Closed-loop EEG study on visual recognition during driving

Ruslan Aydarkhanov[1], Marija Ušćumlić[2], Ricardo Chavarriaga[3,4], Lucian Gheorghe[5], José del R Millán[3,6,7]

[1]Medical Image Processing Laboratory, Center for Neuroprosthetics, Interschool Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Campus Biotech H4, 1202 Geneva, Switzerland
[2]Nissan International SA, La Pièce 12, 1180 Rolle, Switzerland
[3]École Polytechnique Fédérale de Lausanne (EPFL), Campus Biotech H4, 1202 Geneva, Switzerland
[4]ZHAW Datalab, Zurich University of Applied Sciences, Winterthur, Switzerland
[5] Advanced Materials and Processing Laboratory, Nissan Research Center, Nissan Motors Co. LTD, 1, Natsushima, Yokosuka-shi, Kanagawa-ken, 237-8523, Japan
[6]Dept. of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA
[7]Dept. of Neurology, The University of Texas at Austin, Austin, TX 78712, USA

E-mail: `ruslan.aydarkhanov@alumni.epfl.ch`

Abstract. *Objective.* In contrast to the classical visual BCI paradigms, which adhere to a rigid trial structure and restricted user behavior, EEG-based visual recognition decoding during our daily activities remains challenging. The objective of this study is to explore the feasibility of decoding the EEG signature of visual recognition in experimental conditions promoting our natural ocular behavior when interacting with our dynamic environment. *Approach.* In our experiment, subjects visually search for a target object among suddenly appearing objects in the environment while driving a car-simulator. Given that subjects exhibit an unconstrained overt visual behavior, we based our study on eye fixation-related potentials (EFRP). We report on gaze behavior and single-trial EFRP decoding performance (fixations on visually similar target vs. non-target objects). In addition, we demonstrate the application of our approach in a closed-loop BCI setup. *Main results.* To identify the target out of four symbol types along a road segment, the BCI system integrated decoding probabilities of multiple EFRP and achieved the average online accuracy of 0.37 ± 0.06 (12 subjects), statistically significantly above the chance level. Using the acquired data, we performed a comparative study of classification algorithms (discriminating target vs. non-target) and feature spaces in a simulated online scenario. The EEG approaches yielded similar moderate performances of at most 0.6 AUC, yet statistically significantly above the chance level. In addition, the gaze duration (dwell time) appears to be an additional informative feature in this context. *Significance.* These results show that visual recognition of sudden events can be decoded during active driving. Therefore, this study lays a foundation for assistive and recommender systems based on the driver's brain signals.

Keywords: Brain-computer interfaces, Electroencephalography, Eye tracking, Driving, Visual recognition

1. Introduction

Brain-computer interfaces (BCI), especially those based on non-invasive electroencephalogram (EEG) signals, are not only proving their value as assistive tools for people with disabilities [1, 2, 3, 4, 5, 6] and potential rehabilitation tools for neurological patients [7, 8, 9, 10] but also open opportunities to augment interaction for users without disabilities [11, 12, 13]. In this latter respect, a promising possibility is to decode neural correlates of perceptual and cognitive processes while people overtly interact with real-world environments [14, 15, 16]. In this work, we explore the feasibility of decoding visual recognition from EEG in experimental conditions mimicking daily activities, which involve overt visual search. In particular, we performed a study aiming at EEG decoding of visual recognition during driving. Here the driver is primarily engaged in controlling a car simulator, actively exploring the visual environment to complete a visual search task.

Previous works have reported EEG signatures of visual recognition in simple, well-controlled experimental setups [17, 18, 19, 15]. The classical restrictive conditions require the stimuli to be static and undergo sharp transitions (e.g., flashing) and subjects to sit still in front of a screen [18]. Such experiments often use an oddball paradigm where subjects had to recognize rare target stimuli in a sequence. In these cases the recognition process is reflected in EEG as the well-known P300 component of Event-Related Potential (ERP). P300 is a positive deflection in the parietal region, which occurs typically between 250 and 500 ms after the stimulus onset [20]. Various BCI applications such as P300-based speller have successfully used this signal. With a typing speed of 10 characters per minute [21], such spellers' home use may improve the quality of life of people with severe motor disabilities, such as ALS [22, 23, 24]. For people without motor impairments, however, this setup and typing performance does not bring much value.

Successful decoding of visual recognition in free viewing tasks can provide novel opportunities for BCI application for healthy users. Car drivers are exposed to a richer environment and constantly interact with it by controlling the car, visually exploring the surroundings, and planning upcoming actions. A potential application could use BCI to decode the driver's recognition of road signs of a particular target category, e.g., parking. The driver's interest in parking signs decoded from the EEG will allow to provide timely recommendations for the nearest parking.

Visual search tasks in free viewing conditions require us to fixate static or track moving objects to perceive them in all details. Fixations evoke a type of ERP called Eye Fixation Related Potentials (EFRP) as opposed to external events in more classical ERP designs. Early EEG deflections in EFRP reflect the processing of low-level features of visual input projected from the retina to the occipital lobe and are manifested mainly

in high-amplitude of the P100 component, also referred to as lambda component [25]. When fixating at a target object in free viewing visual search tasks, later EEG deflections in EFRP resemble the P300 component in the oddball paradigm [26, 27, 28].

Visual stimuli used in previous ERP and EFRP studies range from simple geometric shapes in static scenes to static natural images or dynamic scenes with geometric shapes [29, 30]. Only a few attempts on decoding ERP or EFRP in videos or virtual reality (VR) simulations have been reported, e.g., [19, 15]. However, the experimental conditions in those studies did not fully reflect the real-world dynamics. For instance, in human action recognition from a cartoon animation, the video playback was sped up to limit the time of the recognition process [19]. The classification performance of Target vs. Non-Target event reached average AUC > 0.8. In another study involving maze navigation, subjects experienced fast autonomous driving and had to press a button whenever the car in front braked suddenly [15]. The braking task required constant monitoring of the car in front, leading to a brief visual attendance of the stimuli. It resembled the attentional load of real driving, however, unnaturally high speed and lack of full car control limit the transfer of the results to real driving. The classification performance based on a single EFRP was around 0.65 AUC on average across subjects.

In this paper, we report a study on visual recognition decoding during driving in a car simulator based on the Eye Fixation Related Potentials and gaze behavior. In our study, 13 subjects visually attended signboards along the roads while primarily engaged in active driving through an urban environment with a natural speed of $\sim$ 80 km/h. We considered the sudden appearance of the task-relevant visual content (the boards) in the driving environment. Furthermore, the boards were invisible until the driver got close enough, then they popped up at random locations above the sidewalks. Such a stimulus presentation ensured that drivers can recognize the content once they gaze at a board. This setting was chosen to prevent two challenges in decoding visual recognition due to fixations on distant and poorly recognizable boards: 1) to locate the recognition timing within long attendance, 2) to select one of multiple returning fixations when the recognition occurs. The latter impedes the collection of clean data while the former challenges the EFRP decoding itself. Future studies should complement the findings presented here to consider these situations.

The experiment consisted of two phases: offline and online. The goal of the offline phase was two-fold: (1) to analyze ocular and neural correlates of visual recognition and (2) to gather data to train a single-trial classifier that differentiates between attending boards with the target and non-target symbols to be used in the BCI application (the online phase). Then we conducted the online phase, which allows us to examine the single-trial classifier performance in a closed-loop setting and, as a major objective, to verify the corresponding BCI application's feasibility for detecting the drivers' interest along a road segment.

Our protocol closely resembles the real-life scenario of traffic sign recognition on the roads. In the application, the BCI system that decodes the cognitive response on the target sign recognition can be integrated into the advanced driver-assistance system

Figure 1. Visualization of the experimental setup and protocol. Top left: The experimental setup with the driving simulator (car seat, steering wheel, pedals and 3 3D screens), Eye-tracker and subject with EEG cap. Middle left: A screenshot from the beginning of the road with the cue on the floor and the first board on the left-hand side. Bottom left: A screenshot (in the online phase) showing the feedback after the finish line for the correct target symbol detection – a green circle and the 2D board of the inferred target symbol. Right: The schematic drawing of a road segment with the target cue, boards, indications left/right turn at the crossings and finish line. Note that the boards are invisible until they are approached to a close distance.

(ADAS). In this case, the target is naturally chosen by the driver according to the current driving situation. The accurate decoding of the selection of target traffic signs from the driver's brainwaves can shed light on the driver's intentions and goals, from which the ADAS system could derive useful recommendations and assistance.

