



OPEN

DATA DESCRIPTOR

# Mobile BCI dataset of scalp- and ear-EEGs with ERP and SSVEP paradigms while standing, walking, and running

Young-Eun Lee <sup>1</sup>, Gi-Hwan Shin <sup>1</sup>, Minji Lee <sup>1</sup> & Seong-Whan Lee <sup>1,2</sup> ✉

We present a mobile dataset obtained from electroencephalography (EEG) of the scalp and around the ear as well as from locomotion sensors by 24 participants moving at four different speeds while performing two brain-computer interface (BCI) tasks. The data were collected from 32-channel scalp-EEG, 14-channel ear-EEG, 4-channel electrooculography, and 9-channel inertial measurement units placed at the forehead, left ankle, and right ankle. The recording conditions were as follows: standing, slow walking, fast walking, and slight running at speeds of 0, 0.8, 1.6, and 2.0 m/s, respectively. For each speed, two different BCI paradigms, event-related potential and steady-state visual evoked potential, were recorded. To evaluate the signal quality, scalp- and ear-EEG data were qualitatively and quantitatively validated during each speed. We believe that the dataset will facilitate BCIs in diverse mobile environments to analyze brain activities and evaluate the performance quantitatively for expanding the use of practical BCIs.

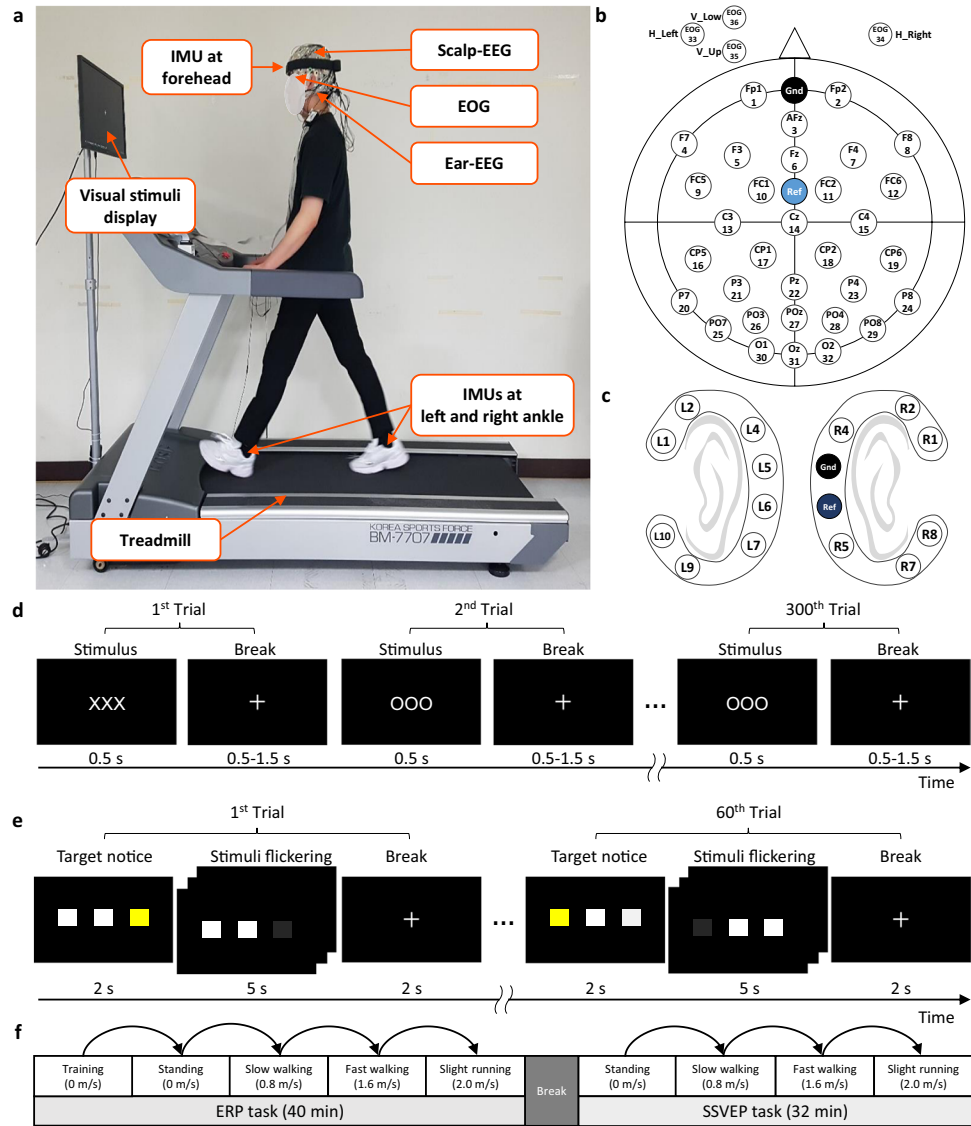
## Background & Summary

Human movement is a complex process that requires the integration of the central and peripheral nervous systems, and therefore, researchers have analyzed human locomotion using brain activity<sup>1,2</sup>. Brain-computer interface (BCI) has been studied based on the communication between human thoughts and external devices to recover the motor sensory function of disabled patients and support daily life of healthy people<sup>3–5</sup>. In particular, electroencephalography (EEG) has been used as the most common method for measuring brain activity with high time resolution, portability, and ease of use<sup>6</sup>; in addition, several attempts have been made to increase its practicality<sup>7–10</sup>. However, EEG recording in a mobile environment can cause artifacts and signal distortion, further resulting in loss of accuracy and signal quality<sup>11,12</sup>. Therefore, research in the mobile environment is necessary to study the brain activity during movements to mitigate limitations such as loss of accuracy and signal quality and to improve the practical BCI technology<sup>9,12</sup>. Moreover, several studies have developed software techniques to realize the practical BCI, such as preprocessing algorithm removing artifacts<sup>13–16</sup> or novel classification algorithm improving user intention performance<sup>12,17</sup>.

To recognize the human intention, two representative exogenous BCI paradigms, event-related potential (ERP)<sup>18</sup> and steady-state visual evoked potential (SSVEP)<sup>19</sup>, are commonly used in the mobile environment owing to their strong responses to brain activity. ERP is a time-locked brain response to stimuli (i.e., visual, auditory, etc.), including a positive peak response (P300) that occurs 300 ms after the stimulus appears. The ERP has a relatively high performance in both scalp-EEG and ear-EEG, with accuracies of 85–95% for scalp-EEG<sup>20,21</sup> and approximately 70% for ear-EEG<sup>9</sup> in a static state. SSVEP is a period brain response in the occipital area to stimuli flickering at a particular frequency. The performance of SSVEP is reliable in terms of the accuracy and signal-to-noise ratio (SNR), with 80–95% accuracy for scalp-EEG<sup>17,20</sup>, but 40–70% accuracy for ear-EEG as it is located far from the occipital cortex<sup>22,23</sup>. Brain signal data obtained by performing BCI paradigms can be used to quantitatively evaluate signal quality in a mobile environment<sup>24</sup>.

