: Review. *JAMA* **2019**, *322*, 1589–1599. [[CrossRef](#)] [[PubMed](#)]
6. Bature, F.; Guinn, B.A.; Pang, D.; Pappas, Y. Signs and symptoms preceding the diagnosis of Alzheimer's disease: A systematic scoping review of literature from 1937 to 2016. *BMJ Open* **2017**, *7*, e015746. [[CrossRef](#)]
7. Maresova, P.; Hruska, J.; Klimova, B.; Barakovic, S.; Krejcar, O. Activities of daily living and associated costs in the most widespread neurodegenerative diseases: A systematic review. *Clin. Interv. Aging* **2020**, *15*, 1841–1862. [[CrossRef](#)] [[PubMed](#)]
8. World Health Organization. *Neurological Disorders: Public Health Challenges*; World Health Organization: Geneva, Switzerland, 2006.
9. Debù, B.; De Oliveira Godeiro, C.; Lino, J.C.; Moro, E. Managing gait, balance, and posture in Parkinson's disease. *Curr. Neurol. Neurosci. Rep.* **2018**, *18*, 23. [[CrossRef](#)] [[PubMed](#)]
10. Mok, V.C.T.; Pendlebury, S.; Wong, A.; Alladi, S.; Au, L.; Bath, P.M.; Biessels, G.J.; Chen, C.; Cordonnier, C.; Dichgans, M.; et al. Tackling challenges in care of Alzheimer's disease and other dementias amid the COVID-19 pandemic, now and in the future. *Alzheimer's Dement.* **2020**, *16*, 1571–1581. [[CrossRef](#)]
11. Cummings, J. Correction to: New approaches to symptomatic treatments for Alzheimer's disease. *Mol. Neurodegener.* **2021**, *16*, 21. [[CrossRef](#)]
12. Armstrong, M.J.; Okun, M.S. Diagnosis and treatment of Parkinson disease: A Review. *JAMA* **2020**, *323*, 548–560. [[CrossRef](#)] [[PubMed](#)]
13. Rafferty, M.R.; Nettnin, E.; Goldman, J.G.; Macdonald, J. Frameworks for Parkinson's Disease Rehabilitation Addressing When, What, and How. *Curr. Neurol. Neurosci. Rep.* **2021**, *21*, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Gupta, A.; Prakash, N.B.; Sannyasi, G. Rehabilitation in dementia. *Indian J. Psychol. Med.* **2021**, *43*, S37–S47. [[CrossRef](#)] [[PubMed](#)]

15. Formisano, R.; Pratesi, L.; Modarelli, F.T.; Bonifati, V.; Meco, G. Rehabilitation and Parkinson's disease. *Scand. J. Rehabil. Med.* **1992**, *24*, 157–160. [[PubMed](#)]
16. Fayyaz, M.; Jaffery, S.S.; Anwer, F.; Zil-E-Ali, A.; Anjum, I. The effect of physical activity in Parkinson's disease: A mini-review. *Cureus* **2018**, *10*, e2995. [[CrossRef](#)] [[PubMed](#)]
17. Chromiec, P.A.; Urbaś, Z.K.; Jacko, M.; Kaczor, J.J. The proper diet and regular physical activity slow down the development of Parkinson disease. *Aging Dis.* **2021**, *12*, 1605–1623. [[CrossRef](#)]
18. Cusso, M.E.; Donald, K.J.; Khoo, T.K. The impact of physical activity on non-motor symptoms in Parkinson's disease: A systematic review. *Front. Med.* **2016**, *3*, 35. [[CrossRef](#)]
19. Speelman, A.D.; van de Warrenburg, B.P.; van Nimwegen, M.; Petzinger, G.M.; Munneke, M.; Bloem, B.R. How might physical activity benefit patients with Parkinson disease? *Nat. Rev. Neurol.* **2011**, *7*, 528–534. [[CrossRef](#)]
20. Abbruzzese, G.; Marchese, R.; Avanzino, L.; Pelosin, E. Rehabilitation for Parkinson's disease: Current outlook and future challenges. *Park. Relat. Disord.* **2016**, *22* (Suppl. S1), S60–S64. [[CrossRef](#)]
21. Mak, M.K.; Wong-Yu, I.S.; Shen, X.; Chung, C.L. Long-term effects of exercise and physical therapy in people with Parkinson disease. *Nat. Rev. Neurol.* **2017**, *13*, 689–703. [[CrossRef](#)]