2. Experimental setup and protocol

2.1. Data collection

Thirteen volunteers (age of 28 ± 6, 3 female, 2 left-handed) with normal or corrected-to-normal vision (8 with contact lenses) participated in the study. One of the participants had a consistently poor eye tracker calibration quality, which led to poor fixation

extraction. This participant was excluded from the analyses. The study was approved by the ethical committees of the Cantons of Vaud and Geneva, Switzerland (Commission cantonale d'éthique de la recherche sur l'être humain, study no. PB_2017-00295), and all the participants provided written consent. The experiment lasted 3 hours, including the set up ($\sim$ 1 h), the offline phase ($\sim$ 45 min), one pause period ($\sim$ 30 min), and the online phase ($\sim$ 45 min). The extended pause period was necessary to process all the data from the offline phase and to train the classifier for the online phase. A run represents a ride through the city. The offline phase consisted of 3 runs, whereas the online phase could comprise 3 to 5 runs depending on the available time and the subject's fatigue. One run included 20 road segments with 12 boards each, resulting in 240 boards per run. Before each run, the subjects were asked to move their eyes up-down and left-right for one minute in order to collect the data for eye movement artifact removal.

The EEG was acquired with the Biosemi ActiveTwo system (Biosemi, the Netherlands) with 64 electrodes at a 2 kHz sampling rate. Additionally, we recorded 3 EOG channels to collect the eye movement data: two electrodes next to the outer canthi of the eyes and one above the nasion. The real-time processing of EEG in the online phase was done on a dedicated computer.

The eye gaze was recorded with the SMI RED (SensoMotoric Instruments, Teltow, Germany) Eye tracking system with a sampling rate of 120 Hz. The chair and eye tracker positions were adjusted for each subject. The eye tracker was calibrated with 13 points once after the EEG setup and before the beginning of the experiment.

The driving simulator logged the car location in the virtual environment, the controllers' state (gas and brake pedals, steering wheel, buttons on the wheel), and the 2D position of boards on the screen at a sampling rate of 256 Hz. In order to synchronize the data acquisition on three separate machines (EEG, eye tracking, and driving simulator) at different sampling rates, a square pulse of 4 Hz was generated by the driving simulator and sent to the eye tracker through TCP connection and to BioSemi through the parallel port.

2.2. Driving simulator

In our study, we used the driving simulator previously used in [12, 11, 31]. It allows for an immersive driving experience by utilizing a real Nissan driving chair with a steering wheel and two pedals (gas and brake). The simulated car has an automatic transmission, which excludes actions for the manual gear shift. The visual input is provided with three 3D monitors, which create multiple renders for different angles. The virtual environment is implemented using the open-source driving simulator project VDrift. The environment resembles a regular grid city with static objects, i.e., buildings, traffic lights, fields. The task-related items include direction indications on the road, target cues, boards with symbols, and finish lines (Figure 1).

2.3. Tasks

The experimental session had 2 phases: offline and online. In both phases, the subject is instructed to drive through the city (without stopping) while following indications left or right turn at the crossings (Figure 1). At the beginning of a road segment, a symbol (the cue) is depicted on the ground. Subjects must memorize and search for it among the boards appearing through the road segment. The data collected in the offline phase allows us to train a single-trial classifier that differentiates between attending the target and non-target symbols. The online phase illustrates how this classifier can be used in the closed-loop BCI application for detecting the driver's symbol of interest within a road segment. In particular, the single-trial classifier is used to decode the driver's brain responses to multiple instances of symbols along a road segment. The BCI system infers the target symbol based on this accumulated evidence and shows it to the driver as feedback at the end of a road segment.

Offline phase. Subjects are asked to search and count the number of boards with the target symbol along a road segment. To keep subjects engaged in the task, they have to report the number of targets at the end of a road segment. After crossing the finish line, the subject presses a button located at the steering wheel as many times as the number of target boards he/she counted along the current road.

Online phase. Subjects are instructed to search for boards with the target symbol and pay attention to the feedback at the end of a road segment – after crossing the finish line, the inferred target symbol is rendered at the bottom of the screen as a 2D object (Figure 1). If the target identification was correct, a green circle is also shown. Otherwise, the circle color is red. The online phase is a closed-loop interaction because the subjects perceive the decoder result shortly after looking at the stimuli, within 15 seconds after the first board on the road.

In both phases, we included empty road segments (without boards and cues) where subjects have no additional tasks except driving, allowing them to rest from the visual recognition task. In total, a quarter of the road segments were empty.

2.4. Stimuli presentation

2.4.1. Stimuli visibility. The board presentation is carefully adjusted to guide the visual behavior of subjects. First of all, boards are invisible unless the driver approaches them close enough to make them pop up. Their positions are generated using the following rules. The boards appear along the road at an equal distance between them, randomly on either side of the road above the sidewalks, with a maximum of 2 boards on the same side in a row (e.g., 2 boards on the left in Figure 1). The number of boards on the left and right sides is balanced. The pop-up distance was greater than the distance between the boards along the road, creating a time overlap in the presence of multiple

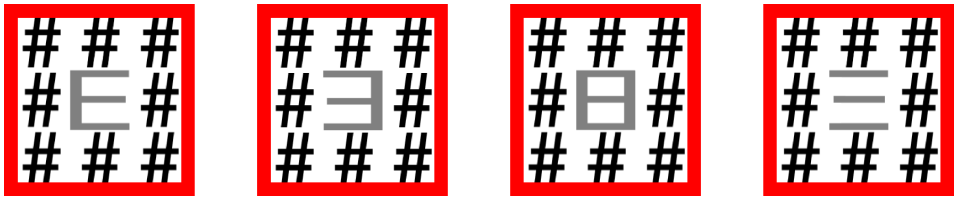

Figure 2. The boards with the used symbols. With one of them being a target symbol per road segment subjects were instructed to recognize and count targets. In the offline phase, we used only boards with symbol E and flipped E whereas in the online phase we used all 4 boards.

boards on the same side. For this reason, the predefined boards' positions (horizontal and vertical coordinates) were chosen to avoid the overlap for the driver view.

The car's maximum speed was limited to 80 km/h to ensure that all the targets could be attended. The subjects were allowed to slow down if necessary to attend all the boards and count the targets. Nonetheless, all the subjects practiced until they felt comfortable completing the recognition task at the maximum speed during the EEG setup. Due to nearly constant speed and regular placement of the boards, there was an interval of about 900 ms between board onsets.

2.4.2. Stimuli content. In order to link the perception of the symbols on the board with the eye fixations, the recognition by peripheral vision must be avoided. Therefore, the target and non-target symbols were similar and surrounded by the # character to create a crowding effect and force the foveating on the main symbol as done in a previous study [32]. Additionally, we added a bright red border around the board, similar to the traffic signs, to create a contrast with the environment and facilitate their identification (Figure 2).

Symbols in the offline phase. In the offline phase, one of the two symbols were depicted on each board: ∃ and ∃. One of them was randomly chosen as a target per road segment and was presented as the cue at the beginning of the road. There were 2-4 targets out of 12 boards on each road segment, with the average proportion of the targets of 0.25 in total.

Symbols in the online phase. In the online scenario that corresponds to the BCI application, we included four different symbols. It is done to mimic the diversity of the natural visual environment (e.g., traffic signs) and to estimate the generalization of the training with the oddball paradigm. Since the BCI application detects the symbol drivers are interested in along a road segment, 4 symbols allow an equal occurrence rate of symbols and equal evidence accumulation, while the target symbol appearance is still a rare event. There were between 2 and 4 boards of each symbol resulting in 12 boards on the road. Only one symbol was a target on each road segment. Therefore,

Table 1. The protocol differences between offline and online phases.

	Offline	Online
Types of boards	2	4
Task	Count silently Button press at the end	Count silently Observe the feedback

the proportion of the targets was 0.25 on average per road section in a run as in the offline phase.

In order to infer the target symbol at the end of the road, each EFRP associated with the symbol attendance is assigned a probability to belong to the target class. At the end of the road segment, the probabilities of all EFRPs are aggregated symbol-wise by averaging. The symbol with the highest average probability of being a target class is selected as an identified target.

3. Methods

The offline and online phases of our experiment have different objectives. As earlier explained, the offline phase is necessary for collecting good quality data for training a classifier that differentiates between attending target and non-target objects to be used in the closed-loop BCI in the online phase. However, it also allows us to study EEG correlates of visual recognition and perform a comparative study of different classification algorithms. The online phase allows the estimation of the closed-loop BCI application performance. In the online phase, we apply a chosen classifier trained on all the offline data.