Portable and non-hair-bearing EEG have been frequently investigated to enhance the applicability of practical BCI in the real world<sup>25–29</sup>. In particular, ear-EEG, which comprises electrodes placed inside or around the

<sup>1</sup>Korea University, Department of Brain and Cognitive Engineering, Seoul, 02841, Republic of Korea. <sup>2</sup>Korea University, Department of Artificial Intelligence, Seoul, 02841, Republic of Korea. ✉e-mail: [sw.lee@korea.ac.kr](mailto:sw.lee@korea.ac.kr)

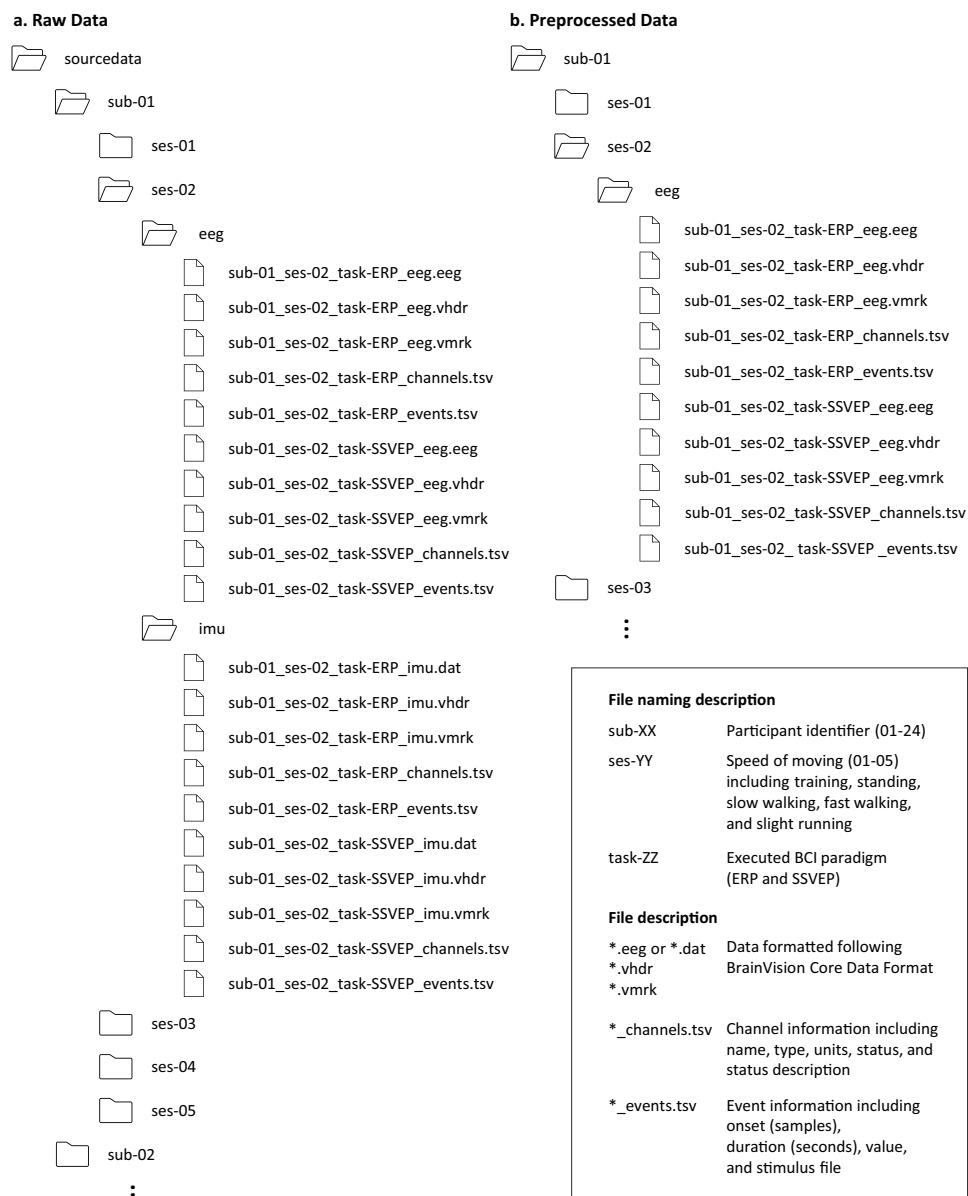


**Fig. 1** Experimental design. **(a)** Experimental setup while standing (0 m/s), slow walking (0.8 m/s), fast walking (1.6 m/s), and slight running (2.0 m/s) on the treadmill, wearing scalp-EEG, ear-EEG, EOG, and IMUs. Informed consent was obtained from the participant for publishing the figure. Channel placement of **(b)** scalp-EEG with EOG and **(c)** ear-EEG. Experimental paradigms for **(d)** ERP paradigm with 300 trials and **(e)** SSVEP paradigm with 60 trials. **(f)** Experimental procedure.

ear, has several advantages over conventional scalp-EEG in terms of stability, portability, and unobtrusiveness<sup>9,30</sup>. Moreover, the signal quality of ear-EEG has been validated for recognizing human intention using several BCI paradigms, including ERP<sup>9,29,31</sup>, SSVEP<sup>22,23</sup>, and others<sup>32</sup>.

Recently, EEG datasets for mobile environments have been published, including motion information and different mobile environments. He *et al.*<sup>33</sup> recorded signals from 60-channel scalp-EEG, 4-channel electrooculography (EOG), and 6 goniometers from 8 participants while walking slowly at a constant speed of 0.45 m/s. They used a BCI paradigm, avatar control, by predicting the joint angle of goniometers from the EEG while walking. Brantley *et al.*<sup>34</sup> collected a dataset consisting of full-body locomotion from 10 participants on stairs, ramps, and level grounds without BCI paradigms, by recording from 60-channel scalp-EEG, 4-channel EOG, 12-channel electromyogram (EMG), and 17 inertial measurement units (IMUs). Wagner *et al.*<sup>35</sup> recorded signals from 108-channel scalp-EEG, 2-channel EMG, 2 pressure sensors, and 3 goniometers from 20 participants without BCI paradigms while walking at a constant speed. Consequently, although these datasets were collected from the mobile environment, the movement condition was kept constant, and two of the datasets were performed without BCI paradigms. In addition, because only scalp-EEG signals were measured, the application of practical BCI was restricted.

In this study, we present a mobile BCI dataset with scalp- and ear-EEGs collected from 24 participants with BCI paradigms at different speeds. Data from 32-channel scalp-EEG, 14-channel ear-EEG, 4-channel EOG, and 27-channel IMUs were recorded simultaneously. The experimental environment involved movements of