22. Yau, S.Y.; Gil-Mohapel, J.; Christie, B.R.; So, K.F. Physical exercise-induced adult neurogenesis: A good strategy to prevent cognitive decline in neurodegenerative diseases? *Biomed Res. Int.* **2014**, *2014*, 403120. [[CrossRef](#)]
23. Marques-Aleixo, I.; Belez, J.; Sampaio, A.; Stevanović, J.; Coxito, P.; Gonçalves, I.; Ascens ao, A.; Magalhães, J. Preventive and therapeutic potential of physical exercise in neurodegenerative diseases. *Antioxid. Redox Signal.* **2021**, *34*, 674–693. [[CrossRef](#)]
24. Vecchio, L.M.; Meng, Y.; Xhima, K.; Lipsman, N.; Hamani, C.; Aubert, I. The neuroprotective effects of exercise: Maintaining a healthy brain throughout aging. *Brain Plast.* **2018**, *4*, 17–52. [[CrossRef](#)] [[PubMed](#)]
25. Docu Axelerad, A.; Stroe, A.Z.; Muja, L.F.; Docu Axelerad, S.; Chita, D.S.; Frecus, C.E.; Mihai, C.M. Benefits of Tango Therapy in Alleviating the Motor and Non-Motor Symptoms of Parkinson's Disease Patients—A Narrative Review. *Brain Sci.* **2022**, *12*, 448. [[CrossRef](#)] [[PubMed](#)]
26. Dahmen-Zimmer, K.; Jansen, P. Karate and dance training to improve balance and stabilize mood in patients with Parkinson's disease: A feasibility study. *Front. Med.* **2017**, *4*, 237. [[CrossRef](#)] [[PubMed](#)]
27. Ferrazzoli, D.; Ortelli, P.; Madeo, G.; Giladi, N.; Petzinger, G.M.; Frazzitta, G. Basal ganglia and beyond: The interplay between motor and cognitive aspects in Parkinson's disease rehabilitation. *Neurosci. Biobehav. Rev.* **2018**, *90*, 294–308. [[CrossRef](#)] [[PubMed](#)]
28. Irazoki, E.; Contreras-Somoza, L.M.; Toribio-Guzmán, J.M.; Jenaro-Río, C.; van der Roest, H.; Franco-Martín, M.A. Technologies for cognitive training and cognitive rehabilitation for people with mild cognitive impairment and dementia. A systematic review. *Front. Psychol.* **2020**, *11*, 648. [[CrossRef](#)]
29. Dimyan, M.A.; Cohen, L.G. Neuroplasticity in the context of motor rehabilitation after stroke. *Nat. Rev. Neurol.* **2011**, *7*, 76–85. [[CrossRef](#)]
30. Cations, M.; Laver, K.E.; Crotty, M.; Cameron, I.D. Rehabilitation in dementia care. *Age Ageing* **2018**, *47*, 171–174. [[CrossRef](#)]
31. Cotelli, M.; Manenti, R.; Brambilla, M.; Gobbi, E.; Ferrari, C.; Binetti, G.; Cappa, S.F. Cognitive telerehabilitation in mild cognitive impairment, Alzheimer's disease and frontotemporal dementia: A systematic review. *J. Telemed. Telecare* **2019**, *25*, 67–79. [[CrossRef](#)]
32. Ferraris, C.; Ronga, I.; Pratola, R.; Coppo, G.; Bosso, T.; Falco, S.; Amprimo, G.; Pettiti, G.; Lo Priore, S.; Priano, L.; et al. Usability of the REHOME solution for the telerehabilitation in neurological diseases: Preliminary results on motor and cognitive platforms. *Sensors* **2022**, *22*, 9467. [[CrossRef](#)] [[PubMed](#)]
33. Reis, E.; Postolache, G.; Teixeira, L.; Arriaga, P.; Lima, M.L.; Postolache, O. Exergames for motor rehabilitation in older adults: An umbrella review. *Phys. Ther. Rev.* **2019**, *24*, 84–99. [[CrossRef](#)]