To make a comparative analysis of different decoding (classifications) approaches under online conditions, we performed a simulated online analysis. Namely, we trained different classifiers on all the offline data and evaluated their performance on the data recorded in the online phase.

The data processing for offline and online phases as well as for simulated online analysis is summarized in Figure 3.

3.1. Fixation extraction and analysis

There exist numerous methods to extract eye movement events from the eye gaze direction [33, 34]. We used a robust and accurate method of Identification by 2-Means Clustering (I2MC) [35] for the offline phase data. In the online phase and simulated online analysis, we used Identification by Dispersion-Threshold (IDT), a simpler and easy-to-implement method for real-time processing [36].

Offline phase. We relied on the I2MC implementation provided by the authors of the method. The main idea behind this method is to find the transition between two consecutive fixations by applying 2-mean clustering in a sliding window manner. During

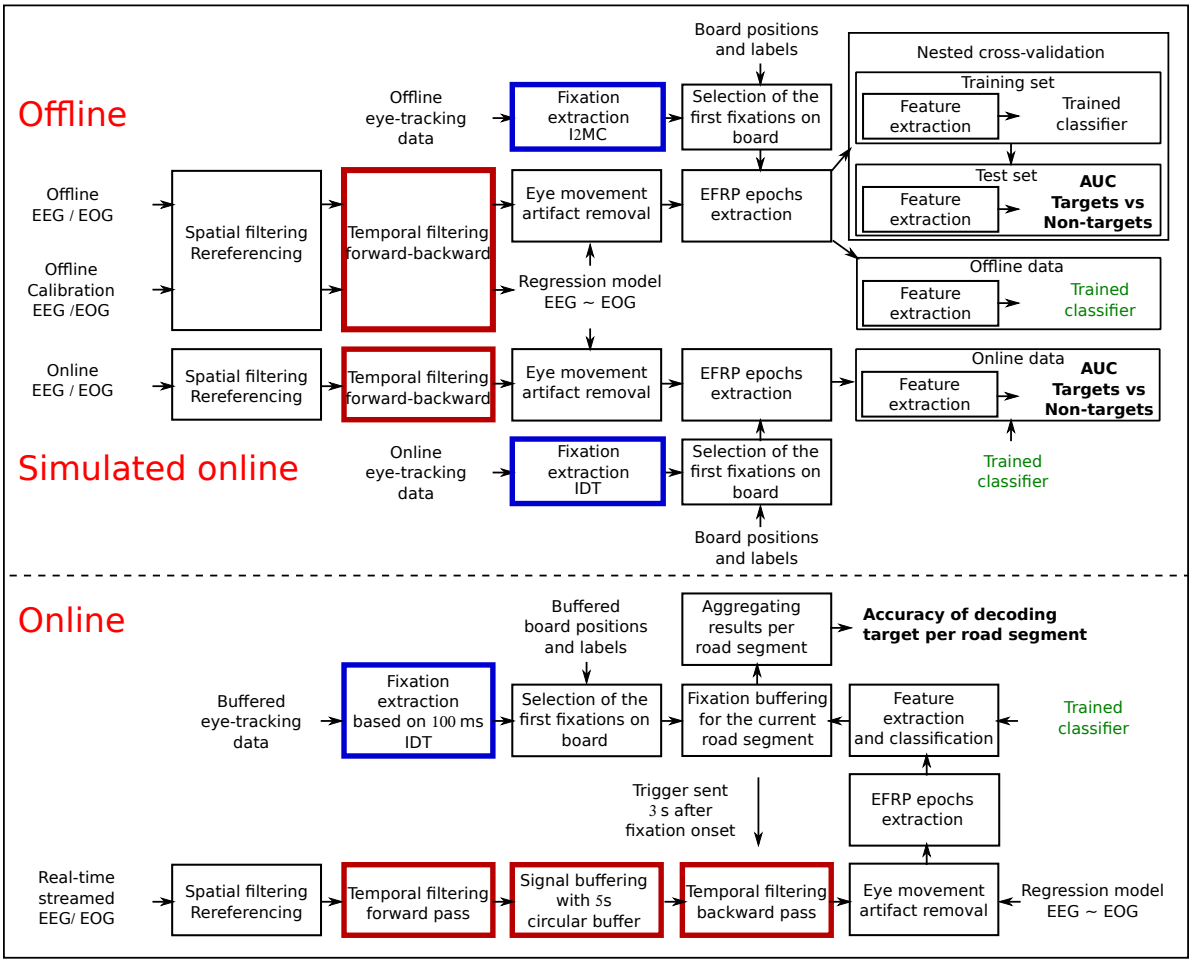

Figure 3. The summary of the data processing in offline and online phases and in simulated online analysis. The boxes represent actions whereas the text without boxes represents data.

a fixation, the eyes do not move, so if we can clearly detect 2 clusters of gaze direction within the sliding window, it means that they correspond to two fixations. This method is more precise and robust to noisy outliers, leading to a higher quality training dataset. We used the values suggested by the authors for most of the numerous parameters. The major changes include the minimum duration of the fixation adjusted to match the value used for the IDT method (100 ms) and a removal of the downsampling step due to a lower sampling rate than in the original work.

Online phase. Since the available implementation of the I2MC method cannot be applied in real time to extract eye fixations, we used the Identification by Dispersion-Threshold (IDT) supplied with our eye-tracking system. Fixation in IDT is extracted when the signals lie within the dispersion thresholds for at least a minimum fixation duration. It requires two parameters: we used 100 ms for the minimum fixation duration and 200 pixels for the maximum dispersion.

Simulated online phase. The same procedure for fixation extraction is applied as in the online phase.

Fixation analysis. The P300-like components of cognitive response are stronger when the stimulus is perceived and recognized for the first time [30]. We assume that subjects categorized the symbol at the first attendance, so we use only the first fixations on the boards for our analysis. As mentioned in Section 2.4, the popping up effect and the boards' design ensure that recognition occurs after fixating on the symbol.

The visual input during the task is dynamic. Due to driving through the virtual environment, the objects, including the boards, are also moving on the screen. To judge whether the board contains the target symbol, the subjects had to fixate it and follow it by gaze until they made a decision. So we assume that most of the board attendances are done with smooth pursuit rather than fixations. However, the onset of the first fixation on a board will coincide with the onset of smooth pursuit. To the best of our knowledge, there is no available algorithm for efficient extraction of smooth pursuit for eye movement data sampled at 120 Hz (the upper sampling rate limit of the eye-tracker model that we used). The only consequence of extracting fixation from smooth pursuit is that a single smooth pursuit may be oversegmented into multiple fixations. For the sake of our analysis, we do not need to differentiate between fixations and smooth pursuit movements. The onset of the first fixation on a board will coincide with the onset of smooth pursuit.

Each detected fixation was assigned to one of the three categories: a target board fixation, a non-target board fixation, or a non-board fixation. Considering that (i) the visual span (the angular span) within which our vision is sharp enough to perform a given task extends up to several degrees ($\sim 3-10^\circ$ in central vision) [37], (ii) the driver's motion relative to the boards imposes visual tracking to perform the search task, and (iii) the noise in eye-tracking (sudden and brief disruptions spanning several data samples are occasionally observed in the eye-tracking data), we applied the following approach to assigning the boards to fixations. For each eye gaze sample, we estimate the probability of fixating eyes on the board's center according to a normal distribution. After averaging log-probabilities across the fixation time window, we apply a hard threshold to assign the fixation to a board or a non-board category. This procedure was implemented in the analysis of the data. The real-time assignment of fixations on boards in the online phase was modified for computational efficiency. It was based on the first 100 ms after fixation onset and used on the average fixation direction within 100 ms.

For the gaze analysis, we estimate the dwell time – the uninterrupted time which the driver spent looking at a board. The dwells were created by merging all the fixations on the same board with the inter-fixation interval between them below 50 ms. We used the IDT method to extract fixations for the sake of a valid comparison between offline and online phases in gaze analysis.

3.2. EEG processing

EEG processing of offline data. All the EEG and EOG were downsampled from 2 kHz to 256 Hz using FIR antialiasing lowpass filter and further filtered with a Butterworth band-pass filter of order 4 within the band [1, 10] Hz in a forward-backward way (to eliminate phase distortion). The filtering range is characteristic for the P300 component [38]. Due to the low conductivity of the skull and the skin, the EEG signal is spatially smoothed, so a high contrast between nearby channels results from noise and movement artifacts. We remove this noise by keeping only low spatial frequency components after decomposition EEG with SPHARA [39]. Horizontal and vertical components of eye movement were estimated, which allowed removing the eye movement artifacts from EEG using multiple regression as in [40] using only horizontal and vertical components as well as the intercept. The coefficients of multiple regression were estimated from the one-minute session of eye movements before the corresponding run. Then the signal is spatially filtered with common-average-reference (CAR).