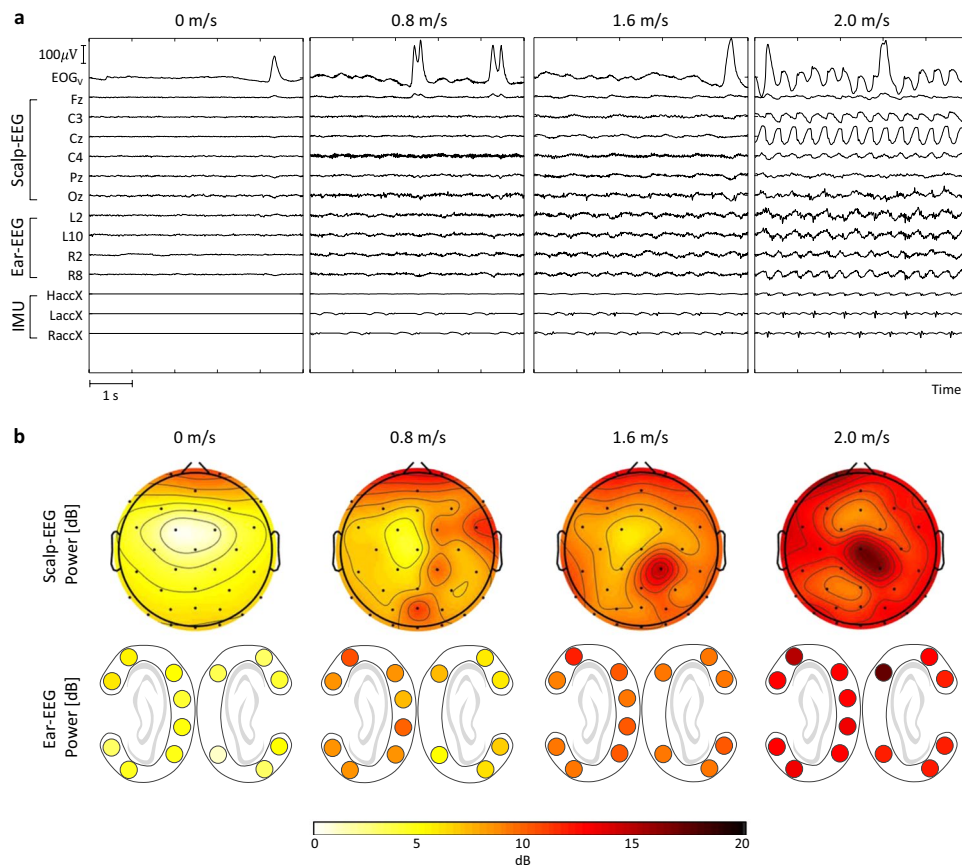


**Fig. 2** Data folder structure. The folders and files are described, including (a) raw data and (b) preprocessed data in the data repository. The folder was named with ‘sub-XX’, ‘ses-YY’, and modality, and the file was named with ‘sub-XX\_ses-YY\_task-ZZ\_WW’. The ‘sub-XX’ indicated the participant identifier, including 1–24, the ‘ses-YY’ indicated the session number, including training(01), standing(02), slow walking(03), fast walking(04), and slight running(05), the ‘task-ZZ’ indicated executed BCI paradigms including ERP and SSVEP, and the modality ‘WW’ indicated the modality of each data, including EEG (scalp-EEG and ear-EEG) and IMU.

Modality	The number of channels	Description
Scalp-EEG	36	32 EEG channels and 4 EOG channels
Ear-EEG	14	8 channels on the left side and 6 channels on the right side
IMU	27	3 devices (head, left ankle, and right ankle)
		9-axis device containing accelerometers, gyroscopes, and magnetometers

**Table 1.** Description of channel types and information of three modalities.

participants at different speeds of 0, 0.8, 1.6, and 2.0 m/s on a treadmill. For each speed, two BCI paradigms were used to evaluate signal quality, which facilitated diverse analysis, including time-domain analysis using ERP data and frequency-domain analysis using SSVEP data. Therefore, we believe that the dataset facilitates addressing the issues of brain dynamics in diverse mobile environments in terms of the cognitive level of multitasking,



**Fig. 3** Examples of the signals and topography at different speeds. **(a)** Time-synchronized subset of scalp-EEG, ear-EEG, and IMU data for 5 s while moving at different speeds of 0, 0.8, 1.6, and 2.0 m/s. The  $EOG_V$  channel was calculated by subtracting lower VEOG from upper VEOG. **(b)** EEG power topography in each channel of scalp-EEG and ear-EEG.

locomotion complexity, and quantitative evaluation of artifact removal methods or classifiers for BCI tasks in the mobile environment.

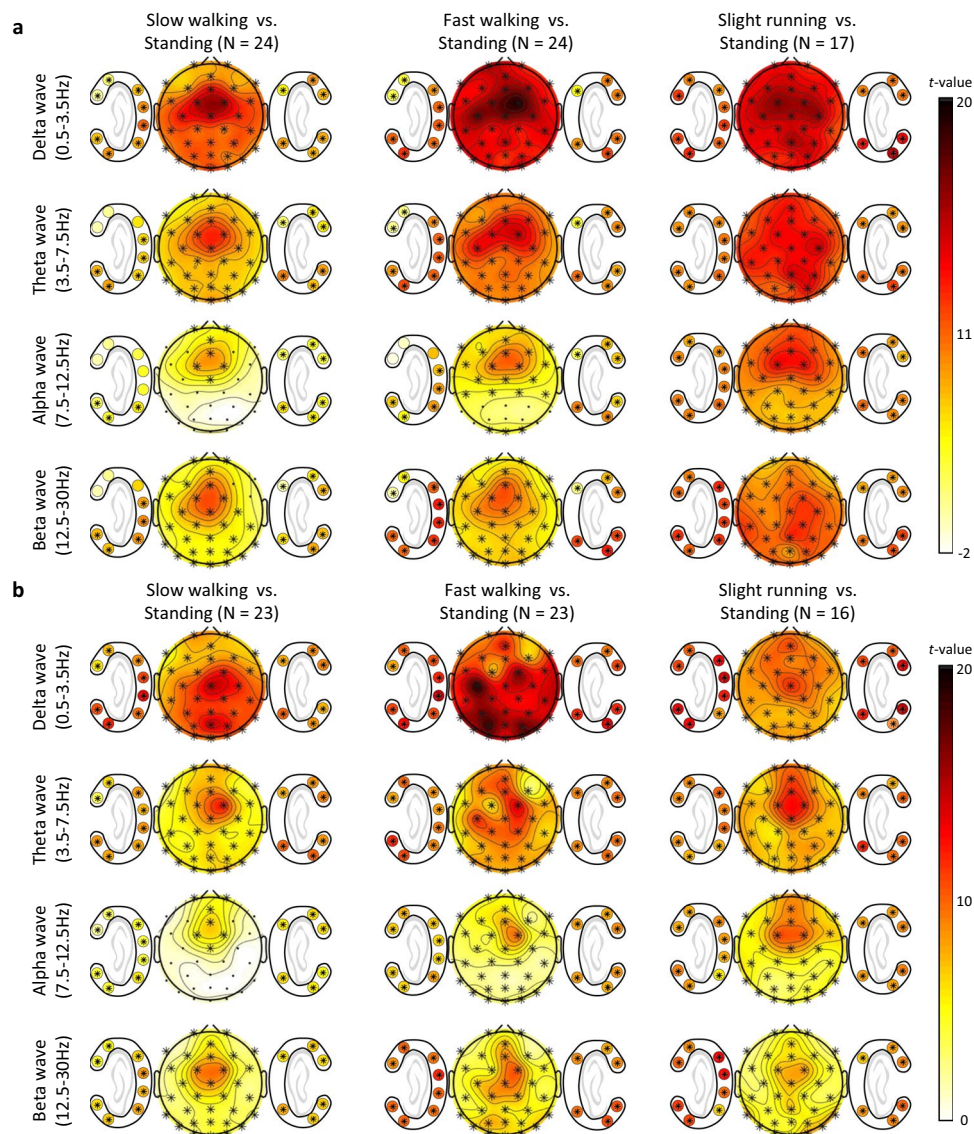
## Methods

**Participants.** Twenty-four healthy individuals (14 men and 10 women,  $24.5 \pm 2.9$  years of age) without any history of neurological or lower limb pathology participated in this experiment. In the ERP tasks, all of them participated, and 17 participants performed a slight running session at a speed of 2.0 m/s. In the SSVEP tasks, 23 of them participated, excluding one because of a personal problem unrelated to the experimental procedure, and 16 participants performed a slight running session at a speed of 2.0 m/s. The participants were provided the option to perform a slight running session. This study was approved by the Institutional Review Board of Korea University (KUIRB-2019-0194-01), and all participants provided written informed consent before the experiments. All experiments were conducted in accordance with the Declaration of Helsinki.

**Data acquisition.** For conducting the experiment, we simultaneously collected data from three different modalities: scalp-EEG, ear-EEG, and IMU (Fig. 1a–c). All data can be accessed from here<sup>36</sup>. To synchronize the three devices, triggers were sent to the recording system of each device simultaneously while presenting a paradigm in MATLAB.

The head circumference of each participant was measured to select an appropriately sized cap of scalp-EEG. We obtained signals from the scalp with 32 EEG Ag/AgCl electrodes according to the 10/20 international system using BrainAmp (Brain Product GmbH). The ground and reference electrodes were placed on Fpz and FCz. In addition, we used four EOG channels to capture dynamically changing eye movements such as blinking. The EOG channels were placed above and below the left eye to measure the vertical eye artifacts ( $VEOG_u$  and  $VEOG_l$ ) as well as at the left and right temples to measure the horizontal eye artifacts ( $HEOG_l$  and  $HEOG_r$ ) (Fig. 1b). The sampling rate of the scalp-EEG and EOG was 500 Hz with a resolution of 32 bits. All electrode impedances were maintained below 50 k $\Omega$ , and most channels were reduced to below 20 k $\Omega$ <sup>33,34</sup>.

