



# Article A Hybrid Asynchronous Brain–Computer Interface Based on SSVEP and Eye-Tracking for Threatening Pedestrian Identification in Driving

Jianxiang Sun and Yadong Liu \*

College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China

\* Correspondence: liuyadong@nudt.edu.cn

Abstract: A brain-computer interface (BCI) based on steady-state visual evoked potential (SSVEP) has achieved remarkable performance in the field of automatic driving. Prolonged SSVEP stimuli can cause driver fatigue and reduce the efficiency of interaction. In this paper, a multi-modal hybrid asynchronous BCI system combining eye-tracking and EEG signals is proposed for dynamic threatening pedestrian identification in driving. Stimuli arrows of different frequencies and directions are randomly superimposed on pedestrian targets. Subjects scan the stimuli according to the direction of arrows until the threatening pedestrian is selected. The thresholds determined by offline experiments are used to distinguish between working and idle states of the asynchronous online experiments. Subjects need to judge and select potentially threatening pedestrians in online experiments according to their own subjective experience. The three proposed decisions filter out the results with low confidence and effectively improve the selection accuracy of hybrid BCI. The experimental results of six subjects show that the proposed hybrid asynchronous BCI system achieves better performance compared with a single SSVEP-BCI, with an average selection time of 1.33 s, an average selection accuracy of 95.83%, and an average information transfer rate (ITR) of 67.50 bits/min. These results indicate that our hybrid asynchronous BCI has great application potential in dynamic threatening pedestrian identification in driving.

**Keywords:** brain–computer interface (BCI); steady-state visual evoked potential (SSVEP); electroencephalography (EEG); threatening pedestrians; eye-tracking

## 1. Introduction

In recent years, the brain–computer interface (BCI) has become a research hotspot in the field of artificial intelligence, aiming at building communication between the human brain and external devices. Electroencephalography (EEG), reflecting brain activity, is the common signal source of BCI applications. As a non-invasive and low-cost signal, EEG has shown high levels of reliability [1,2]. As a new interactive mode, BCI has been widely used in the fields of medical assistance [3] automobile driving [4], robot control [5], etc.

As a complex BCI application, there is a direct control pathway between the brain and the vehicle in Brain-Controlled Vehicles (BCV). At present, the BCI paradigms adopted by BCV systems are mainly P300 [6], motor imagery (MI) [7], and steady-state visual evoked potential (SSVEP) [8]. P300, which is always evoked by a visual stimulus with poor real-time performance, can only be used for the control of static targets, such as switches, wipers, etc. The real-time performance of MI is also poor, and the degrees of freedom available is limited (generally less than four), which makes it impossible to complete the overall driving task. SSVEP, which is an electrophysiological response to a repetitive visual stimulus, has a high information transfer rate (ITR) and good real-time performance. When subjects focus their attention on a stimulus, the corresponding frequency appears in the representation of the EEG signals recorded mainly in occipital regions [9]. Studies [10,11] have shown that the



**Citation:** Sun, J.; Liu, Y. A Hybrid Asynchronous Brain–Computer Interface Based on SSVEP and Eye-Tracking for Threatening Pedestrian Identification in Driving. *Electronics* **2022**, *11*, 3171. https:// doi.org/10.3390/electronics11193171

Academic Editor: Jose Eugenio Naranjo

Received: 3 September 2022 Accepted: 30 September 2022 Published: 2 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/). human cerebral cortex will produce SSVEP characteristic components at the fundamental or multiplicative frequency of the target stimuli when exposed to a fixed-frequency visual stimulus. The target stimuli can be identified by detecting the dominant frequency of SSVEP. Based on its high applicability, simplicity, and high accuracy, BCI adopting SSVEP is conducive to the selection of threat targets in the process of automatic driving. However, it is easy to cause fatigue by long visual flicker.