34. López-Nava, I.H.; Rodríguez, M.D.; García-Vázquez, J.P.; Perez-Sanpablo, A.I.; Quiñones-Urióstegui, I.; Meneses-Peñaloza, A.; Castillo, V.; Cuaya-Simbro, G.; Armenta, J.S.; Martínez, A.; et al. Current state and trends of the research in exergames for the elderly and their impact on health outcomes: A scoping review. *J. Ambient Intell. Humaniz. Comput.* **2022**, 1–33. [[CrossRef](#)]
35. Amprimo, G.; Masi, G.; Priano, L.; Azzaro, C.; Galli, F.; Pettiti, G.; Mauro, A.; Ferraris, C. Assessment tasks and virtual exergames for remote monitoring of Parkinson's disease: An integrated approach based on Azure Kinect. *Sensors* **2022**, *22*, 8173. [[CrossRef](#)] [[PubMed](#)]
36. de Melo Cerqueira, T.M.; de Moura, J.A.; de Lira, J.O.; Leal, J.C.; D'Amelio, M.; do Santos Mendes, F.A. Cognitive and motor effects of Kinect-based games training in people with and without Parkinson disease: A preliminary study. *Physiother. Res. Int.* **2020**, *25*, e1807. [[CrossRef](#)] [[PubMed](#)]
37. Liu, H.; Xing, Y.; Wu, Y. Effect of Wii Fit exercise with balance and lower limb muscle strength in older adults: A meta-analysis. *Front. Med.* **2022**, *9*, 812570. [[CrossRef](#)] [[PubMed](#)]
38. Chen, W.; Bang, M.; Krivonos, D.; Schimek, H.; Naval, A. An immersive virtual reality exergame for people with Parkinson's disease. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2020; pp. 138–145.
39. Eisapour, M.; Cao, S.; Domenicucci, L.; Boger, J. Virtual reality exergames for people living with dementia based on exercise therapy best practices. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2018**, *62*, 528–532. [[CrossRef](#)]
40. Chu, C.H.; Biss, R.K.; Cooper, L.; Quan, A.M.L.; Matulis, H. Exergaming platform for older adults residing in long-term care homes: User-centered design, development, and usability study. *JMIR Serious Games* **2021**, *9*, e22370. [[CrossRef](#)] [[PubMed](#)]

41. Park, C.; Mishra, R.K.; York, M.K.; Enriquez, A.; Lindsay, A.; Barchard, G.; Vaziri, A.; Najafi, B. Tele-medicine based and self-administered interactive exercise program (Tele-exergame) to improve cognition in older adults with mild cognitive impairment or dementia: A feasibility, acceptability, and proof-of-concept study. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16361. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Barry, G.; Galna, B.; Rochester, L. The role of exergaming in Parkinson's disease rehabilitation: A systematic review of the evidence. *J. Neuroeng. Rehabil.* **2014**, *11*, 33. [\[CrossRef\]](#) [\[PubMed\]](#)
43. van Santen, J.; Dröes, R.M.; Holstege, M.; Henkemans, O.B.; van Rijn, A.; de Vries, R.; van Straten, A.; Meiland, F. Effects of exergaming in people with dementia: Results of a systematic literature review. *J. Alzheimer's Dis.* **2018**, *63*, 741–760. [\[CrossRef\]](#)
44. Nonnekes, J.; Nieuwboer, A. Towards personalized rehabilitation for gait impairments in Parkinson's disease. *J. Parkinson's Dis.* **2018**, *8*, S101–S106. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Omboni, S.; Padwal, R.S.; Alessa, T.; Benczúr, B.; Green, B.B.; Hubbard, I.; Kario, K.; Khan, N.A.; Konradi, A.; Logan, A.G.; et al. The worldwide impact of telemedicine during COVID-19: Current evidence and recommendations for the future. *Connect. Health* **2022**, *1*, 7. [\[CrossRef\]](#)