The ear-EEG consists of cEEGrids electrodes located around each ear of the participant, with eight channels on the left, six channels on the right, and ground and reference channels on the right at the center (Fig. 1c)<sup>9</sup>. Two cEEGrids were connected to a wireless mobile DC EEG amplifier (SMARTING, mBrainTrain, Belgrade, Serbia).



**Fig. 4** Statistical differences of PSD in each frequency band for scalp- and ear-EEGs between standing and other speeds while (a) ERP and (b) SSVEP. The colored topological maps indicate  $t$ -values and the electrodes in cluster showing a statistically significant effect on spectral power between the data of corresponding speeds are marked with black asterisk ( $p < 0.05$ , cluster-based correction for multiple comparison).

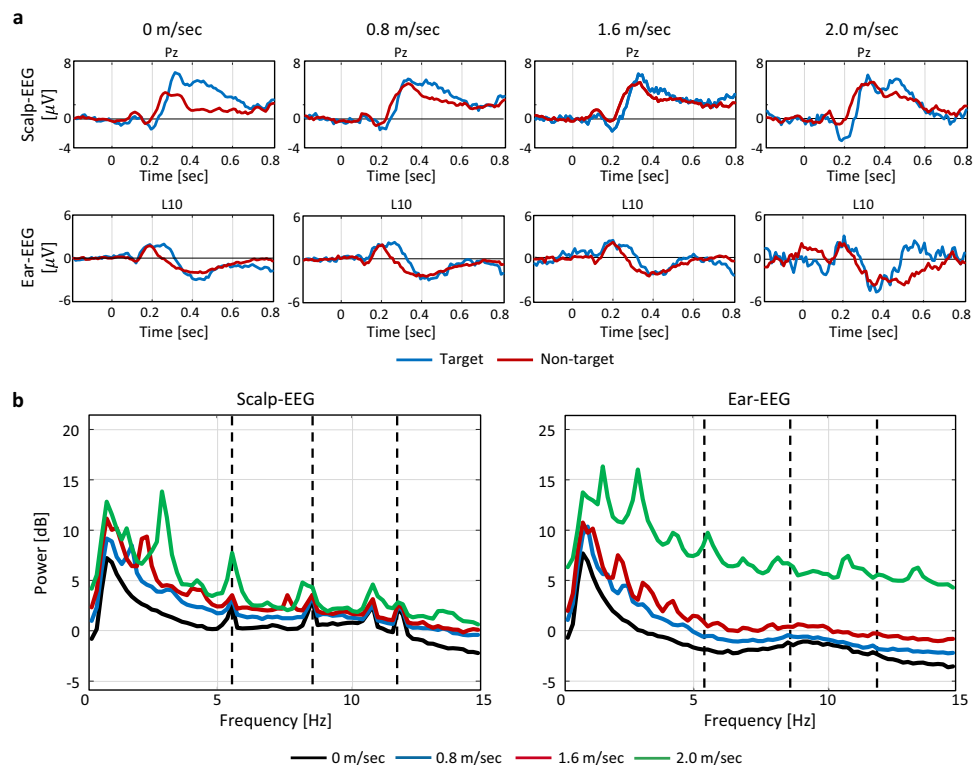
Data were recorded at a sampling rate of 500 Hz and a resolution of 24 bits. All impedances of ear-EEG were maintained below 50 k $\Omega$ , and most channels were reduced to below 20 k $\Omega$ <sup>33,34</sup>.

To measure the locomotion of the participants, three wearable IMU sensors (APDM wearable technologies) were placed at the head, left ankle, and right ankle. An IMU consisted of 9-channel sensors, including a 3-axis accelerometer, 3-axis gyroscope, and 3-axis magnetometer. Therefore, 27-channel IMU signals were collected and recorded at a sampling rate of 128 Hz with a resolution of 32 bits.

**Experimental paradigm.** They performed tasks under the ERP and SSVEP paradigms, in which stimuli were displayed on the monitor during each session at different speeds. Two BCI paradigms were developed based on the OpenBMI (<http://openbmi.org>)<sup>20</sup> and Psychtoolbox (<http://psychtoolbox.org>)<sup>37</sup> in MATLAB (The Mathworks, Natick, MA).

During the ERP task, target ('OOO') and non-target ('XXX') characters were presented on the monitor as visual stimuli. All characters were displayed with a black background at the center of the monitor screen. The proportion of the target was 0.2, and the total number of trials was 300. In a trial, one of the stimuli was presented for 0.5 s, and a fixation cross ('+') was presented to take a break for randomly 0.5–1.5 s (Fig. 1d).

During the SSVEP task, three target SSVEP visual stimuli were displayed at three positions (left, center, and right) on an LCD monitor<sup>19,38</sup>. The frequency range of stimuli containing 5–30 Hz is known to be appropriate for obtaining SSVEP responses<sup>39</sup>. It is also known that movement artifacts have a significant impact on frequency



**Fig. 5** Grand average of all participants of ERP and SSVEP waveforms according to four different speeds of 0, 0.8, 1.6, and 2.0 m/s. **(a)** Grand average baseline corrected waves of all participants for ERP of target and non-target in scalp-EEG at Pz and ear-EEG at L10 for 1 s from  $-200$  to  $800$  ms according to the trigger at different speeds. **(b)** Grand average PSD of all participants for SSVEP in scalp-EEG at Oz (left) and ear-EEG at L10 (right) at different speeds. The dashed line indicated the target frequency, such as 5.45, 8.57, and 12 Hz.

spectrum below 12 Hz<sup>40</sup>. Based on these studies, the stimuli were designed to flicker at 5.45, 8.57, and 12 Hz, which were calculated by dividing the monitor frame rate of 60 Hz by an integer (i.e.,  $60/11$ ,  $60/7$ , and  $60/5$ )<sup>20</sup>. The participants were asked to gaze in the direction of the target stimulus highlighted in yellow. In each trial, the target of a random sequence was noticed for 2 s, after which all stimuli blinked for 5 s, with a break time for 2 s. The SSVEP experiment consisted of 20 trials for each frequency, a total of 60 trials in a session (Fig. 1e).

**Experimental protocol and procedure.** Figure 1a depicts the experimental setup of this study. The participants stood on the treadmill in a lab and were instructed to look at a 24-inch monitor (refresh rate: 60 Hz, resolution:  $1920 \times 1080$  pixels) placed  $80 (\pm 10)$  cm in front of the participants. The participants were monitored and instructed to minimize other movements, such as that of neck or arm, to avoid any artifacts that might occur by movements other than walking. The mobile environment included standing, slow walking, fast walking, and slight running at speeds of 0, 0.8, 1.6, and 2.0 m/s, respectively on the treadmill ( $0^\circ$  inclination)<sup>41,42</sup>.

To proceed with the experiment under the same conditions for each participant, the experimental procedures were sequentially performed (Fig. 1f). They conducted two BCI tasks while standing, slow walking, fast walking, and slight running on the treadmill. Training sessions for the ERP task were conducted at a speed of 0 m/s to train the ERP classifier prior to all ERP tasks. Duration of a session of ERP and SSVEP tasks consisted of 7–8 min, with a total of 40 min and 32 min for all sessions. Each session involved the same procedure with a random sequence of targets. All sessions for one participant were performed on a single day. Moreover, due to the possibility of fatigue and habituation, which could be induced in the sequence of the experiment<sup>43</sup>, the following actions were considered. At first, participants were allowed sufficient breaks between and within sessions when they needed. Furthermore, the participants became familiar with the paradigm stimuli by being fully exposed to them before starting the experiment.