With the maturity of eye-tracking technology and the continuous improvement of human requirements for interaction comfort, interaction based on eye-tracking has attracted more and more attention. In contrast to EEG, interaction based on eye-tracking is more natural, which can further reduce fatigue. In addition, eye-tracking interaction learning is inexpensive, and most users can operate it without special training [12]. However, there are still some drawbacks to eye-tracking. Some eye movements are not guided by volitional attention. If the system does not distinguish between these eye movements, it is likely to misunderstand human intentions and cause false triggering, which is called the "Midas Touch" problem [13]. In addition, eye-tracking technology is not completely reliable. In addition, some random instability factors can cause system errors. Several eye-movement interactions have been applied to text spelling [14] and robot control [15].

A hybrid BCI system is generally composed of one BCI and another system (which might be another BCI) and can perform better than a conventional BCI [16]. Some studies [17] adopt hybrid systems to recognize characters, combining EEG and EOG. In addition, eye-tracking, which is a popular technology in the field of computer vision, has been gradually adopted to combine with BCI to control games [18], robotic arms [19], and drones [20].

At present, the significant improvement of computer information fusion capability is constantly promoting the development of automatic driving. Autonomous driving is gradually moving from specific scenarios (such as highways, experimental parks) to complex urban traffic. Urban traffic conditions are relatively complex, with many dynamic pedestrian targets and variable trajectories. In such a complex road situation, the environment perception approach based on computer vision technology cannot predict a threatening pedestrian target quickly and accurately. Driver intention is integrated into the vehicle's environment perception through BCI, which can help to improve the comfort and safety of driving.

In this work, a multi-modal hybrid BCI combining SSVEP with eye-tracking is proposed for the selection of potentially threatening pedestrians. The arrows in different directions are randomly superimposed on pedestrian targets. SSVEP is evoked by the stimuli of the corresponding frequency while subjects scan the threatening pedestrian target according to the direction of arrows. I-VT filter is applied to process eye-movement tracks, and canonical correlation analysis (CCA) is adopted to detect EEG signals. The combination of eye-tracking and EEG can not only be used to distinguish between working and idle states, but also shorten target selection time and improve accuracy. The experimental results of six subjects show that the proposed hybrid asynchronous BCI system of eye-tracking and SSVEP achieves better performance compared with a single SSVEP-BCI, with an average selection time of 1.33 s, an average selection accuracy of 95.83%, and an average information transfer rate of 67.50 bits/min.

The remainder of this paper is presented as follows: Section 2 introduces a hybrid BCI system, target detection and tracking, graphical stimuli interface, participants, signal acquisition and preprocessing. Section 3 presents the process of experiments, evaluation metrics, and the results of experiments. Section 4 is the discussion of the hybrid BCI system, and Section 5 summarizes the main work of this paper.

#### 2. Materials and Methods

Figure 1 shows the overall framework for threatening pedestrian identification. Yolov5 is introduced to detect pedestrian targets, and DeepSORT is used to track pedestrians. SSVEP stimuli of different frequencies are superimposed on the obtained pedestrian coor-

dinates. Subjects scan pedestrians according to the direction of superimposed arrowhead stimuli. The three decisions effectively reduce the false positives and improve the reliability of threatening pedestrian identification.



Figure 1. Hybrid asynchronous BCI system for dynamic pedestrian detection.

## 2.1. System Description

The purpose of this study is to evaluate the performance of a multi-modal BCI that combines eye-tracking and SSVEP for pedestrian tracking and selection. First, the ZED2 camera collects real-time video of driving foreground road conditions and performs multitarget detection and tracking. The coordinates and IDs of the pedestrians are transmitted to the remote computer through the LAN. Second, flashing arrows of different stimuli appear on the targets after receiving data, and follow their movements. The arrows point to a random distribution. After calibrating the eye tracker, participants gaze at the stimulation interface. The eye tracker and the EEG acquisition instrument begin to collect the corresponding signals at the same time. The flow of online signal processing is shown in Figure 1. The sampling frequency of the eye tracker is 60 Hz. In the processing of eye-movement data, I-VT filter is introduced to process visual trajectories. Decision I: When the confidence of the trajectory change over 60 consecutive sampling points exceeds 70%, the result  $\{r1, r2 \dots\}$  of eye-tracking is output. In the processing of EEG signals, the canonical correlation analysis (CCA) algorithm performs feature extraction on 1000 ms of EEG data and outputs the maximum correlation coefficient ( $\rho$ ). Decision II: Output the result  $\{s1\}$  of EEG selection when  $\rho$  exceeds the pre-set threshold. Decision III: Output selection target when  $\{r1, r2 \dots\} \cap \{s1\} \neq \emptyset$ . Otherwise, no result output is considered an idle state. Slide the window forward 200 ms to acquire eye-tracking data and EEG data for the next 1 s and process until the target result is output.