EEG processing in the simulated online analysis. The EEG was band-pass filtered by the same filters as in the offline phase. Besides, the eye movement artifacts removal is done using the multiple regression model obtained from the calibration data recorded in the offline phase.

EEG processing in the online analysis. The closed-loop interaction in the online phase required real-time processing. SMI system provides real-time eye fixation detection through the IDT method. The fixations were buffered by a parallel process within the driving simulator, matched with the boards, and a trigger was sent to the BioSemi system 3 s after the onset of each fixation on a board. We chose 3 s delay because we applied a non-causal filter on EEG data and needed to diminish the edge effect of the backward filter pass. EEG processing was identical to the offline procedure except for two steps:

- the spectral filtering was done on a 5 s buffer of data, around [-2, 3] s in reference to the fixation onset;
- the eye movement artifacts were removed based on the multiple regression coefficients trained with the eye movement data from the offline phase.

3.3. Decoding approaches

To decode visual target recognition from EEG signals, we solve the binary classification problem to classify EFRP epochs into Target vs. Non-target. We extracted the EEG epochs around the fixation onset – more precisely, the EEG activity within the time interval from 200 ms to 1000ms following the fixation onset, which is equal to 205 time points. Given the 64 channel montage, a single EEG epoch has a dimensionality of $\mathbf{x} \in \mathbb{R}^{205\times 64}$ or $\mathbf{x} \in \mathbb{R}^{13120}$ after vectorization. In addition, we extracted dwell time as

eye-gaze based features to detect visual recognition (see Section 3.1). We investigate and compare different decoding approaches, which consist of various feature sets and classifiers.

Offline and simulated online decoding. The following combination of features and classifiers were applied in the offline and simulated online decoding:

- **PLR-Waveform**. Penalized logistic regression (PLR) trained on waveform features (i.e., the signal's value at each channel and time point) after reducing the dimensionality with PCA. Only the components which explain 90% of the variance are kept (transformed feature vector $\mathbf{z} \in \mathbb{R}^{86\pm11}$).
- **PLR-Dwell**. PLR trained on dwell time on the boards ($\mathbf{z} \in \mathbb{R}^{1}$).
- **PLR-Combined**. PLR trained on the combination of waveform features with dwell time. We concatenate the dwell time and EEG waveform features after applying PCA to keep 95% of variance ($\mathbf{z} \in \mathbb{R}^{87\pm11}$).
- **RF-Waveform**. Random forest trained on waveform features. We use 100 decision trees and with a maximum depth of 5 ($\mathbf{z} \in \mathbb{R}^{13120}$).
- **PLR-Riemann**. PLR trained on Riemannian features from simple EEG epochs. To build Riemannian features, we estimate a spatial covariance matrix with shrinkage and project it to the tangent space according to the classical Riemannian geometry on SPD matrices [41]. We subselected 8 channels based on the mean Fisher score across the epoch ($\mathbf{z} \in \mathbb{R}^{36}$).
- **PLR-Riemann+**. PLR trained on Riemannian features from augmented epochs. Before computing the covariance matrix, we augment the epoch with the averaged ERP for each class (target and non-target). Otherwise, it is identical to the previous approach ($\mathbf{z} \in \mathbb{R}^{300}$).

Since PLR is a linear regularized classifier, we standardize all the features to z-score when using PLR. For PLR, we applied elastic net regularization with a major contribution from ℓ^2-norm (1000 times greater than ℓ^1-norm; $\lambda_2 >> \lambda_1$).

Online decoding. The preliminary tests on two subjects showed that the RF-Waveform classifier performed better than the PLR-Waveform classifier. We did not consider the PLR-Riemann classifier at this stage because its training was too slow to apply between offline and online phases. Therefore, we trained an RF-Waveform classifier for each subject individually on the data obtained in the offline phase and applied it in real time in the online phase. The probability of being a target was estimated on each instance of every symbol along a road segment and sent back to the driving simulator. Based on the probabilities averaged over instances for each of four symbols separately, the BCI system will infer the target symbol at the end of a road segment.

3.4. Performance evaluation

Offline performance evaluation. We employ nested cross-validation to adjust various hyperparameters in the inner loop: regularization term for PLR and the tree depth in Random Forest. The purpose of the outer loop is to obtain an unbiased performance estimation, so it is critical to avoid training and testing on correlated data. We achieve it by performing leave-one-run-out for the outer loop, although we had only 3 offline runs. The inner loop is implemented with 4-fold cross-validation while keeping the temporal order of the trials before the split. Since the classes of target and non-target eye fixations are unbalanced, we utilized AUC to measure the classification performance. The group-level average performance was tested with a non-parametric one-sided one-sample Wilcoxon signed-rank test against the theoretical chance level of 0.5 AUC for each classifier. The p-values are corrected with the Benjamini-Hochberg correction [42]. The differences between classifiers are estimated with a non-parametric Friedman test.

Simulated online performance evaluation. After training the classifiers on all the offline data, we applied them to all the online data and obtained a single AUC value. The applied statistical procedure is the same as for the offline phase data.

Online performance evaluation. During the online phase, we predicted the target symbol from the EFRP classification. We assess the overall performance with accuracy and confusion matrices for 4 symbols. Accuracies were estimated per subject and averaged across subjects. The significance of single-subject performance was tested with the exact binomial test against 0.25 (the theoretical chance level with 4 balanced classes) [43]. Due to the high number of tests, we applied the Benjamini-Hochberg correction [42]. The confusion matrices were tested for independence with Pearson's chi-square test [44] separately for each subject with Benjamini-Hochberg correction.

As for the level of statistical significance, we select the standard level of $p = 0.05$.

4. Results

4.1. Gaze analysis

Comparison of fixation extraction methods I2MC vs. IDT We used two different methods to extract fixations in the offline and online phase of the experiment. We compared them using the offline phase data via two measures: 1) the average number of all extracted dwells on boards per run and 2) the average dwell time (Table 2). The number of dwells obtained with IDT method is 10% smaller (260 vs. 284), which is a statistically significant difference ($p < 0.0001$ with one-tailed t-test). Since we used the IDT method in the online phase, we missed 10% of the fixations, which led to either missed attendance or labeling some of the second fixations as first fixations. The dwell duration is slightly longer with the IDT method (433 vs. 408 ms), and it is not significantly different from the I2MC method ($p = 0.06$ with one-tailed t-test).

Table 2. The average number of dwells on boards and the average dwell time per run based on fixation extraction with I2MC and IDT methods in the offline phase. The numbers represent the average and standard deviation across subjects.

	# of dwells	Dwell time [ms]
I2MC	284 ± 59	408 ± 54
IDT	260 ± 64	433 ± 86

Table 3. The board attendance rate for targets and non-targets in the offline and online phases with IDT method of fixation extraction. The values represent the average and standard deviation across 12 subjects. The values in parentheses represent the number of boards per subject.

	Offline	Online
Targets	0.92 ± 0.07 (173)	0.47 ± 0.06 (213 ± 39)
Non-Targets	0.92 ± 0.07 (547)	0.44 ± 0.09 (635 ± 120)

Board attendance For the analysis of attendance rate, we used the IDT method to extract fixations for the data from both offline and online phases. In the offline phase, subjects visually attended most of the boards. The attendance rate is almost halved in the online phase. The average attendance rate is shown in Table 3. Two-way repeated measures ANOVA shows a minor significant effect in the interaction (experimental phase x board type) with $F_{(1,11)} = 8.927$, $p = 0.012$. The main effect of board type (targets vs. non-target) is also minor but significant with $F_{(1,11)} = 4.923$, $p = 0.0485$. However, the experimental phase (offline vs. online) shows a major and significant effect with $F_{(1,11)} = 405$, $p < 0.0001$. Since the effect of board type is statistically significant, we can assume that subjects could sometimes differentiate between targets and non-targets without directly looking at the board. However, due to the small difference in the attendance rate, this effect is negligible in the context of our study.

Counting The total number of targets in the offline phase is 173. We analyzed the button presses after each road segment, which should be equal to the number of targets on each road. The average number of incorrect counts (both missed and extra counts) was 5, the worst performance was at 15 errors (see the number of mis-counts in Figure 7).