**Preprocessing.** We preprocessed the data using an open-source toolbox for EEG data, such as BBCI ([https://github.com/bbci/bbci\\_public](https://github.com/bbci/bbci_public))<sup>44</sup>, BCILAB (<https://github.com/scn/BCILAB>)<sup>45</sup>, and EEGLAB (<https://scn.ucsd.edu/eeGLAB>)<sup>46</sup> in MATLAB. At first, the data were preprocessed using a high-pass filter that was set above 0.5 Hz using a fifth-order Butterworth filter. Thereafter, three procedures: EOG removal, line noise removal, and interpolation were performed. Vertical EOG components were removed from the scalp-EEG using the *flt\_eog* function in BCILAB<sup>47</sup>. The line noise removal method automatically removed artifacts that contained noise for extended periods of time with several parameters. This method removed bad channels that carried abnormal signals with standard deviations above the threshold of z-score using the function *flt\_clean\_channels* in BCILAB with threshold of 4 and window length of 5 s. The removed bad channels were interpolated using the super-fast

Measure	AUC				SNR			
	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s
s1	0.99	0.83	0.77	0.59	1.03	1.01	0.96	0.99
s2	0.85	0.68	0.58	0.61	0.89	0.91	0.95	0.97
s3	0.78	0.76	0.68	0.56	0.77	0.83	0.88	0.96
s4	0.86	0.84	0.69	0.44	0.93	0.92	0.96	0.95
s5	0.94	0.75	0.60	0.63	1.05	1.06	1.00	1.03
s6	0.98	0.86	0.80	0.63	1.15	1.60	1.47	1.11
s7	0.93	0.81	0.59	0.62	1.01	1.12	1.00	0.97
s8	0.89	0.62	0.68	0.59	0.99	1.18	1.35	1.11
s9	0.92	0.81	0.70	0.54	0.99	0.99	1.02	0.90
s10	0.97	0.70	0.58	0.56	1.05	1.18	1.13	1.04
s11	0.80	0.73	0.59	0.50	0.85	0.89	0.99	0.96
s12	0.80	0.81	0.64	0.54	1.01	0.92	1.01	0.99
s13	0.80	0.75	0.70	0.75	0.91	1.30	1.33	1.03
s14	0.91	0.66	0.57	0.59	1.08	1.25	1.05	0.93
s15	0.94	0.78	0.50	0.57	0.88	0.92	0.91	0.99
s16	0.94	0.85	0.73	0.55	0.88	1.06	1.04	0.95
s17	0.98	0.83	0.78	0.59	0.95	1.04	1.02	1.11
s18	0.92	0.84	0.69	—	1.07	1.01	1.04	—
s19	0.82	0.79	0.68	—	0.94	1.18	1.04	—
s20	0.90	0.71	0.68	—	0.92	1.23	1.10	—
s21	0.75	0.69	0.71	—	0.74	0.92	0.92	—
s22	0.95	0.85	0.72	—	0.93	0.98	0.99	—
s23	0.94	0.77	0.64	—	0.98	1.28	1.20	—
s24	1.00	0.86	0.72	—	0.90	1.00	1.00	—
AVG	0.90	0.77*	0.67*	0.58*	0.95	1.07*	1.06*	1.00
STD	0.07	0.07	0.07	0.06	0.09	0.17	0.14	0.06

**Table 2.** AUC and SNR of ERP in scalp-EEG. Asterisk indicates significance levels of 5% between the performance at 0 m/s and corresponding speed.

spherical method to avoid losing any channel information. On average,  $2.38 \pm 1.94$  channels in the scalp-EEG and  $1.35 \pm 1.18$  channels in the ear-EEG were removed and interpolated for all participants in all sessions. All channels in scalp-EEG and ear-EEG were each re-referenced to a common average reference. We down-sampled the data from scalp-EEG, ear-EEG, and IMU sensors to 100 Hz. The continuous signal was segmented into epoched signals according to the designated trigger timing of each paradigm. For the ERP, each trial was segmented from  $-200$  to  $800$  ms based on the stimulus presentation time. For the SSVEP, each trial was segmented from  $0$ – $5$  s based on the starting time of the stimulus flickering.

### Data Records

All data files are available in Open Science Framework repository<sup>36</sup> and are available under the terms of Attribution 4.0 International Creative Commons License (<http://creativecommons.org/licenses/by/4.0/>).

**Data format.** All data are provided according to the standardized Brain Imaging Data Structure format for EEG data<sup>48</sup> as shown in Fig. 2. The data format followed BrainVision Core Data Format, developed by Brain Products GmbH. The data file is organized with the following naming convention:

sub-XX\_ses-YY\_task-ZZ\_eeg

where the session number includes 1–5, which session 1 indicates training session for ERP and session 2–5 indicates each speed of 0, 0.8, 1.6, and 2.0 m/s, respectively, and the task includes ERP and SSVEP. In ‘source-data’ folder, the data are separated into EEG and IMU because their sampling frequencies are different. In each subject folder, the sampling frequencies of data are downed to 100 Hz and data from all modalities are in a file. The number of channels for each modality is listed in Table 1.

**Missing data.** *Data missing.* The IMU data of participant 21 for ERP at 0 m/s and that of participant 12 for SSVEP at 0.8 m/s were missing because of a malfunction in the communication of the IMU during data collection.

*Trials missing.* The number of trials for ERP data of participant 11 at every speed and participant 13 and 15 at a speed of 2.0 m/s, and SSVEP data of participant 14 at a speed of 2.0 m/s were approximately two-thirds of the normal number of trials because of the malfunction of device communication.

Measure	AUC				SNR			
	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s
s1	0.52	0.55	0.52	0.49	1.28	0.92	1.06	0.97
s2	0.71	0.69	0.52	0.53	0.93	0.94	0.97	1.04
s3	0.66	0.57	0.57	0.54	0.77	0.92	0.86	0.97
s4	0.54	0.54	0.53	0.58	1.05	1.13	0.98	0.83
s5	0.94	0.75	0.63	0.53	0.97	0.99	1.01	0.87
s6	0.86	0.70	0.65	0.51	1.20	1.17	1.09	1.04
s7	0.85	0.74	0.57	0.56	1.06	1.08	1.01	1.01
s8	0.57	0.58	0.59	0.51	1.20	1.06	1.03	1.01
s9	0.81	0.76	0.70	0.60	0.93	1.01	1.01	0.95
s10	0.72	0.64	0.58	0.49	0.98	0.98	0.98	0.92
s11	0.69	0.47	0.55	0.47	1.07	0.94	0.96	1.11
s12	0.64	0.60	0.56	0.54	0.98	0.94	1.00	0.97
s13	0.46	0.46	0.62	0.49	0.97	1.08	1.06	1.02
s14	0.77	0.58	0.46	0.49	1.09	1.09	0.95	0.97
s15	0.76	0.56	0.51	0.55	0.97	0.97	0.94	0.96
s16	0.79	0.71	0.57	0.51	1.01	1.06	1.00	0.99
s17	0.87	0.80	0.67	0.64	1.01	0.97	1.00	0.99
s18	0.51	0.68	0.58	—	0.99	1.03	0.99	—
s19	0.72	0.69	0.73	—	0.88	1.24	0.94	—
s20	0.51	0.41	0.58	—	2.24	1.08	0.88	—
s21	0.69	0.64	0.52	—	0.88	1.01	0.97	—
s22	0.88	0.68	0.55	—	0.95	1.01	0.98	—
s23	0.80	0.69	0.56	—	0.94	0.97	0.97	—
s24	0.90	0.73	0.64	—	0.95	0.96	0.96	—
AVG	0.72	0.63*	0.58*	0.53*	1.06	1.02	0.98	0.98
STD	0.14	0.10	0.06	0.04	0.27	0.08	0.05	0.06

**Table 3.** AUC and SNR of ERP in ear-EEG. Asterisk indicates significance levels of 5% between the performance at 0 m/s and corresponding speed.