#### 2.2. Target Detection and Tracking

Detecting and tracking pedestrians in the driving environment can reduce the cognitive load of drivers to a certain extent, and assist the vehicle intelligence system in making decisions, which plays a very important role in improving the safety of intelligent vehicles, and is a hot research topic in intelligent driving and computer vision.

In recent years, with the development of big data and the improvement of computer performance, deep learning has been widely applied in the field of computer vision and has achieved good performance. As a representative target-detection algorithm at the present stage, YOLO algorithms have excellent performance in both detection speed and accuracy, which can achieve end-to-end training. YOLO takes the whole image as the input of the network, and directly outputs the coordinates and IDs of the objects after inference. Compared with other algorithms, yolov5s [21] has higher detection accuracy, faster detection speed, and lower consumption of computation, which can better meet the real-time requirements and be easier to apply in the actual systems. However, detecting the position of pedestrians is not enough. Each object must be tracked before being chosen.

Pedestrian detection determines the position and ID of the object in a particular frame, and pedestrian tracking locks the target in consecutive frames. Most application scenarios involve the tracking of multi-targets. DeepSORT [22] extracts the appearance characteristics of targets, and adopts recursive Kalman filtering and frame-by-frame correlation [23] to match the trajectory of multi-objects, which can effectively reduce the number of target IDs transitions. In this study, we use yolov5 and DeepSORT to process the driving foreground video, which realizes multi-object detection and tracking accurately and quickly, and obtains the position coordinates and IDs of pedestrians in real time.

### 2.3. Graphical Stimuli Interface

According to the object positions and IDs obtained by the object-detection and tracking module, flicker stimuli of different frequencies are superimposed on each pedestrian in Figure 2, and participants can achieve their selection by staring at stimuli. The length of arrows that flash alternately in black and white is 60 pixels. The frequency list is set to meet a variable number of pedestrians. Studies [24] have shown that a frequency band of 8~15 Hz can induce a relatively strong SSVEP response. Moreover, each frequency should satisfy that there is no overlap between the fundamental frequency and the frequency doubling. The interval between frequencies is set as large as possible to ensure the distinguishability of signals. Considering the above factors, the frequency list is set to 6.10 Hz, 8.18 Hz, 15.87 Hz, 12.85 Hz, 10.50 Hz, 8.97 Hz, 13.78 Hz, 9.98 Hz, 11.23 Hz, 7.08 Hz, 14.99 Hz, and 11.88 Hz. The frequencies of the superimposed stimuli are sequentially selected from the frequency list according to the coding order of each pedestrian ID. During the experiment, participants find the threatening target and follow his movement until the flicker of the target stack stops and turn yellow.



(a)

(b)

**Figure 2.** Stimuli presentation interface of hybrid BCI system on trial 6 of a block. (**a**) Arrows in different directions are randomly superimposed on pedestrians according to the frequency list corresponding to the ID order; (**b**) Subjects select the threatening object, and the arrow stops flashing and turns yellow.

## 2.4. Participants

Six healthy subjects (22–25 years; four men, two women) with an average age of 23.2 years were recruited from the campus and participated in the study. No one was left-handed. In addition, each subject reported no history of any psychiatric deficits. Following the Declaration of Helsinki, all subjects signed a letter of commitment after receiving a detailed description of the procedure.