Dwell time We analyzed the dwell time distributions on targets vs. non-targets in the offline and online phases (Figure 4). Most of the dwells are limited to the time between the boards pop up equal to 900 ms. The dwell times are identical for non-targets in both phases and statistically significantly shorter than for targets ($p < 0.0001$). The median dwell time for targets is statistically significantly longer in the online phase ($p < 0.0001$).

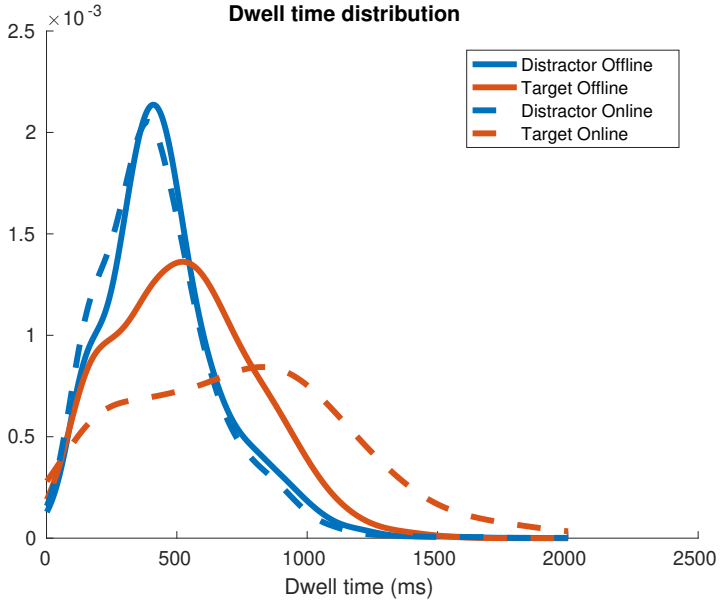

Figure 4. Dwell time distribution for first dwells on targets vs. non-targets in the offline and online phases. The single-subject distributions were smoothed, normalized and averaged across subjects.

4.2. EFRP waveform

We present the analysis of the aggregated EFRP waveform for the 4 subjects (S1-S4) who demonstrated the highest classification performance in the offline phase with the PLR-Waveform classifier because it allows for direct representation of the features used by a linear classifier (see Section 4.3, Figure 7).

We visualized the representative Cz channel signal for each eye fixation while ordering them by the dwell time (Figure 5). The amplitude of the presented EFRP is limited to the range [-1.5, 1.5] uV. The complex of components right after the fixation onset ranging from 100 to 300 ms reflects the evoked potentials from the fixation itself. It contains negative and positive deflections. We can observe the same complex of components after the dwell offset. The shift of gaze happens right where we expect the P300-like component, so it can be masked by this evoked activity. The positive deflection occurs at the end of the dwell for both targets and non-targets, however, it has a greater amplitude for targets as seen on the averaged EFRP.

The spatio-temporal distribution of discriminant power is visualized with signed R-squared (r^2) value obtained on aggregated epochs of the 4 best performing subjects (Figure 6). The results are qualitatively similar between the offline and online phases. The greatest values are mainly confined within the region between 100 and 700 ms. There is a spatio-temporal evolution of discriminant activity spreading across the whole scalp. It is expected as the visual processing involves different brain regions. It is important to emphasize that the interpretation of discriminant activity's spatial distribution depends on the chosen referencing method (CAR in our case). Nonetheless, the P300-like component of EFRP can be seen at 500 ms after the fixation onset.

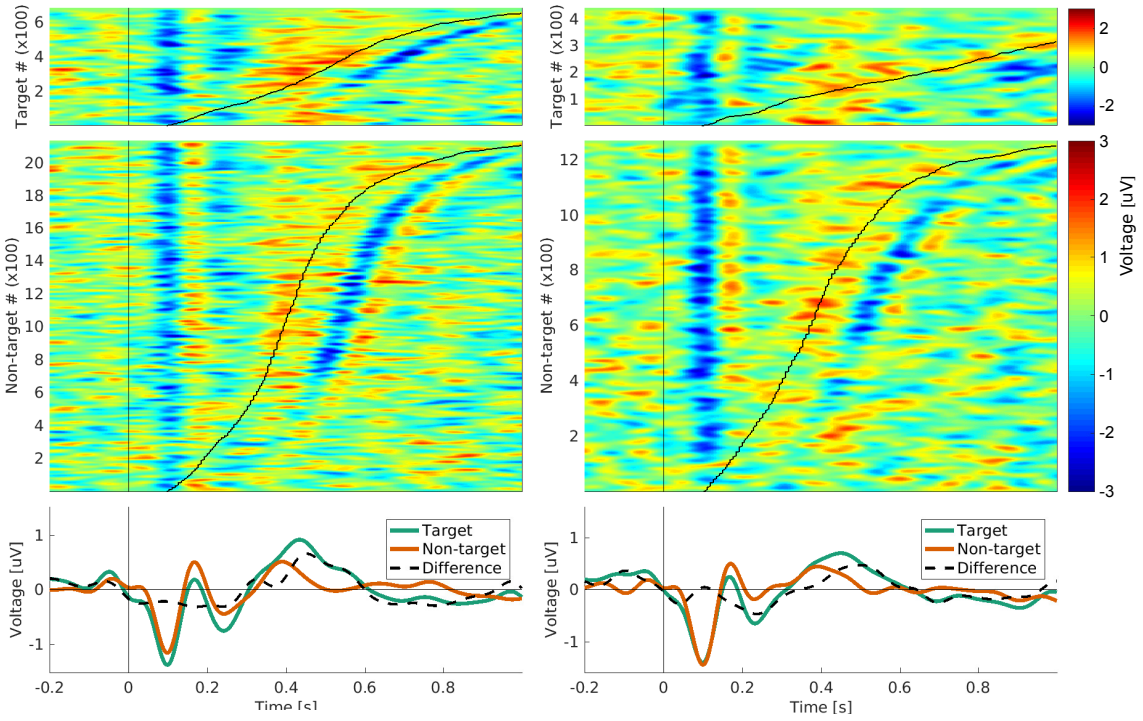

Figure 5. The signal of Cz channel for extracted EFRP epochs aggregated for the 4 best performing subjects with the offline AUC above 0.6 with the PLR-Waveform approach. Left: offline phase. Right: online phase. Top panel: target EFRP epochs. Middle panel: non-target EFRP epochs. Bottom line plot: average signal across target and non-target EFRP epochs. The epochs are ordered according to the dwell time shown with the black S-shaped curve.

4.3. Comparison of decoding approaches

Offline All the classification methods yielded a single-trial performance between 0.52 and 0.60 AUC on average (Figure 7), which is statistically significant against the chance level of 0.5 for all methods ($p < 0.05$ with Wilcoxon tests after Benjamini-Hochberg correction for 6 methods). Five out of twelve subjects achieve performance above 0.6 for at least one of the approaches based on EEG features. The performance of different approaches has a different ranking per subject, and the differences between the approaches are statistically significant ($\chi^2_5 = 6.5$, $p = 0.0001$ with repeated measures ANOVA). It is worth noticing that the combination of both dwell time and EEG waveform is better than just one of these feature sets with the post hoc one-sided Wilcoxon signed-rank providing $p = 0.04$. The *Riemann* feature set has poorer performance than the *Riemann+* feature set with the post hoc one-sided Wilcoxon signed-rank providing $p = 0.001$.

Simulated online The average AUC values lie between 0.55 and 0.73 for each approach, which is statistically significant against 0.5 for all approaches. ($p < 0.05$ with Wilcoxon

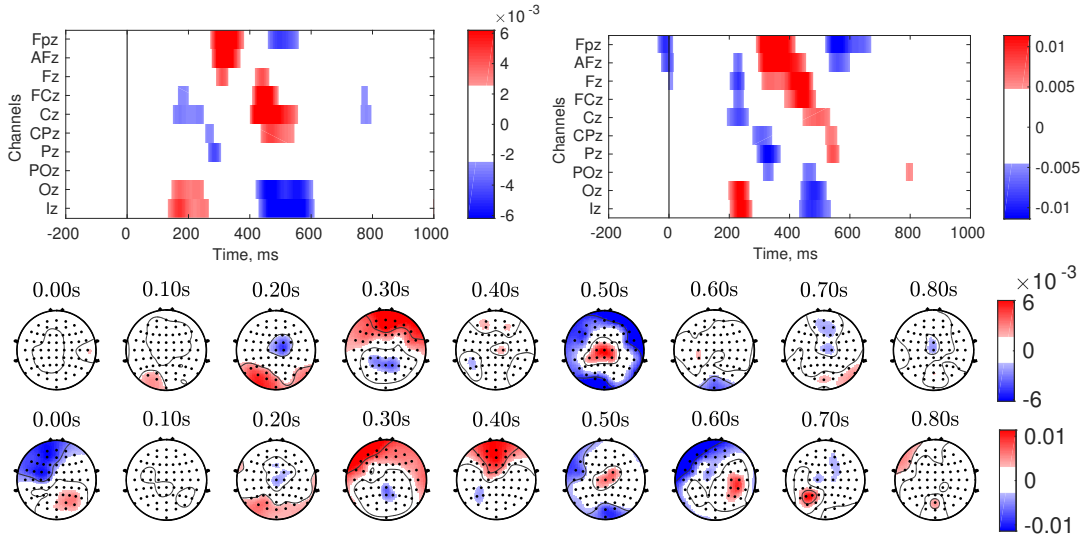

Figure 6. The discriminant power for the aggregated epochs of the 4 best performing subjects with the offline AUC above 0.6 with the PLR-Waveform approach. Signed R^2 is demonstrated for midline channels around the eye fixation onset (top left: offline, top right: online) and on topographic maps (top map: offline, bottom map: online).