*Excluded data.* The data of participant 17 at 2.0 m/s for SSVEP, participant 19 at 2.0 m/s for ERP and SSVEP, and participant 20 at 2.0 m/s for ERP were excluded since the electrodes did not adhere well during data recording, resulting in loss of more than 50% in a session.

### Technical Validation

Figure 3a depicts an example of the scalp-EEG, ear-EEG, and IMU signals for 5 s at speeds of 0, 0.8, 1.6, and 2.0 m/s. The amplitudes of the scalp-EEG, ear-EEG, and IMU increased as the speed increased<sup>41,49,50</sup>. Figure 3b depicts an example of the topography for the scalp-EEG and ear-EEG at different speeds of 0, 0.8, 1.6, and 2.0 m/s. The powers of the scalp-EEG and ear-EEG increased as the speed increased. To evaluate the dataset, statistical analysis was conducted using a one-tailed paired *t*-test to compare the performance at each moving speed to the performance at standing, as indicated by the asterisk at a confidence level of 95%. We quantitatively evaluated the preprocessed data of the ERP and SSVEP paradigms in terms of accuracy and SNR. Moreover, the baseline corrected waves for ERP and power spectral density (PSD) for SSVEP were plotted to evaluate the signal quality.

**Statistical verification.** To verify the dataset, we performed statistical verification to demonstrate significant differences between the speeds across every channel of scalp-EEG and ear-EEG. Figure 4a for the ERP and Fig. 4b for the SSVEP depicts the topological map of *t*-values in particular frequency bands, including delta waves (0.5–3.5 Hz), theta waves (3.5–7.5 Hz), alpha waves (7.5–12.5 Hz), and beta waves (12.5–30 Hz)<sup>51</sup>. In particular, PSDs in each frequency band were analyzed using cluster-based correction with non-parametric permutation testing for multiple comparisons to verify the difference between the data at four speeds, including 0, 0.8, 1.6, and 2.0 m/s. The significance probabilities and critical values of permutation distribution are estimated using Monte-Carlo method with iterations of 10,000.

Significant channels could indicate that noise signals are included in corresponding frequency bands and speeds. The topography of the delta band depicts that step frequencies, which was mostly in the range of 0.5–3.5 Hz, affect most channels at all speeds. During slight running session, all channels, including scalp-EEG and ear-EEG, were significantly different in entire frequency band. In addition, paradigm-related areas such as the occipital area during SSVEP tasks and the central area during ERP tasks showed the large *t*-values in delta band, resulting in low concentration on the tasks due to the workload of multi-tasking.



Measure	Accuracy (%)				SNR			
	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s
s1	100	98.33	98.33	100	4.60	3.82	2.27	2.48
s2	100	100	98.33	90.00	3.93	3.42	3.01	2.91
s3	96.67	90.00	83.33	33.33	1.52	1.69	1.35	1.59
s4	100	100	100	93.33	2.43	2.09	2.06	1.79
s5	96.67	96.67	96.67	31.67	3.79	3.42	3.42	1.63
s6	100	81.67	75.00	35.00	2.35	2.18	1.88	1.36
s7	100	98.33	96.67	35.00	4.12	3.68	2.87	2.66
s8	96.67	81.67	65.00	41.67	1.14	1.31	1.01	1.20
s9	98.33	81.67	98.33	33.33	2.46	1.66	1.90	2.31
s10	100	86.67	78.33	35.00	3.42	1.88	1.90	1.74
s11	100	86.67	81.67	50.00	2.19	1.71	1.99	2.05
s12	100	96.67	90.00	60.00	2.33	2.16	1.91	1.45
s13	80.00	66.67	80.00	71.67	1.32	1.23	1.11	1.45
s14	100	96.67	90.00	36.17	2.87	2.55	2.22	2.19
s15	100	100	100	98.33	2.87	2.82	2.39	2.57
s16	78.33	93.33	96.67	—	2.72	2.22	2.09	—
s17	80.00	70.00	56.67	—	1.84	1.35	1.13	—
s18	91.67	63.33	48.33	31.67	1.43	1.32	1.17	1.63
s19	33.33	33.33	31.67	—	1.17	1.15	0.99	—
s20	45.00	48.33	50.00	—	1.32	1.27	1.19	—
s21	96.67	91.67	98.33	—	2.74	2.60	2.81	—
s22	100	100	95.00	—	2.81	2.67	2.29	—
s23	46.67	50.00	46.67	—	1.24	1.02	1.13	—
AVG	88.70	83.12*	80.65*	54.76*	2.64	2.14*	1.92*	1.94*
STD	19.52	18.68	20.38	25.84	0.99	0.84	0.68	0.51

**Table 4.** Accuracy and SNR of SSVEP in scalp-EEG. Asterisk indicates significance levels of 5% between the performance at 0 m/s and corresponding speed.

**Evaluation of ERP.** The ERP dataset was evaluated by demonstrating ERP waves and metrics using the area under the receiver operating characteristic curve (AUC) and approximate SNR at each speed. AUC indicates the true positive rate over the false positive rate of the results. To acquire the AUC, the features of the ERP were extracted by the power over time intervals of every 50 ms from 200 ms to 450 ms. For the classification, we used a conventional classifier, regularized linear discriminant analysis, to evaluate the ERP performance. The data from the training session at a speed of 0 m/s were used for the training set, and the other dataset containing different speeds was used for the testing set. The SNR can indicate the quality of signals, and approximate SNR of ERP was calculated by the root mean square (RMS) of the amplitude of the peaks at P300 divided by the RMS of the average amplitude of the pre-stimulus baseline (−200 to 0 ms) at channel Pz<sup>52,53</sup>.

Figure 5a depicts the baseline corrected waves of the target and non-target in the scalp- and ear-EEGs at channels Pz and L10 at each speed. The higher the speed, the lower the amplitude of the P300 components of the target in both the scalp- and ear-EEGs. Tables 2 and 3 list the performance of ERP in the scalp-EEG and ear-EEG, respectively. The grand average AUCs of ERP for all participants were  $0.90 \pm 0.07$ , and  $0.67 \pm 0.07$  ( $p < 0.05$ ) in the scalp-EEG at speeds of 0 and 1.6 m/s, respectively, and  $0.72 \pm 0.14$  and  $0.58 \pm 0.06$  ( $p < 0.05$ ) in the ear-EEG at speeds of 0 and 1.6 m/s, respectively. The grand average SNRs of ERP for all participants were  $0.95 \pm 0.09$  and  $1.06 \pm 0.14$  ( $p < 0.05$ ) for the scalp-EEG at speeds of 0 and 1.6 m/s, respectively, and  $1.06 \pm 0.27$  and  $0.98 \pm 0.05$  for the ear-EEG at speeds of 0 and 1.6 m/s, respectively.

**Evaluation of SSVEP.** The SSVEP dataset was evaluated by implementing statistical analysis to measure the signal properties using PSD, and the metrics using accuracy and approximate SNR at each speed. Accuracy was measured as the percentage of correct predictions in the total number of cases. A canonical correlation analysis was used for the classification that does not require the training datasets. The SNR of SSVEP was calculated using the ratio of the power of the target frequencies to the power of the neighboring frequencies (resolution: 0.25 Hz, number of neighbors: 12)<sup>54</sup>.

Figure 5b depicts the PSD of the SSVEP for the scalp-EEG and ear-EEG at channels Oz and L10 at each speed. The higher the speed, the greater the power in all frequency spectra for both the scalp- and ear-EEGs. Tables 4 and 5 list the performance of the SSVEP for scalp-EEG and ear-EEG, respectively. The grand average accuracies of SSVEP for all participants were  $88.70 \pm 19.52\%$  and  $80.65 \pm 20.38\%$  ( $p < 0.05$ ) for scalp-EEG at speeds of 0 and 1.6 m/s, respectively, and  $53.19 \pm 13.93$  and  $39.57 \pm 6.39$  ( $p < 0.05$ ) for ear-EEG at speeds of 0 and 1.6 m/s, respectively. The grand average SNRs of SSVEP for all participants were  $2.64 \pm 0.99$  and  $1.92 \pm 0.68$  ( $p < 0.05$ ) for the scalp-EEG at speeds of 0 and 1.6 m/s, respectively, and  $1.21 \pm 0.23$  and  $1.03 \pm 0.10$  ( $p < 0.05$ ) for the ear-EEG at speeds of 0 and 1.6 m/s, respectively.