## 2.5. Signal Acquisition and Processing

As shown in Figure 3, the eye-movement data were collected by Tobii Pro Nano at a frequency of 60 Hz and an operating distance of 80 cm. Subjects were required to calibrate their eye trackers before participating in experiments. An LCD screen (LEGION Y27gq-25, 1920  $\times$  1080 pixels) was used to present stimuli with a refresh rate of 240 Hz.



Figure 3. Placement of data acquisition equipment.

A 64-channel extended international 10/20 system was used to record the EEG signals in this experiment. Figure 4 shows the placement of 9 electrodes for EEG collection, which were placed in Pz, PO7, PO3, POz, PO4, PO8, O1, Oz and O2. The reference electrode was placed behind the right ear, and the ground electrode was placed on the forehead. Before data acquisition with BrainAmp DC amplifier (Brain Products GmbH, Germany), the impedance of each electrode was reduced to less than 10 k $\Omega$ . The sampling frequency was 200 Hz and was filtered by a 4–35 Hz bandpass and notch at 50 Hz. BCI2000 [25] served as the control platform to collect EEG signals, the PyGame [26], a Python expansion package, presented the stimuli interface, and MATLAB was responsible for real-time signal processing. The display interface and the control platform were connected through the TCP/IP protocol.



**Figure 4.** Placement of electrodes. The blue circles are the placements of the sampling electrode. The reference electrode is placed on the green circle behind the right ear, and the ground electrode is placed on the forehead.

Canonical correlation analysis (CCA) [27] was applied to extract features of preprocessed EEG signals, which fuses multi-channel data and identifies the target by calculating the correlation coefficient between multi-channel EEG signal and stimuli frequency. The target was the option corresponding to the maximum SSVEP response score. Periodic stimuli were represented as square-wave periodic signals that could be decomposed into Fourier harmonic series:

where *N* is the number of harmonics, *t* is the current time, *L* is the number of sampling points of the original signals, and *S* is the sampling rate of EEG. CCA is a multivariate

statistical analysis method, which calculates the maximum correlation coefficient ( $\rho$ ) of the linear combination of variables ( $x = X^T W_x$ ,  $y = Y^T W_y$ ) in two data sets (X, Y), to reflect the correlation of the two groups of signals. The calculation formula for  $\rho$  is as follows:

$$\rho(x,y) = \max_{\omega_x,\omega_y} \frac{E[x^T y]}{\sqrt{E[x^T x]E[y^T y]}} = \max_{\omega_x,\omega_y} \frac{E[\omega_x^T X Y^T \omega]_y}{\sqrt{E[\omega_x^T X X^T \omega_x]E[\omega_y^T Y Y^T \omega_y]}}, \quad (2)$$

The velocity threshold recognition (I-VT) filter is a popular speed-based eye-tracking method [28], which realizes the classification of eye tracks by analyzing the speed of eye movement. As shown in Formula (3), the eye-movement velocity can be obtained by the ratio of the distance between the two sampling points to the corresponding sampling time. Speed is commonly expressed in visual degrees per second (°/s). When the speed is higher than the set threshold, the sample associated with the speed is determined to be a saccade, and below the threshold is fixation.

$$v_x = \frac{x_2 - x_1}{t_2 - t_1}, v_y = \frac{y_2 - y_1}{t_2 - t_1},$$
(3)

where  $v_x$  represents the velocity in the *x* direction,  $v_y$  represents the velocity in the *y* direction, and  $(x_1, y_1)$  is the coordinate of the eyeball's position at the moment of  $t_1$ . Similarly,  $(x_2, y_2)$  is the coordinate of the eyeball position at  $t_2$  moment.