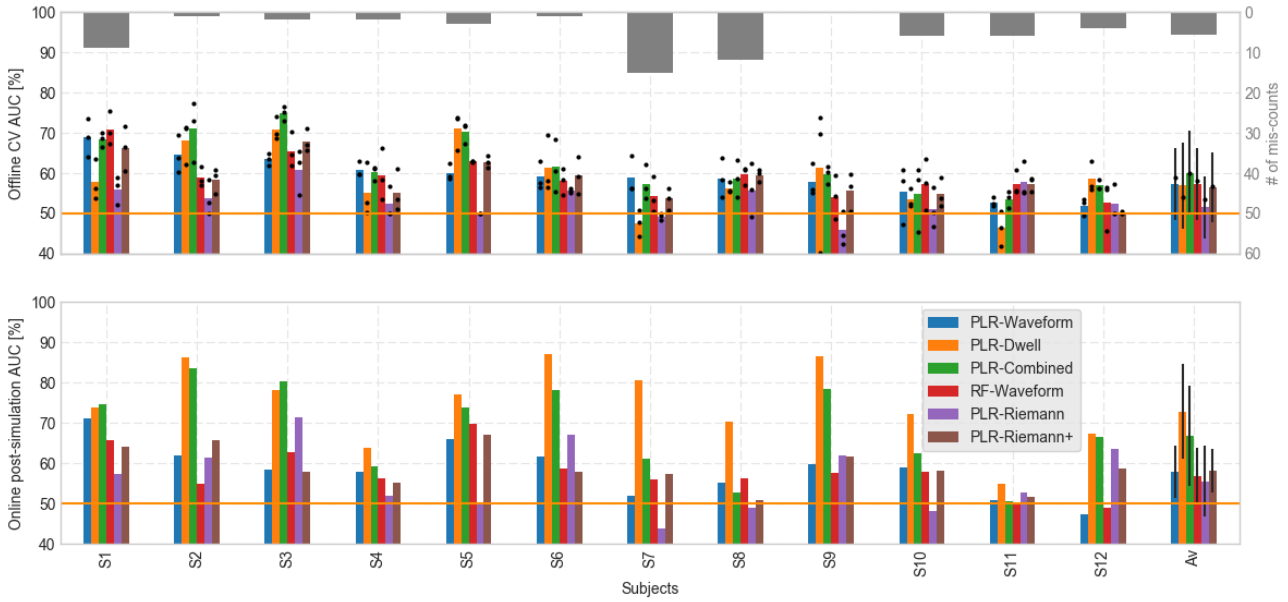

Figure 7. Performance of EFRP classification with various approaches in the offline analysis (top) and the simulated online analysis (bottom) as a percentage of AUC. Each dot shows single fold performance in a leave-one-run-out cross-validation for the corresponding classification approach. The overlaid gray bars at the top show the behavioral performance: the number of mis-counts in the offline phase with the respect to the right-hand y-axis.

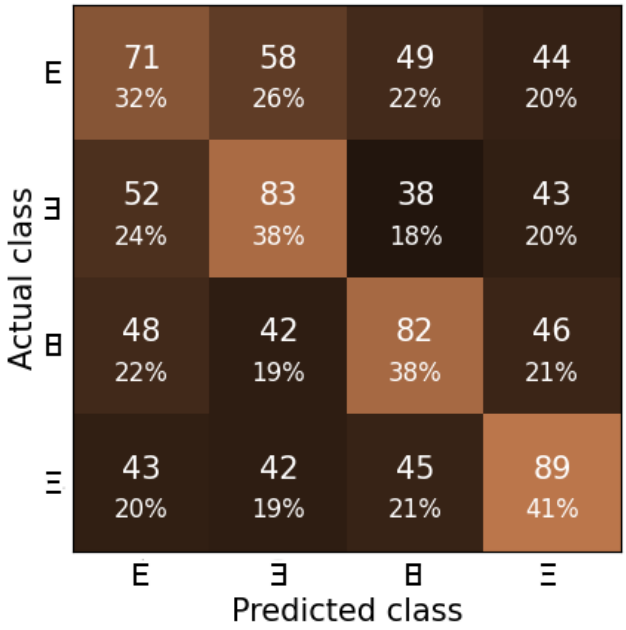

Figure 8. The aggregated confusion matrix of online decoding for each type of symbol being a target class. The absolute number of cases is presented and normalized as a percentage relative to the actual class (row-wise normalization).

tests after Benjamini-Hochberg correction for 6 approaches).

It is worth noting that the performance of EEG-based approaches on online data is consistent with the training performance on offline data. However, for approaches relying on the dwell time, the performance drastically improved for several subjects compared to offline performance. The average AUC for the *Dwell* classifier increased from 57 to 73 ($p = 0.0002$ with one-sided Wilcoxon signed-rank) and for the *PLR-Combined* classifier from 60 to 67 ($p = 0.008$ with one-sided Wilcoxon signed-rank). This improvement is a direct consequence of the increased target dwell time together while non-target dwell time remained unchanged, see Figure 4. However, the underlying question of why the dwell time changed between the offline and the online phases remains.

4.4. Online performance

We assessed each subject's task performance as the accuracy of detecting the target symbol at the end of a road segment based on the evidence obtained by applying an RF-Waveform classifier to decode driver's brain responses to multiple instances of symbols along a road segment (Table 4). The averaged accuracy equals 0.37, and it is statistically significantly above the chance level of 0.25 for a multi-class classification with 4 balanced classes ($p < 0.002$ with one-sided one-sample Wilcoxon signed-rank test). Additionally, we applied a statistical test to assess the accuracy per subject. After adjusting p-values with the Benjamini-Hochberg procedure, 8 out of 12 subjects performed statistically above chance level.

Table 4. Summary of online performance: accuracy of target decoding per road segment and p-values for two statistical tests. The accuracy is tested with binomial distribution. The recognition of a target symbol independent of the actual symbol is tested with a chi-square test of independence. The p-values are corrected with Benjamini-Hochberg method per row. Statistically significant results are typed in bold.

Subjects	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Accuracy	0.44	0.38	0.45	0.29	0.44	0.35	0.39	0.25	0.31	0.38	0.38	0.4
Accuracy test	**0.001**	**0.04**	**0.001**	0.48	**0.001**	0.07	**0.02**	1.0	0.24	**0.04**	**0.04**	**0.03**
Independence test	**0.004**	0.29	**0.03**	0.74	0.08	0.17	0.1	0.41	0.48	0.19	0.22	0.26

To verify the independence of the 4 classes in the online phase, we computed the aggregated confusion matrices across all symbols for all subjects (Figure 8). We performed an independence test for confusion matrices per subject. After adjusting p-values with the Benjamini-Hochberg procedure, only two subjects show a significant imbalance in the confusion matrix. For these two subjects, the accuracy of the symbol Ξ was higher, and the accuracy of the symbol E was lower than of the other symbols.

5. Discussion

The integration of the BCI systems in the daily life of healthy/able-bodied users requires the system to be built around the experimental paradigms supporting natural human behavior. To this end, EFRP-based decoding of cognitive processes in overt visual search has the potential to augment human-machine-interaction. In this study, we investigated the decoding of visual recognition in a driving scenario. It resembles one of the typical everyday activities and provides the associated challenges in decoding visual recognition: free eye gazes, dynamic visual input, primary tasks. For this purpose, we limit the driving task to following the simple route at a comfortable and natural speed. To avoid overloading subjects' attention with too many distractors, we did not include other participants on the road nor moving objects. Nonetheless, due to the car's movement, the drivers were subjected to a dynamic visual input and perceived a natural optic flow.