Measure	Accuracy (%)				SNR			
	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s	0 m/s	0.8 m/s	1.6 m/s	2.0 m/s
s1	96.67	73.33	46.67	26.67	2.06	1.63	1.32	1.09
s2	61.67	41.67	36.67	33.33	1.37	1.10	0.86	1.10
s3	53.33	46.67	36.67	33.33	1.15	1.25	0.97	2.04
s4	55.00	41.67	33.33	43.33	1.11	0.96	0.94	1.50
s5	68.33	68.33	53.33	28.33	1.57	1.27	1.08	1.31
s6	41.67	38.33	40.00	33.33	1.04	1.08	1.07	1.19
s7	58.33	48.33	41.67	33.33	1.18	1.21	1.11	1.29
s8	58.33	55.00	41.67	25.00	1.08	1.11	1.14	1.12
s9	60.00	51.67	38.33	35.00	1.17	1.02	0.92	1.55
s10	51.67	40.00	38.33	31.67	1.11	1.06	1.05	1.14
s11	35.00	33.33	36.67	35.00	1.07	0.98	0.95	1.21
s12	75.00	55.00	38.33	36.67	1.46	1.40	1.12	1.10
s13	46.67	38.33	50.00	36.67	1.09	1.09	1.03	1.22
s14	51.67	41.67	35.00	36.17	1.19	0.94	0.95	1.24
s15	50.00	38.33	35.00	36.67	1.42	1.18	0.97	1.01
s16	48.33	48.33	46.67	—	1.14	1.08	1.06	—
s17	46.67	38.33	35.00	—	1.05	0.93	0.93	—
s18	48.33	45.00	30.00	28.33	1.13	0.87	1.13	1.31
s19	40.00	25.00	35.00	—	0.98	0.89	1.04	—
s20	40.00	30.00	35.00	—	1.02	0.93	0.93	—
s21	63.33	50.00	55.00	—	1.06	1.14	1.11	—
s22	43.33	50.00	36.67	—	1.27	1.05	1.01	—
s23	30.00	31.67	35.00	—	1.05	1.00	1.08	—
AVG	53.19	44.78*	39.57*	33.30*	1.21	1.09*	1.03*	1.27
STD	13.93	11.11	6.39	4.43	0.23	0.17	0.10	0.24

**Table 5.** Accuracy and SNR of SSVEP in ear-EEG. Asterisk indicates significance levels of 5% between the performance at 0 m/s and corresponding speed.

## Usage Notes

This mobile dataset is available in the BrainVision Core Data Format. For analyzing the dataset, we recommend using a common open-source toolbox for EEG data, such as BBCI ([https://github.com/bbci/bbci\\_public](https://github.com/bbci/bbci_public))<sup>44</sup>, OpenBMI (<http://openbmi.org>)<sup>20</sup>, and EEGLAB (<https://scn.ucsd.edu/eeGLab>)<sup>46</sup> in the MATLAB environment, or MNE (<https://martinos.org/mne>)<sup>55</sup> in the Python environment. The supporting code is available on GitHub (<https://github.com/DeepBCI/Deep-BCI>). For the preprocessing, we recommend performing down-sampling to give all signals equal sampling frequency, filtering out extremely low frequency below 0.1 Hz at least to remove the DC drift using a high-pass filter, and interpolating the high distributed channels among all channels. This dataset can be used for the performance evaluation of artifact removal methods and analysis of mental states with quantitative evaluation via BCI paradigms in a mobile environment.

## Code availability

The MATLAB scripts are available for loading data, for evaluating classification performance or signal quality, and for plotting figures at [https://github.com/youngun1209/MobileBCI\\_Data](https://github.com/youngun1209/MobileBCI_Data).

Received: 20 January 2021; Accepted: 8 November 2021;

Published: 20 December 2021

## References

1. Barthélemy, D., Grey, M. J., Nielsen, J. B. & Bouyer, L. Involvement of the corticospinal tract in the control of human gait. *Prog. Brain Res.* **192**, 181–197 (2011).
2. Jahn, K. *et al.* Brain activation patterns during imagined stance and locomotion in functional magnetic resonance imaging. *Neuroimage* **22**, 1722–1731 (2004).
3. Lee, M.-H., Fazli, S., Mehnert, J. & Lee, S.-W. Subject-dependent classification for robust idle state detection using multi-modal neuroimaging and data-fusion techniques in BCI. *Pattern Recognit.* **48**, 2725–2737 (2015).
4. Jeong, J.-H., Shim, K.-H., Kim, D.-J. & Lee, S.-W. Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 1226–1238 (2020).
5. Kwon, O.-Y., Lee, M.-H., Guan, C. & Lee, S.-W. Subject-independent brain-computer interfaces based on deep convolutional neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **31**, 3839–3852 (2019).
6. Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G. & Vaughan, T. M. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **113**, 767–791 (2002).
7. Artoni, F. *et al.* Unidirectional brain to muscle connectivity reveals motor cortex control of leg muscles during stereotyped walking. *Neuroimage* **159**, 403–416 (2017).
8. Luu, T. P., Nakagome, S., He, Y. & Contreras-Vidal, J. L. Real-time EEG-based brain-computer interface to a virtual avatar enhances cortical involvement in human treadmill walking. *Sci. Rep.* **7**, 8895 (2017).