## 3. Results

#### 3.1. Evaluation Metrics

The performance of hybrid BCI selection is evaluated by accuracy and information transfer rate (ITR). In addition, ITR is calculated as follows (bits per minute):

$$ITR = \frac{\left(\log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1}\right)\right) * 60}{T},$$
(4)

$$T = t_s + t_b, \tag{5}$$

where *N* represents the total number of targets, *P* is the target selection accuracy, and *T* represents the time of target selection, including the stimuli flicker time of the target ( $t_s$ ) and flicker interval time ( $t_b$ ). It can be seen that the ITR is not only related to the classification accuracy, but also related to the number of selected targets.

## 3.2. Performance of the Offline Experiment

The threshold is set for the output of the SSVEP to distinguish between idle and working states in online experiments. If the maximum correlation coefficient is higher than the threshold, it is considered to be the working state; otherwise, it is considered to be the idle state. The goal is that the results are not output when the subjects are not staring at the target. In one trial of the offline experiment, participants tend to select threatening pedestrians by staring at flickering stimuli blocks according to cues. Each trial consists of an interval time of 2 s and a stimuli time of 4 s. Each participant participants are given a 5-min break.

Since the SSVEP responses of the participants are individually different, a specific threshold is set for each participant. Ten correct choices of each subject in the offline experiment are randomly selected to calculate the SSVEP response score, and the minimum value is taken as the threshold. As shown in Figure 5, the SSVEP response scores of S5 in 10 correct selection tasks are 0.6387, 0.5647, 0.7696, 0.7065, 0.5630, 0.6896, 0.7323, 0.5981, 0.4541, and 0.6721, respectively. In addition, the minimum response score (0.45) is set as the threshold for Subject 5. The SSVEP response scores of the 6 subjects in the random correct



selection tasks for ten times are shown in Table 1, and the statistical thresholds are 0.56, 0.62, 0.51, 0.47, 0.45, and 0.39, respectively.

**Figure 5.** The SSVEP response scores of S5 for selecting correctly for 10 trials. Multi-colored lines represent different stimuli frequencies. The minimum response score of 10 trials is 0.4541, which is set as the threshold for Subject 5.

Table 1. SSVEP response threshold of 6 subjects.

Subject	<b>S</b> 1	S2	<b>S</b> 3	<b>S</b> 4	<b>S</b> 5	<b>S</b> 6
Threshold	0.56	0.62	0.51	0.47	0.45	0.39

#### 3.3. Performance of Asynchronous Online Experiment

Thresholds obtained from offline experiments are used for online experiments. In the online experiment, the subjects choose threatening pedestrians according to their subjective cognition instead of prompts. There is no time limit for the subjects to complete the experiment. The system continuously outputs control commands to realize the relative real-time selection of threatening pedestrians. The other settings are the same as those for offline experiments. EEG collection and eye-tracking acquisition are performed simultaneously. At the beginning of the experiment, the subjects saccade the stimuli according to the direction of arrows until the color of the stimuli flicker turns yellow and lasts for 0.5 s. If there is a result output, the subjects proceed to the next trial.

Figure 6 shows the change in sight of the subject scanning arrows in the hybrid BCI system. In the one-second time window of the selected target, one of the coordinates of the saccade points remains basically unchanged, and the absolute value of the other coordinate change is about equal to the length of the arrows (60 pixels).

SSVEP-BCI is introduced to verify the effectiveness and availability of the hybrid BCI structure. In Table 2, several evaluation metrics such as the accuracy, target selection time, and ITR are shown to evaluate the performance of two models in which six subjects select dynamic threatening pedestrians. Hybrid BCI achieves a higher selection accuracy (95.83%), shorter selection time (1.33 s), and higher ITR (67.5 bits/min). Compared to SSVEP-BCI, the selection time is shortened by 0.69 s, the accuracy is improved by 5%, and the ITR is increased by 25.2 bits/min. Subject 2 performs perfectly in both SSVEP-BCI and hybrid BCI, with a selection accuracy of 100%. It is worth mentioning that Subject 2 selects threatening pedestrians within 1 s, and the ITR reaches 92.88 bits/min in hybrid BCI. Subject 5 performs