46. Dong, S.; Reder, L.M.; Yao, Y.; Liu, Y.; Chen, F. Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Res.* **2015**, *1616*, 146–156. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Pope, A.T.; Bogart, E.H.; Bartolome, D.S. Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* **1995**, *40*, 187–195. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Eldenfria, A.; Al-Samarraie, H. Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: An EEG study. *Int. J. Hum.–Comput. Interact.* **2019**, *35*, 1960–1974. [\[CrossRef\]](#)
49. Gola, M.; Magnuski, M.; Szumska, I.; Wróbel, A. EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *Int. J. Psychophysiol.* **2013**, *89*, 334–341. [\[CrossRef\]](#)
50. Kay, L.M. Theta oscillations and sensorimotor performance. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3863–3868. [\[CrossRef\]](#)
51. Brauns, I.; Teixeira, S.; Velasques, B.; Bittencourt, J.; Machado, S.; Cagy, M.; Gongora, M.; Bastos, V.H.; Machado, D.; Sandoval-Carrillo, A.; et al. Changes in the theta band coherence during motor task after hand immobilization. *Int. Arch. Med.* **2014**, *7*, 51. [\[CrossRef\]](#)
52. Edwards, L.L.; King, E.M.; Buetefisch, C.M.; Borich, M.R. Putting the “sensory” into sensorimotor control: The role of sensorimotor integration in goal-directed hand movements after stroke. *Front. Integr. Neurosci.* **2019**, *13*, 16. [\[CrossRef\]](#)
53. Nakayashiki, K.; Saeki, M.; Takata, Y.; Hayashi, Y.; Kondo, T. Modulation of event-related desynchronization during kinematic and kinetic hand movements. *J. Neuroeng. Rehabil.* **2014**, *11*, 90. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Szafir, D.; Mutlu, B. Pay attention! Designing adaptive agents that monitor and improve user engagement. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 11–20.
55. Szafir, D.; Mutlu, B. ARTful: Adaptive review technology for flipped learning. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; pp. 1001–1010.
56. Nuamah, J.; Seong, Y. Support vector machine (SVM) classification of cognitive tasks based on electroencephalography (EEG) engagement index. *Brain-Comput. Interfaces* **2018**, *5*, 1–12. [\[CrossRef\]](#)
57. McMahan, T.; Parberry, I.; Parsons, T.D. Evaluating player task engagement and arousal using electroencephalography. *Procedia Manuf.* **2015**, *3*, 2303–2310. [\[CrossRef\]](#)
58. Yücel, Z.; Koyama, S.; Monden, A.; Sasakura, M. Estimating level of engagement from ocular landmarks. *Int. J. Hum. Comput. Interact.* **2020**, *36*, 1527–1539. [\[CrossRef\]](#)
59. Ranti, C.; Jones, W.; Klin, A.; Shultz, S. Blink rate patterns provide a reliable measure of individual engagement with scene content. *Sci. Rep.* **2020**, *10*, 8267. [\[CrossRef\]](#)
60. Daza, R.; DeAlcala, D.; Morales, A.; Tolosana, R.; Cobos, R.; Fierrez, J. ALEBk: Feasibility study of attention level estimation via blink detection applied to e-learning. *arXiv* **2021**, arXiv:2112.09165.
61. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.
62. Planinc, R.; Nake, I.; Kampel, M. Exergame design guidelines for enhancing elderly's physical and social activities. In Proceedings of the AMBIENT 2013, The Third International Conference on Ambient Computing, Applications, Services and Technologies, Porto, Portugal, 29 September–3 October 2013; pp. 58–63.