Ocular behavior in driving. We considered the random pop-up appearance of the task-relevant stimuli above the sidewalks. In the previous study on active search in a dynamic scene [29], this type of appearance created sufficient locking of the cognitive EFRP's components to the fixation onset. At the same time, their results showed that the time spent on the stimuli was not informative about the stimuli type (target vs. non-target) in contrast to the more attentionally demanding motion appearance conditions. Interestingly, we observed the time spent on the stimuli to be discriminative in our experimental scenario, although the same pop-up stimuli appearance was considered. It can be explained by the effect of the user's/driver's motion relative to the objects on

the visual information processing and the attentional load. Moreover, in the online phase, subjects looked longer on targets compared to the offline phase. Although the subjects' dwell time was not decoded directly, they were aware that they could potentially influence the decoding quality. It might lead to deliberate or unconscious changes in their behavior. Some subjects could achieve a high decoding performance based only on dwell time in the offline phase. But with the changes in the behavior during the online phase, most of the subjects drastically improve in their decoding performance based on the dwell time.

The board attendance rate reveals two aspects of the task completion: (1) how well the subjects coped with the pace and attentional load and (2) whether they can differentiate between targets and non-targets using their peripheral vision. We observed a statistically significant difference in the board attendance between offline and online phases despite the equal number of boards. On the one hand, we expected the subjects to be more engaged in the task due to the interactive feedback part. However, the observed increase of the dwell time on target boards in the online phase made it more challenging to attend all the boards within the limited time. On the other hand, behavioral response based on the number of recognized targets was not required, which might relax the cognitive load.

We carefully designed stimuli to ensure that their recognition requires foveal vision. It is confirmed by the similar attendance rate on targets and non-targets. However, the difference is statistically significant, suggesting that peripheral recognition is not completely excluded with the provided stimuli design. We were concerned that introducing new symbols in the online phase may change the behavior or the cognitive response due to the novelty. The balanced confusion matrices of decoding target recognition in the online phase confirm that subjects perceived all the symbols equally with regard to the task. However, the presence of additional types of stimuli may have contributed to the increased dwell time on the target boards because the target identification among 4 symbols closely resembling each other is more challenging than among only 2 symbols [45, 46].

Decoding visual recognition in driving. The discriminant analysis of EFRP shows similar results for both offline and online phases. Most of the relevant features lie within the [200, 700] ms window, which coincides with the dwell durations. The spatial localization of relevant features is consistent with the typical spatial distribution of the P300 component in the oddball paradigm. The EFRP waveforms are known to contain a strong P1 component at the occipital area that reflects the beginning of the visual processing of a stable visual input after the saccade. In the analysis of the Cz channel, it corresponds to the negative deflection at 100 ms. It is clearly present in most of the first fixations on boards. Moreover, we could also see it for the following fixations. During the dwell time, multiple fixations can be detected, leading to the overlap of the P300-like component locked to the first fixation and the following early EFRP component. In contrast to the classical P300 paradigms where the the visual input timing is controlled

(with a constant frequency), here, the overlap of EFRP is variable across trials. It is an additional challenge as compared to typical event-related potential detection approaches. Removing the activity related to previous and subsequent fixations from the EEG was attempted by modeling it from various characteristics of the previous and subsequent fixations. [30, 28, 47]. We did not choose to apply this approach as there is a risk of distorting the signal of interest. The precise EFRP shape depends on preceding and following eye behavior (e.g., amplitude and direction of the saccades) as well as low-level features of the visual input [48].

We compared multiple EFRP-based classification approaches on the offline data. In addition, we did a comparative study using recorded online data (simulated online analysis). All approaches, including waveform-based linear and non-linear and covariance-based methods, resulted in a similar performance on average across subjects, which is significantly above the chance level for all the approaches. The combination of waveform features with the dwell time outperforms other feature sets on average. However, there is no single best approach for all subjects.

The actual online closed-loop performance measured by the target symbol identification (a four-class problem) is significantly above the chance level for eight subjects. Moreover, on average across subjects, the accuracy of 0.37 is statistically significantly higher than the chance level.

Considering the observed dwell-time difference between target and non-target boards, one can argue that the relatively good performance of EFRP classifiers is due to the early evoked components of the following fixations rather than due to the later cognitive components of the analyzed EFRP. Therefore, these classifiers may indirectly rely on the difference between the target and non-targets dwell times. However, this was not the case as the simulated online performance of the EEG-based decoders did not exhibit an improvement similar to the classification models based on dwell time.

A critical part of real-time systems such as the one implemented in our experiments is the computational cost of the used algorithms. The approaches used in the online phase proved to be time-efficient without causing delays. Eye fixation extraction was successfully done with IDT provided by SMI. I2MC method requires more computation due to a clustering step on a sliding window manner, however, it might also be adjusted for a real-time application. Spatial and temporal filtering, as well as eye movement artifact removal, are fast linear operations. Regarding feature extraction and inference of the classifier, these steps are the least demanding in computational time because they are applied only as frequently as the relevant eye fixations are detected. All mentioned pairs of features and classifiers were applicable in the current setup, therefore, it did not affect the choice of the approach in the online phase. Nonetheless, the training even on a single subject data required a considerable amount of time, especially when applied multiple times within nested cross-validation. The most demanding steps included the estimation of parameters for the Riemannian features and the Random Forest classifier. This led to an extended pause between the offline and online phases of the experimental session.

One of the challenges in building a real-world BCI application is an adequate training protocol design. The training protocol has to secure good quality data for training a classifier. Thus, the practice is to create highly controlled trials in terms of stimuli and the subjects' behavior. At the same time, the introduced constraints should not compromise the data not to be representative of the actual phenomenon in real-world conditions when operating the BCI application. It is known that stimuli – particularly their semantic richness, ambiguity, and diversity – may affect visual recognition and the corresponding ERP components. We believe that the approach to address this challenge should include two actions: (i) to create a training protocol as close as possible to the real-world setup, and (ii) to consider decoding algorithms able to deal with variability in neural responses [49]. We plan to further investigate in this direction and encourage the other researchers in this field to address this challenge.

Finally, in this study, our focus was on detecting the EEG correlates of object recognition in free viewing visual search tasks during driving. Ocular correlates such as pupil dilation along with multimodal integration of neural and ocular correlates may be further investigated in future studies. Multimodal integration may improve decoding performance [15].

6. Conclusion

In this study, we have demonstrated the feasibility of decoding visual target recognition from brain signals in challenging conditions of a naturalistic driving scenario. By decoding the brain signals, the smart car can acquire information about the objects of drivers' interest and enhance their interaction by creating a tailored recommendation. Considering that no extra effort is required from the driver and that he could quickly accept or ignore the recommendations, the extremely high decoding performance is not a prerequisite. Our results are, therefore, promising for bringing this type of BCIs to smart cars. Finally, our approach can benefit from improvements in different components, such as decoders robust to the natural temporal variability of EEG potentials, reliable and computationally efficient eye fixation algorithms, longer training of subjects over multiple sessions, and multimodal integration of neural and ocular data. Bigger datasets would also allow the application of more powerful machine learning models, e.g., deep neural networks.

Acknowledgments

The authors thank Nissan Motor Co. Ltd for supporting this work.

References

[1] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398(6725):297–298, March 1999.

[2] J. R. Wolpaw and D. J. McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51):17849–17854, December 2004.

[3] E. M. Holz, J. Höhne, P. Staiger-Sälzer, M. Tangermann, and A. Kübler. Brain–computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users. *Artificial Intelligence in Medicine*, 59(2):111–120, October 2013.

[4] R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson, and J. d. R. Millán. Towards independence: A BCI telepresence robot for people with severe motor disabilities. *Proceedings of the IEEE*, 103(6):969–982, June 2015.

[5] S. Saeedi, R. Chavarriaga, and J. d. R. Millán. Long-term stable control of motor-imagery BCI by a locked-in user through adaptive assistance. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(4):380–391, April 2017.

[6] S. Perdikis, L. Tonin, S. Saeedi, C. Schneider, and J. d. R. Millán. The Cybathlon BCI race: Successful longitudinal mutual learning with two tetraplegic users. *PLOS Biology*, 16(5):e2003787, May 2018.