poorly in SSVEP-BCI with an accuracy of 80% and an ITR of 25.71 bits/min. By combining eye tracks with EEG data, the accuracy of target selection is significantly improved to 90% and the selection time is shortened from 2.3 s to 1.6 s. These results show that in the hybrid SSVEP architecture, the selection time and accuracy of subjects selecting dynamic threatening pedestrians meet the requirements of online experimental tasks. The advantage of hybrid BCI lies in the addition of eye tracks, which effectively avoids the wrong results caused by inattention. At the same time, the multi-modal fusion of eye movements and EEG enables subjects to make choices in a shorter time. The single eye-tracking system is not stable, and the phenomenon of "Midas Touch" often occurs. In actual traffic scenarios, the wrong choice of threatening targets will lead to traffic accidents caused by the inaccurate operation of self-driving vehicles. The stability and robustness of the hybrid BCI can ensure that the drivers can make the judgment and choose the threatening targets quickly and accurately in assisted driving.



**Figure 6.** The change of eye movements during saccade in the direction of arrows. (**a**) Top-to-bottom saccade; (**b**) Bottom-to-top saccade; (**c**) Right-to-left scanning; (**d**) Left-to-right saccade. The blue line represents the change in the *x* direction and the orange line represents the change in the *y* direction. The distance between the two gray lines is 60 pixels.

Subject	SSVEP			Hybrid		
	Mean Time (s)	Accuracy (%)	ITR (bits/min)	Mean Time (s)	Accuracy (%)	ITR (bits/min)
S1	1.9 + 0.5	95	48.39	1.2 + 0.5	95	68.31
S2	1.8 + 0.5	100	60.57	1.0 + 0.5	100	92.88
S3	2.0 + 0.5	95	46.45	1.3 + 0.5	100	77.39
S4	1.9 + 0.5	90	41.32	1.4 + 0.5	95	61.12
S5	2.3 + 0.5	80	25.71	1.6 + 0.5	90	47.23
S6	2.2 + 0.5	85	31.38	1.5 + 0.5	95	58.07
Mean	2.02 + 0.5	90.83	42.30	1.33 + 0.5	95.83	67.50
Std	0.18	6.72	11.43	0.20	3.44	14.62

**Table 2.** Results of asynchronous online selection of threatening pedestrians by SSVEP-BCI and hybrid BCI.

### 4. Discussion

In complex road environments, pedestrians have a great impact on the safety of vehicle driving. The threat to driving safety is usually only a few pedestrians with special locations or trajectories. However, they significantly interfere with the driving route, and even directly determine whether the vehicle can pass safely. Therefore, marking potential threats from many pedestrian targets and feedbacking the location information of these pedestrians to the computer can help vehicles make safer decisions in subsequent control.

This paper proposes a hybrid BCI paradigm for threatening pedestrian selection based on object detection and tracking. The object-detection and tracking method based on deep learning obtains the coordinates and IDs of pedestrian targets, providing initial information for hybrid BCI. This study takes the traffic scenes as the background and combines computer vision with hybrid BCI, aiming at the judgment of dynamic threatening pedestrians. Participants need to judge and select pedestrians who pose a threat to driving safety according to their own subjective experience. Six subjects participated in offline experiments and asynchronous online experiments. The thresholds determined by offline experiments are used to distinguish between the working and idle states of the online experiments. In asynchronous online experiments, the average selection time is 1.33 s, average accuracy reaches 95.83%, and an average ITR reaches 67.5 bits/min. These results show that hybrid BCI has great application potential in dynamic threatening pedestrian selection.