63. Hébert, S.; Béland, R.; Dionne-Fournelle, O.; Crête, M.; Lupien, S.J. Physiological stress response to video-game playing: The contribution of built-in music. *Life Sci.* **2005**, *76*, 2371–2380. [\[CrossRef\]](#)
64. Amprimo, G.; Ferraris, C.; Masi, G.; Pettiti, G.; Priano, L. GMH-D: Combining Google MediaPipe and RGB-depth cameras for hand motor skills remote assessment. In Proceedings of the 2022 IEEE International Conference on Digital Health (ICDH), Barcelona, Spain, 11–15 July 2022.
65. Stroop, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **1935**, *18*, 643. [\[CrossRef\]](#)
66. Guan, J.; Wade, M.G. The effect of aging on adaptive eye-hand coordination. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2000**, *55*, P151–P162. [\[CrossRef\]](#)
67. Boisseau, E.; Scherzer, P.; Cohen, H. Eye-hand coordination in aging and in Parkinson's disease. *Aging Neuropsychol. Cogn.* **2002**, *9*, 266–275. [\[CrossRef\]](#)

68. Hart, S.G. NASA Task Load Index (TLX). 1986. Available online: <https://ntrs.nasa.gov/citations/20000021487> (accessed on 15 March 2022).
69. Hart, S.G. NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Sydney, Australia, 20–22 November 2006; Sage Publications: Los Angeles, CA, USA, 2006; Volume 50, pp. 904–908.
70. Hendy, K.C.; Hamilton, K.M.; Landry, L.N. Measuring subjective workload: When is one scale better than many? *Hum. Factors* **1993**, *35*, 579–601. [\[CrossRef\]](#)
71. Said, S.; Gozdziak, M.; Roche, T.R.; Braun, J.; Rössler, J.; Kaserer, A.; Spahn, D.R.; Nöthiger, C.B.; Tscholl, D.W. Validation of the raw National Aeronautics and Space Administration Task Load Index (NASA-TLX) questionnaire to assess perceived workload in patient monitoring tasks: Pooled analysis study using mixed models. *J. Med. Internet Res.* **2020**, *22*, e19472. [\[CrossRef\]](#) [\[PubMed\]](#)
72. Prabaswari, A.D.; Basumerda, C.; Utomo, B.W. The mental workload analysis of staff in study program of private educational organization. In *Proceedings of the IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019; Volume 528, p. 012018.
73. Barry, G.; van Schaik, P.; MacSween, A.; Dixon, J.; Martin, D. Exergaming (XBOX Kinect™) versus traditional gym-based exercise for postural control, flow and technology acceptance in healthy adults: A randomised controlled trial. *BMC Sports Sci. Med. Rehabil.* **2016**, *8*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Lange, B.; Chang, C.Y.; Suma, E.; Newman, B.; Rizzo, A.S.; Bolas, M. Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 1831–1834.
75. Mocanu, I.; Marian, C.; Rusu, L.; Arba, R. A Kinect based adaptive exergame. In Proceedings of the 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 8–10 September 2016.
76. Tölgyessy, M.; Dekan, M.; Chovanec, L.; Hubinský, P. Evaluation of the Azure Kinect and its comparison to Kinect V1 and Kinect V2. *Sensors* **2021**, *21*, 413. [\[CrossRef\]](#)
77. Face Mesh. Available online: [https://google.github.io/mediapipe/solutions/face\\_mesh.html](https://google.github.io/mediapipe/solutions/face_mesh.html) (accessed on 14 December 2022).
78. Thorey, V.; Guillot, A.; El Kanbi, K.; Harris, M.; Arnal, P. 1211 Assessing the Accuracy of a Dry-EEG Headband for Measuring Brain Activity, Heart Rate, Breathing and Automatic Sleep Staging. *Sleep* **2020**, *43*, A463. [\[CrossRef\]](#)
79. Li, R.; Principe, J.C. Blinking artifact removal in cognitive EEG data using ICA. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 5273–5276.
