scientific data

OPEN

DATA DESCRIPTOR

A multi-subject and multi-session EEG dataset for modelling human visual object recognition

Shuning Xue[1,2], Bu Jin[2], Jie Jiang[2], Longteng Guo[2], Jin Zhou[3], Changyong Wang[3] & Jing Liu[1,2] ✉

We share a multi-subject and multi-session (MSS) dataset with 122-channel electroencephalographic (EEG) signals collected from 32 human participants. The data was obtained during serial visual presentation experiments in two paradigms. Dataset of first paradigm consists of around 800,000 trials presenting stimulus sequences at 5 Hz. Dataset of second paradigm comprises around 40,000 trials displaying each image for 1 second. Each participant completed between 1 to 5 sessions on different days, and each session lasted for approximately 1.5 hours of EEG recording. The stimulus set used in the experiments included 10,000 images, with 500 images per class, manually selected from PASCAL and ImageNet image databases. The MSS dataset can be useful for various studies, including but not limited to (1) exploring the characteristics of EEG visual response, (2) comparing the differences in EEG response of different visual paradigms, and (3) designing machine learning algorithms for cross-subject and cross-session brain-computer interfaces (BCIs) using EEG data from multiple subjects and sessions.

Background & Summary

Humans can identify objects within a few hundred milliseconds with remarkable accuracy[1,2]. The study of human visual processing mechanisms is a vital subject of neuroscience[3,4]. Over the years, researchers have developed various paradigms to explore the response properties of the visual cortex[5–7]. Despite the progress made, many unknown factors still exist in visual processing mechanisms, and modelling time-varying visual perception functions remains a challeng. To gain deeper insights into the cognitive basis of human vision, recent studies have addressed the significance of large-scale neuroimaging datasets[8–12].

Electroencephalography (EEG) is an essential neuroimaging technique that captures the brain's electrical activity. Numerous studies have revealed valuable insights into how EEG responds to visual stimuli[13–17]. To facilitate further study of human visual processing mechanisms, there is a growing interest in collecting EEG signals using naturalistic visual stimuli[10,11,18,19].

To our knowledge, we present below five datasets of visually evoked EEG responses.

The first dataset[20,21], referred to as EEG-Classification in[19], includes 12,000 trials stimulated by 2,000 images from the ImageNet[22] database, collected from six subjects. This dataset features a block-based design, meaning that stimuli of the same class are presented as a single block. Each of the 50 images in each category is displayed sequentially for a duration of 500 milliseconds. The second dataset[18] includes 40,000 trials stimulated by real-world images from the ImageNet database, collected from a single male subject with a low-speed random-design paradigm. The third dataset[11], called THINGS-EEG, records over one million EEG responses to 22,248 images from the THINGS stimulus set[23], measured with a 10-Hz rapid serial visual presentation (RSVP) paradigm on 50 subjects. The fourth dataset[10] also features the THINGS stimulus set, including 820,160 trials on ten participants using a 5-Hz RSVP paradigm. The fifth dataset, known as EEG-ImageNet[19], contains 64,000 trials stimulated by 4,000 images from the ImageNet database, collected from 16 subjects. The EEG-ImageNet dataset features stimuli with multi-granularity labels, including 40 coarse-grained categories and 40 fine-grained categories, with each category consisting of 50 manually curated images. During the experiment, the 4,000

[1]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China. [2]Zidongtaichu Foundation Model Research Center, Institute of Automation, Chinese Academy of Sciences, Beijing, China. [3]Department of Advanced Interdisciplinary Studies, Institute of Basic Medical Sciences and Tissue Engineering Research Center, Beijing, China. ✉e-mail: jliu@nlpr.ia.ac.cn

Parameter	Value
Number of subjects	32
Number of sessions	100
Number of RSVP trials	~800,000
Number of low-speed trials	~40,000
Number of image categories	20
Number of images	10,000
Number of channels	122-channel EEG, 2-bipolar EOG
Sampling rate	1,000 Hz
Bandpass filter	0.1 to 100 Hz
Notch filter	50 Hz
Epoch length	500 ms

Table 1. The key information for the dataset.

images were displayed in bursts for 0.5 seconds each. Researchers have leveraged these datasets to explore human visual processing mechanisms and to develop visual decoding models using deep learning techniques[24–27].

To gain more insights into the complex visual processing system and to develop robust artificial neural networks, it is essential to gather additional EEG datasets using a variety of new image stimuli. In this study, we present a multi-subject and multi-session (MSS) dataset that includes EEG responses to 10,000 images from 32 subjects, collected across multiple independent sessions. The image stimuli are sourced from the PASCAL[28] and ImageNet[22] databases, including 20 semantic categories with 500 images for each category.

Our dataset offers a significant advantage compared to previously datasets: each category contains a larger number of samples. This abundance is beneficial for studying classification using deep learning models and helps mitigate the overfitting problem that can occur with too few samples. Consequently, this dataset is a suitable resource for researchers interested in applying deep learning algorithms to EEG visual decoding. Additionally, our dataset includes data from two different paradigms: Rapid Serial Visual Presentation (RSVP) and low-speed paradigms. These signals can be used to investigate the differences in EEG responses triggered by varying presentation speeds. Furthermore, our dataset utilizes a new source of visual stimuli compared to earlier datasets, and its volume is comparable to the largest dataset currently available for EEG visual recognition. These features fosters further research in EEG visual analysis.