#### 5. Conclusions

This paper designs a hybrid BCI that combines eye-tracking and EEG for threatening pedestrian recognition in the driving environment. The experimental results of six subjects show that hybrid BCI achieves better performance compared with a single SSVEP-BCI, with an average selection time of 1.33 s, an average selection accuracy of 95.83%, and an average information transfer rate (ITR) of 67.50 bits/min. The three proposed decisions filter out the results with low confidence, which effectively improves the selection accuracy of hybrid BCI. The driver's understanding of the environment is fed back to the machine, and humanmachine collaborative driving is realized to a certain extent. Compared with methods that rely solely on computer vision, this method has more advanced environmental semantic understanding ability and is safer and more reliable in driving. The system has been verified online in several specific experimental scenarios, but its applicability needs to be further enhanced in scenarios where multiple threatening pedestrians exist or threatening pedestrians suddenly appear. In future work, we will develop more rapid and accurate signal-processing methods to analyze SSVEP, and combine Bayesian probability to decide on threatening pedestrians in different scenarios.

**Author Contributions:** Conceptualization, Y.L.; methodology, J.S.; software, J.S.; validation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, J.S. and Y.L.; supervision, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by National Natural Science Foundation of China, grant number U19A2083.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

#### References

- 1. Sánchez-Reyes, L.M.; Rodríguez-Reséndiz, J.; Avecilla-Ramírez, G.N.; García-Gomar, M.L.; Robles-Ocampo, J.B. Impact of eeg parameters detecting dementia diseases: A systematic review. *IEEE Access* **2021**, *9*, 78060–78074. [CrossRef]
- Ortiz-Echeverri, C.; Paredes, O.; Salazar-Colores, J.S.; Rodríguez-Reséndiz, J.; Romo-Vázquez, R. A Comparative Study of Time and Frequency Features for EEG Classification. *Lat. Am. Conf. Biomed. Eng.* 2019, 75, 91–97.
- Liu, Y.; Liu, Y.; Tang, J.; Yin, E.; Hu, D.; Zhou, Z. A self-paced BCI prototype system based on the in-corporation of an intelligent environment-understanding approach for rehabilitation hospital environmental. *Comput. Biol. Med.* 2020, *118*, 103618. [CrossRef] [PubMed]
- 4. Li, W.; Duan, F.; Sheng, S.; Xu, C.; Liu, R.; Zhang, Z.; Jiang, X. A human-vehicle collaborative sim-ulated driving system based on hybrid brain–computer interfaces and computer vision. *IEEE Trans. Cogn. Dev. Syst.* 2017, *10*, 810–822. [CrossRef]
- Leeb, R.; Tonin, L.; Rohm, M.; Desideri, L.; Carlson, T.; Millan, J.D.R. Towards Independence: A BCI Telepresence Robot for People With Severe Motor Disabilities. *Proc. IEEE* 2015, 103, 969–982. [CrossRef]
- Bi, L.; Fan, X.-A.; Luo, N.; Jie, K.; Li, Y.; Liu, Y. A Head-Up Display-Based P300 Brain–Computer Interface for Destination Selection. IEEE Trans. Intell. Transp. Syst. 2013, 14, 1996–2001. [CrossRef]
- Zhuang, J.; Yin, G. Motion control of a four-wheel-independent-drive electric vehicle by motor imagery EEG based BCI system. In Proceedings of the 2017 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 5449–5454.
- 8. Stawicki, P.; Gembler, F.; Volosyak, I. Driving a Semiautonomous Mobile Robotic Car Controlled by an SSVEP-Based BCI. *Comput. Intell. Neurosci.* **2016**, 2016, 4909685. [CrossRef]
- 9. Fernandez-Fraga, S.M.; Aceves-Fernandez, M.A.; Rodríguez-Resendíz, J.; Pedraza-Ortega, J.C.; Ramos-Arreguín, J.M. Steady-state visual evoked potential (SSEVP) from EEG signal modeling based upon recurrence plots. *Evol. Syst.* 2019, *10*, 97–109. [CrossRef]