80. Motamedi-Fakhr, S.; Moshrefi-Torbati, M.; Hill, M.; Hill, C.M.; White, P.R. Signal processing techniques applied to human sleep EEG signals—A review. *Biomed. Signal Process. Control* **2014**, *10*, 21–33. [\[CrossRef\]](#)
81. Rechichi, I.; Amato, F.; Cicolin, A.; Olmo, G. Single-Channel EEG Detection of REM Sleep Behaviour Disorder: The Influence of REM and Slow Wave Sleep. In Proceedings of the International Work-Conference on Bioinformatics and Biomedical Engineering, Maspalomas, Spain, 27–30 June 2022; pp. 381–394.
82. Stieger, J.R.; Engel, S.A.; He, B. Continuous sensorimotor rhythm based brain computer interface learning in a large population. *Sci. Data* **2021**, *8*, 98. [\[CrossRef\]](#)
83. Yuan, H.; He, B. Brain–computer interfaces using sensorimotor rhythms: Current state and future perspectives. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 1425–1435. [\[CrossRef\]](#)
84. Coelli, S.; Sclocco, R.; Barbieri, R.; Reni, G.; Zucca, C.; Bianchi, A.M. EEG-based index for engagement level monitoring during sustained attention. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 1512–1515.
85. Li, X.; Jiang, Y.; Hong, J.; Dong, Y.; Yao, L. Estimation of cognitive workload by approximate entropy of EEG. *J. Mech. Med. Biol.* **2016**, *16*, 1650077. [\[CrossRef\]](#)
86. Angelakis, E.; Lubar, J.F.; Stathopoulou, S. Electroencephalographic peak alpha frequency correlates of cognitive traits. *Neurosci. Lett.* **2004**, *371*, 60–63. [\[CrossRef\]](#)
87. Akiyama, M.; Tero, A.; Kawasaki, M.; Nishiura, Y.; Yamaguchi, Y. Theta-alpha EEG phase distributions in the frontal area for dissociation of visual and auditory working memory. *Sci. Rep.* **2017**, *7*, 42776. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Remeseiro, B.; Fernández, A.; Lira, M. Automatic eye blink detection using consumer web cameras. In *Advances in Computational Intelligence*; Springer International Publishing: Cham, Switzerland, 2015; pp. 103–114.
89. Jordan, A.A.; Pegatoquet, A.; Castagnetti, A.; Raybaut, J.; Le Coz, P. Deep learning for eye blink detection implemented at the edge. *IEEE Embed. Syst. Lett.* **2021**, *13*, 130–133. [\[CrossRef\]](#)
90. Zdarsky, N.; Treue, S.; Esghaei, M. A deep learning-based approach to video-based eye tracking for human psychophysics. *Front. Hum. Neurosci.* **2021**, *15*, 685830. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Dewi, C.; Chen, R.C.; Jiang, X.; Yu, H. Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks. *PeerJ Comput. Sci.* **2022**, *8*, e943. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Zhu, T.; Zhang, C.; Wu, T.; Ouyang, Z.; Li, H.; Na, X.; Liang, J.; Li, W. Research on a real-time driver fatigue detection algorithm based on facial video sequences. *Appl. Sci.* **2022**, *12*, 2224. [\[CrossRef\]](#)
93. Jamovi—Open Statistical Software for the Desktop and Cloud. Available online: <https://www.jamovi.org> (accessed on 19 December 2022).
94. Urbanowicz, R.J.; Meeker, M.; La Cava, W.; Olson, R.S.; Moore, J.H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **2018**, *85*, 189–203. [\[CrossRef\]](#)



95. Rogers, J.M.; Jensen, J.; Valderrama, J.T.; Johnstone, S.J.; Wilson, P.H. Single-channel EEG measurement of engagement in virtual rehabilitation: A validation study. *Virtual Real.* **2021**, *25*, 357–366. [[CrossRef](#)]
96. Lee, J.C.; Tan, D.S. Using a low-cost electroencephalograph for task classification in HCI research. In Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology, Montreux, Switzerland, 15–18 October 2006; pp. 81–90.
97. Chanel, G.; Rebetez, C.; Bétrancourt, M.; Pun, T. Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2011**, *41*, 1052–1063. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.