9. Debener, S., Emkes, R., De Vos, M. & Bleichner, M. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Sci. Rep.* **5**, 16743 (2015).
10. Jeong, J.-H., Kwak, N.-S., Guan, C. & Lee, S.-W. Decoding Movement-Related Cortical Potentials Based on Subject-Dependent and Section-Wise Spectral Filtering. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 687–698 (2020).
11. Bulea, T. C., Prasad, S., Kilicarslan, A. & Contreras-Vidal, J. L. Sitting and standing intention can be decoded from scalp EEG recorded prior to movement execution. *Front. Neurosci.* **8**, 376 (2014).
12. Kwak, N.-S., Müller, K.-R. & Lee, S.-W. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLoS One* **12**, e0172578 (2017).
13. Gwin, J. T., Gramann, K., Makeig, S. & Ferris, D. P. Removal of movement artifact from high-density EEG recorded during walking and running. *J. Neurophysiol.* **103**, 3526–3534 (2010).
14. Castermans, T. *et al.* Optimizing the performances of a P300-based brain-computer interface in ambulatory conditions. *IEEE J. Emerg. Sel. Topics Circuits Syst.* **1**, 566–577 (2011).
15. Nordin, A. D., Hairston, W. D. & Ferris, D. P. Dual-electrode motion artifact cancellation for mobile electroencephalography. *J. Neural Eng.* **15**, 056024 (2018).
16. Lee, Y.-E., Kwak, N.-S. & Lee, S.-W. A real-time movement artifact removal method for ambulatory brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 2660–2670 (2020).
17. Kwak, N.-S., Müller, K.-R. & Lee, S.-W. A lower limb exoskeleton control system based on steady state visual evoked potentials. *J. Neural Eng.* **12**, 056009 (2015).
18. Lee, M.-H., Williamson, J., Won, D.-O., Fazli, S. & Lee, S.-W. A high performance spelling system based on EEG-EOG signals with visual feedback. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**, 1443–1459 (2018).
19. Won, D.-O., Hwang, H.-J., Dähne, S., Müller, K.-R. & Lee, S.-W. Effect of higher frequency on the classification of steady-state visual evoked potentials. *J. Neural Eng.* **13**, 016014 (2015).
20. Lee, M.-H. *et al.* EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy. *GigaScience* **8**, giz002 (2019).
21. Yeom, S.-K., Fazli, S., Müller, K.-R. & Lee, S.-W. An efficient ERP-based brain-computer interface using random set presentation and face familiarity. *PLoS One* **9**, e111157 (2014).
22. Floriano, A., Diez, P. F. & Bastos-Filho, T. F. Evaluating the influence of chromatic and luminance stimuli on SSVEPs from behind-the-ears and occipital areas. *Sensors* **18**, 615 (2018).
23. Kwak, N.-S. & Lee, S.-W. Error correction regression framework for enhancing the decoding accuracies of ear-EEG brain-computer interfaces. *IEEE Trans. Cybern.* **50**, 3654–3667 (2019).
24. Gramann, K., Gwin, J. T., Bigdely-ShamloDelorme, N., Ferris, D. P. & Makeig, S. Visual evoked responses during standing and walking. *Front. Hum. Neurosci.* **4**, 202 (2010).
25. Wang, Y.-T. *et al.* An online brain-computer interface based on SSVEPs measured from non-hair-bearing areas. *IEEE Trans. Neural Syst. Rehabil. Eng.* **25**, 14–21 (2016).
26. Chen, Y. *et al.* A high-security EEG-based login system with RSVP stimuli and dry electrodes. *IEEE Trans. Inf. Forensic Secur.* **11**, 2635–2647 (2016).
27. Wei, C.-S., Wang, Y.-T., Lin, C.-T. & Jung, T.-P. Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**, 400–406 (2018).
28. Kidmose, P., Looney, D., Ungstrup, M., Rank, M. L. & Mandic, D. P. A study of evoked potentials from ear-EEG. *IEEE Trans. Biomed. Eng.* **60**, 2824–2830 (2013).
29. Bleichner, M. G., Mirkovic, B. & Debener, S. Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison. *J. Neural Eng.* **13**, 066004 (2016).
30. Goverdovsky, V., Looney, D., Kidmose, P. & Mandic, D. P. In-ear EEG from viscoelastic generic earpieces: Robust and unobtrusive 24/7 monitoring. *IEEE Sens. J.* **16**, 271–277 (2015).
31. Bleichner, M. G. & Debener, S. Concealed, unobtrusive ear-centered EEG acquisition: cEEGrids for transparent EEG. *Front. Hum. Neurosci.* **11**, 163 (2017).
32. Mirkovic, B., Bleichner, M. G., De Vos, M. & Debener, S. Target speaker detection with concealed EEG around the ear. *Front. Neurosci.* **10**, 349 (2016).
33. He, Y., Luu, T. P., Nathan, K., Nakagome, S. & Contreras-Vidal, J. L. A mobile brain-body imaging dataset recorded during treadmill walking with a brain-computer interface. *Sci. Data* **5**, 180074 (2018).
34. Brantley, J. A., Luu, T. P., Nakagome, S., Zhu, F. & Contreras-Vidal, J. L. Full body mobile brain-body imaging data during unconstrained locomotion on stairs, ramps, and level ground. *Sci. Data* **5**, 180133 (2018).
35. Wagner, J. *et al.* High-density EEG mobile brain/body imaging data recorded during a challenging auditory gait pacing task. *Sci. Data* **6**, 211 (2019).
36. Lee, Y.-E., Shin, G.-H., Lee, M. & Lee, S.-W. Mobile BCI dataset of scalp- and ear-EEGs with ERP and SSVEP paradigms while standing, walking, and running. *Open Science Framework* <https://doi.org/10.17605/OSF.IO/R7S9B> (2021).
37. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? *Perception* **36**, 14 (2007).
38. Lee, M.-H., Williamson, J., Lee, Y.-E. & Lee, S.-W. Mental fatigue in central-field and peripheral-field steady-state visually evoked potential and its effects on event-related potential responses. *Neuroreport* **29**, 1301–1308 (2018).
39. Parini, S., Maggi, L., Turconi, A. C. & Andreoni, G. A robust and self-paced BCI system based on a four class SSVEP paradigm: algorithms and protocols for a high-transfer-rate direct brain communication. *Comput. Intell. Neurosci.* **2009**, 864564 (2009).
40. Snyder, K. L., Kline, J. E., Huang, H. J. & Ferris, D. P. Independent component analysis of gait-related movement artifact recorded using EEG electrodes during treadmill walking. *Front. Hum. Neurosci.* **9**, 639 (2015).
41. Kline, J. E., Huang, H. J., Snyder, K. L. & Ferris, D. P. Isolating gait-related movement artifacts in electroencephalography during human walking. *J. Neural Eng.* **12**, 046022 (2015).
42. Nordin, A. D., Hairston, W. D. & Ferris, D. P. Human electrocortical dynamics while stepping over obstacles. *Sci. Rep.* **9**, 4693 (2019).
43. Edmunds, K. J. *et al.* Cortical recruitment and functional dynamics in postural control adaptation and habituation during vibratory proprioceptive stimulation. *J. Neural Eng.* **16**, 026037 (2019).
44. Krepki, R., Blankertz, B., Curio, G. & Müller, K.-R. The Berlin Brain-Computer Interface (BBCI)—towards a new communication channel for online control in gaming applications. *Multimed. Tools Appl.* **33**, 73–90 (2007).
45. Kothe, C. A. & Makeig, S. BCILAB: a platform for brain-computer interface development. *J. Neural Eng.* **10**, 056014 (2013).
46. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
47. He, P., Wilson, G. & Russell, C. Removal of ocular artifacts from electro-encephalogram by adaptive filtering. *Med. and Biol. Eng. and Comput.* **42**, 407–412 (2004).
48. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci. Data* **6**, 1–5 (2019).
49. Lin, Y.-P., Wang, Y. & Jung, T.-P. Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset. *J. NeuroEng. Rehabil.* **11**, 119 (2014).
50. Zink, R., Hunyadi, B., Van Huffel, S. & De Vos, M. Mobile EEG on the bike: disentangling attentional and physical contributions to auditory attention tasks. *J. Neural Eng.* **13**, 046017 (2016).

51. Barollo, F. *et al.* Postural control adaptation and habituation during vibratory proprioceptive stimulation: an HD-EEG investigation of cortical recruitment and kinematics. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 1381–1388 (2020).
52. Schimmel, H. The ( $\pm$ ) reference: Accuracy of estimated mean components in average response studies. *Science* **157**, 92–94 (1967).
53. Vos, M. D., Gandras, K. & Debener, S. Towards a truly mobile auditory brain-computer interface: Exploring the P300 to take away. *Int. J. Psychophysiol.* **91**, 46–53 (2014).
54. Nakanishi, M., Wang, Y., Wang, Y.-T., Mitsukura, Y. & Jung, T.-P. Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate. *PLoS One* **9**, e99235 (2014).
55. Gramfort, A. *et al.* MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* **7**, 267 (2013).

## Acknowledgements

This work was supported by the Institute for Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean Government, No. 2017-0-00451: Development of BCI based Brain and Cognitive Computing Technology for Recognizing User's Intentions using Deep Learning, No. 2015-0-00185: Development of Intelligent Pattern Recognition Softwares for Ambulatory Brain-Computer Interface, and No. 2019-0-00079: Artificial Intelligence Graduate School Program (Korea University). We would like to express our sincere gratitude to Y.-H. Kang and D.-Y. Lee for their assistance in data collection, and N.-S. Kwak for his advice while designing the experiment.

## Author contributions

Y.-E.L. contributed to the design of the experiment, data collection, software programming, data validation, and preparation of the manuscript. G.-H.S. contributed to data collection, software programming, data validation, and preparation of the manuscript. M.L. contributed to data validation and preparation of the manuscript. S.-W.L. contributed to supervision of the project and editing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to S.-W.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2021, corrected publication 2022