- 10. Ortner, R.; Allison, B.Z.; Korisek, G.; Gaggl, H.; Pfurtscheller, G. An SSVEP BCI to Control a Hand Orthosis for Persons With Tetraplegia. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2010**, *19*, 784–796. [CrossRef]
- 11. Wu, Z.; Lai, Y.; Xia, Y.; Wu, D.; Yao, D. Stimulator selection in SSVEP-based BCI. Med. Eng. Phys. 2008, 30, 1079–1088. [CrossRef]
- 12. Pavan Kumar, B.N.; Balasubramanyam, A.; Patil, A.K.; Chethana, B.; Chai, Y.H. GazeGuide: An Eye-Gaze-Guided Active Immersive UAV Camera. *Appl. Sci.* 2020, *10*, 1668.
- Jacob, R.J.K. Eye Tracking in Advanced Interface Design. In *Virtual Environments and Advanced Interface Design*; Oxford University Press: Oxford, UK, 1995; Volume 258, p. 288.
- 14. Stawicki, P.; Gembler, F.; Rezeika, A.; Volosyak, I. A Novel Hybrid Mental Spelling Application Based on Eye Tracking and SSVEP-Based BCI. *Brain Sci.* 2017, 7, 35. [CrossRef]
- 15. Kishore, S.; González-Franco, M.; Hintemüller, C.; Kapeller, C.; Guger, C.; Slater, M.; Blom, K.J. Comparison of SSVEP BCI and eye tracking for controlling a humanoid robot in a social environment. *Presence Teleoper. Virtual Environ.* 2014, 23, 242–252. [CrossRef]
- 16. Pfurtscheller, G.; Allison, B.Z.; Bauernfeind, G.; Brunner, C.; Solis Escalante, T.; Scherer, R.; Zander, T.O.; Mueller-Putz, G.; Neuper, C.; Birbaumer, N. The hybrid BCI. *Front. Neurosci.* **2010**, *4*, 3. [CrossRef] [PubMed]
- 17. Zhou, Y.; He, S.; Huang, Q.; Li, Y. A Hybrid Asynchronous Brain-Computer Interface Combining SSVEP and EOG Signals. *IEEE Trans. Biomed. Eng.* **2020**, *67*, 2881–2892. [CrossRef]
- 18. Myna, K.N.; Tarpin-Bernard, F. Evaluation and comparison of a multimodal combination of BCI paradigms and eye tracking with affordable consumer-grade hardware in a gaming context. *IEEE Trans. Comput. Intell. AI Games* **2013**, *5*, 150–154. [CrossRef]
- McMullen, D.P.; Hotson, G.; Katyal, K.D.; Wester, B.A.; Fifer, M.S.; McGee, T.G.; Harris, A.; Johannes, M.S.; Vogelstein, R.J.; Ravitz, A.D.; et al. Demonstration of a semi-autonomous hybrid brain–machine interface using human intracranial EEG, eye tracking, and computer vision to control a robotic upper limb prosthetic. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2013, 22, 784–796. [CrossRef]
- 20. Kim, B.H.; Kim, M.; Jo, S. Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking. *Comput. Biol. Med.* **2014**, *51*, 82–92. [CrossRef]
- 21. Thuan, D. Evolution of Yolo Algorithm and Yolov5: The State-of-the-Art Object Detention Algorithm; Oulu University of Applied Sciences: Oulu, Finland, 2021.
- 22. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3645–3649.
- Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; Upcroft, B. Simple online and realtime tracking. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3464–3468.
- 24. Zhang, N.; Zhou, Z.; Liu, Y.; Yin, E.; Jiang, J.; Hu, D. A Novel Single-Character Visual BCI Paradigm With Multiple Active Cognitive Tasks. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 3119–3128. [CrossRef]

- 25. Schalk, G.; McFarland, D.; Hinterberger, T.; Birbaumer, N.; Wolpaw, J. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1034–1043. [CrossRef]
- 26. Kelly, S. Basic introduction to pygame. In *Python, PyGame and Raspberry Pi Game Development;* Springer: Berlin/Heidelberg, Germany, 2016; pp. 59–65.
- 27. Lin, Z.; Zhang, C.; Wu, W.; Gao, X. Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2610–2614. [CrossRef] [PubMed]
- 28. Olsen, A. The Tobii I-VT Fixation Filter; Tobii Technology: Danderyd Municipality, Sweden, 2012; Volume 21, pp. 4–19.