[7] A. Ramos-Murguialday, D. Broetz, M. Rea, L. Läer, Ö. Yilmaz, F. L. Brasil, G. Liberati, M. R. Curado, E. Garcia-Cossio, A. Vyziotis, W. Cho, M. Agostini, E. Soares, S. Soekadar, A. Caria, L. G. Cohen, and N. Birbaumer. Brain-machine interface in chronic stroke rehabilitation: A controlled study. *Annals of Neurology*, 74(1):100–108, July 2013.

[8] F. Pichiorri, G. Morone, M. Petti, J. Toppi, I. Pisotta, M. Molinari, S. Paolucci, M. Inghilleri, L. Astolfi, F. Cincotti, and D. Mattia. Brain–computer interface boosts motor imagery practice during stroke recovery. *Annals of Neurology*, 77(5):851–865, 2015.

[9] A. Biasiucci, R. Leeb, I. Iturrate, S. Perdikis, A. Al-Khodairy, T. Corbet, A. Schnider, T. Schmidlin, H. Zhang, M. Bassolino, D. Viceic, P. Vuadens, A. G. Guggisberg, and J. d. R. Millán. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. *Nature Communications*, 9(1):1–13, June 2018.

[10] M. A. Cervera, S. R. Soekadar, J. Ushiba, J. d. R. Millán, M. Liu, N. Birbaumer, and G. Garipelli. Brain-computer interfaces for post-stroke motor rehabilitation: A meta-analysis. *Annals of Clinical and Translational Neurology*, 5(5):651–663, March 2018.

[11] H. Zhang, R. Chavarriaga, Z. Khaliliardali, L. Gheorghe, I. Iturrate, and J. d. R. Millán. EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering*, 12(6):066028, 2015.

[12] Z. Khaliliardali, R. Chavarriaga, L. A. Gheorghe, and J. d. R. Millán. Action prediction based on anticipatory brain potentials during simulated driving. *Journal of Neural Engineering*, 12(6):066006, 2015.

[13] R. Chavarriaga, M. Uščumlić, H. Zhang, Z. Khaliliardali, R. Aydarkhanov, S. Saeedi, L. Gheorghe, and J. d. R. Millán. Decoding neural correlates of cognitive states to enhance driving experience. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(4):288–297, August 2018.

[14] M. Uščumlić, R. Chavarriaga, and J. d. R. Millán. An iterative framework for EEG-based image search: Robust retrieval with weak classifiers. *PLOS ONE*, 8(8):e72018, August 2013.

[15] D. C. Jangraw, J. Wang, B. J. Lance, S.-F. Chang, and P. Sajda. Neurally and ocularly informed graph-based models for searching 3D environments. *Journal of Neural Engineering*, 11(4):046003, 2014.

[16] S. Haufe, J.-W. Kim, I.-H. Kim, A. Sonnleitner, M. Schrauf, G. Curio, and B. Blankertz. Electrophysiology-based detection of emergency braking intention in real-world driving. *Journal of Neural Engineering*, 11(5):056011, August 2014.

[17] A. D. Gerson, L. C. Parra, and P. Sajda. Cortically coupled computer vision for rapid image search. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):174–179, June 2006.

[18] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. Toward

enhanced P300 speller performance. *Journal of Neuroscience Methods*, 167(1):15–21, January 2008.

[19] D. Rosenthal, P. DeGuzman, L. C. Parra, and P. Sajda. Evoked neural responses to events in video. *IEEE Journal of Selected Topics in Signal Processing*, 8(3):358–365, June 2014.

[20] J. Polich. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10):2128–2148, October 2007.

[21] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak. Brain–computer interface spellers: A review. *Brain Sciences*, 8(4), March 2018.

[22] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw. A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis*, 11(5):449–455, October 2010.

[23] E. M. Holz, L. Botrel, T. Kaufmann, and A. Kübler. Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: A case study. *Archives of Physical Medicine and Rehabilitation*, 96(3, Supplement):S16–S26, March 2015.

[24] J. R. Wolpaw, R. S. Bedlack, D. J. Reda, R. J. Ringer, P. G. Banks, T. M. Vaughan, S. M. Heckman, L. M. McCane, C. S. Carmack, S. Winden, D. J. McFarland, E. W. Sellers, H. Shi, T. Paine, D. S. Higgins, A. C. Lo, H. S. Patwa, K. J. Hill, G. D. Huang, and R. L. Ruff. Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis. *Neurology*, 91(3):e258–e267, July 2018.

[25] C. Evans. Some further observations on occipital sharp waves (lambda waves). *Electroencephalography and Clinical Neurophysiology*, (4):371, 1952.

[26] A.-M. Brouwer, B. Reuderink, J. Vincent, M. A. J. v. Gerven, and J. B. F. v. Erp. Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. *Journal of Vision*, 13(3):17–17, September 2013.

[27] L. N. Kaunitz, J. E. Kamienkowski, A. Varatharajah, M. Sigman, R. Q. Quiroga, and M. J. Ison. Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task. *NeuroImage*, 89:297–305, April 2014.

[28] H. Devillez, N. Guyader, and A. Guérin-Dugué. An eye fixation–related potentials analysis of the P300 potential for fixations onto a target object when exploring natural scenes. *Journal of Vision*, 15(13):20, September 2015.

[29] M. Uščumlić and B. Blankertz. Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. *Journal of Neural Engineering*, 13(1):016015, 2016.

[30] H. Devillez, E. Kristensen, N. Guyader, B. Rivet, and A. Guérin-Dugué. The P300 potential for fixations onto target object when exploring natural scenes during a visual task after denoising overlapped EFRP. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1024–1027, April 2015.

[31] H. Renold, R. Chavarriaga, L. A. Gheorghe, and J. d. R. Millán. EEG correlates of active visual search during simulated driving: An exploratory study. In *2014 IEEE International Conference on Systems, Man, and Cybernetics*, 2014.

[32] J. E. Kamienkowski, M. J. Ison, R. Q. Quiroga, and M. Sigman. Fixation-related potentials in visual search: A combined EEG and eye tracking study. *Journal of Vision*, 12(7):4–4, July 2012.

[33] Y.-F. Zhang, X.-Y. Gao, J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu. A novel approach to driving fatigue detection using forehead EOG. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 707–710, April 2015.

[34] A. M. Pena, E. H. Nirjhar, A. Pachuilo, T. Chaspari, and E. D. Ragan. Detecting Changes in User Behavior to Understand Interaction Provenance during Visual Data Analysis. In *IUI Workshops*, 2019.

[35] R. S. Hessels, D. C. Niehorster, C. Kemner, and I. T. C. Hooge. Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behavior Research Methods*, 49(5):1802–1823, 2017.

[36] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols.

In *2000 Symposium on Eye Tracking Research & Applications*, pages 71–78, November 2000.

[37] G. E. Legge, J. S. Mansfield, and S. T. L. Chung. Psychophysics of reading: XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, 41(6):725–743, March 2001.

[38] T. Demiralp, A. Ademoglu, M. Comerchero, and J. Polich. Wavelet analysis of P3a and P3b. *Brain Topography*, 13(4):251–267, June 2001.

[39] U. Graichen, R. Eichardt, P. Fiedler, D. Strohmeier, F. Zanow, and J. Haueisen. SPHARA - A generalized spatial Fourier analysis for multi-sensor systems with non-uniformly arranged sensors: Application to EEG. *PLOS ONE*, 10(4):e0121741, April 2015.

[40] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller. A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, 118(1):98–104, January 2007.

[41] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, March 2012.

[42] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.

[43] G. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10(1):52–55, 2008.

[44] K. Pearson. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, July 1900.

[45] A. O. Dick. Processing time for naming and categorization of letters and numbers. *Perception & Psychophysics*, 9(3):350–352, May 1971.

[46] J. Alegria and P. Bertelson. Time uncertainty, number of alternatives and particular signal-response pair as determinants of choice reaction time. *Acta Psychologica*, 33:36–44, January 1970.

[47] E. Kristensen, A. Guerin-Dugué, and B. Rivet. Comparison between adjar and xDawn algorithms to estimate eye-fixation related potentials distorted by overlapping. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 976–979, April 2015.

[48] A. R. Nikolaev, R. N. Meghanathan, and C. van Leeuwen. Combining EEG and eye movement recording in free viewing: Pitfalls and possibilities. *Brain and Cognition*, 107:55–83, August 2016.

[49] R. Aydarkhanov, M. Uščumlić, R. Chavarriaga, L. Gheorghe, and J. d. R. Millán. Spatial covariance improves BCI performance for late ERPs components with high temporal variability. *Journal of Neural Engineering*, 17(3):036030, June 2020.