Due to the substantial variability of EEG signals, neural decoding of visual information from EEG across multiple subjects and multiple sessions is a long-standing challenge. To address this issue, we recruited 32 subjects and collected data from multiple independent sessions. Among the 32 subjects, 14 completed five sessions (around one week apart), 3 conducted three sessions, 5 conducted two sessions, and 10 conducted one session based on their preferences. Each session includes four RSVP runs and four low-speed runs. The MSS dataset, which contains data from multiple subjects and sessions, provides a valuable resource for studying variations in EEG responses among different subjects, as well as signal changes over time within individual subjects. This dataset can also be employed to assess the effectiveness of EEG visual decoding methods across different subjects and sessions. We anticipate this new dataset will advance research in EEG-based brain-computer interfaces (BCIs) by addressing the challenges associated with cross-subject and cross-session decoding.

For validation purposes, we trained the fully connected (FC) model for EEG signal classification. The results demonstrate that the dataset carries intricate neural reactions to images, making it a valuable resource for exploring the mechanisms of visual object recognition.

This paper provides a comprehensive description of the experimental paradigm, the recording process, and the data preprocessing for the MSS dataset. The technical validation results demonstrate that the recorded signals contain distinguishable EEG responses to naturalistic visual stimuli. The published dataset offers both preprocessed and raw data, which can be utilized for future research. Furthermore, our validation codes are accessible online, allowing researchers to reproduce all the experimental results presented in this paper. The key information regarding the MSS dataset is summarized in Table 1.

Methods

Participants. Thirty-two healthy adults (mean age 23.66 years, SD = 2.34; 4 females, 28 males) volunteered to participate in the experiment (see the details in Table 2). Participants were recruited from the college student population in Beijing. All participants reported normal or corrected-to-normal vision and reported no neurological or psychiatric disorders.

Informed written consent was obtained from all participants at the start of the experiments. The experimenter explained the detailed information about the consent and the experiment process. participants carefully reviewed the consent form and signed two copies: one for themselves and one for the experimenter. each participant received monetary compensation at the end of each session.

The informed consent form included the participant's agreement to the disclosure of data for EEG research while ensuring their names would not be disclosed. It also explained the purpose and context of the study, the data collection process, and the rights of the participants, including their right to discontinue participation at any time.

Participant ID	Age	Sex	RSVP-accuracy (%)	low-speed-accuracy (%)
sub-01	22	M	98.75	98.88
sub-02	23	M	99.40	99.25
sub-03	22	M	95.75	98.55
sub-04	24	M	97.13	98.88
sub-05	33	M	98.20	99.40
sub-06	24	M	98.13	97.63
sub-07	24	F	99.00	99.75
sub-08	22	F	99.00	98.88
sub-09	24	M	97.45	98.90
sub-10	26	F	98.75	98.75
sub-11	28	M	99.08	99.42
sub-12	25	M	98.25	98.50
sub-13	23	M	91.50	93.50
sub-14	24	M	98.80	98.60
sub-15	26	M	99.15	98.35
sub-16	22	M	99.25	99.25
sub-17	24	M	98.25	99.05
sub-18	24	M	97.50	98.17
sub-19	22	M	95.80	98.45
sub-20	22	M	98.53	99.30
sub-21	23	M	98.38	99.00
sub-22	27	M	98.08	99.00
sub-23	23	M	98.20	99.00
sub-24	22	M	99.20	99.65
sub-25	23	M	99.25	98.50
sub-26	22	M	97.75	98.50
sub-27	22	M	98.75	96.50
sub-28	23	M	98.80	98.75
sub-29	21	M	98.90	99.00
sub-30	24	M	98.67	94.50
sub-31	22	F	98.75	98.25
sub-32	21	M	91.50	96.00

Table 2. Details of the participants.

The experiment was conducted in accordance with the Declaration of Helsinki. The ethical committee of the Institute of Automation at the Chinese Academy of Sciences approved this study, with the institution number: EC-20230427-1014 and the approval number: IA21-2410-020201.

Experimental design. *General design.* Before the recording, the participants in the study were given a thorough explanation of the experimental methods and procedures. Additionally, all subjects completed two informed consent forms. The experiment was conducted in a quiet and closed laboratory, and the subjects were seated on a chair located a certain distance away from the LCD monitor. To ensure reliability, one experimenter was responsible for supervising the entire process.

The MSS dataset consisted of two types of EEG paradigms: the RSVP experiment[29] and the low-speed experiment[6]. Each paradigm consisted of four runs, and the two paradigms alternated. This means that the subjects completed an RSVP run first and then conducted a low-speed run. Each RSVP run lasted approximately 14 minutes with a break at around 7 minutes, while each low-speed run lasted around 6 minutes. Participants were allowed to take sufficient rests between each run to maintain their physical and mental condition for recording high-quality signals.

A total of 32 subjects participated in the study and data for 100 sessions was recorded. Among them, 14 subjects (sub02, sub03, sub05, sub09, sub14, sub15, sub17, sub19, sub20, sub23, sub24, sub26, sub28, sub29) completed five sessions, 4 subjects (sub11, sub18, sub21, sub22) completed three sessions, 4 subjects (sub01, sub04, sub06, sub08) conducted two different sessions, and 10 subjects (sub07, sub10, sub12, sub13, sub16, sub25, sub27, sub30, sub31, sub32) conducted one session. Most subjects conducted different sessions with an interval of more than a week. The detailed experiment schedule of all subjects is provided in the "derivatives/annotations/Experiment_schedule.csv".

The number of sessions varies because we prioritize participants' willingness to engage in the collection process. We do not pressure participants to complete all sessions if they prefer not to continue. When recruiting participants, we informed them that they could expect a maximum of five sessions. However, there is typically a one-week interval between each session, which means it usually takes over a month for a participant to complete

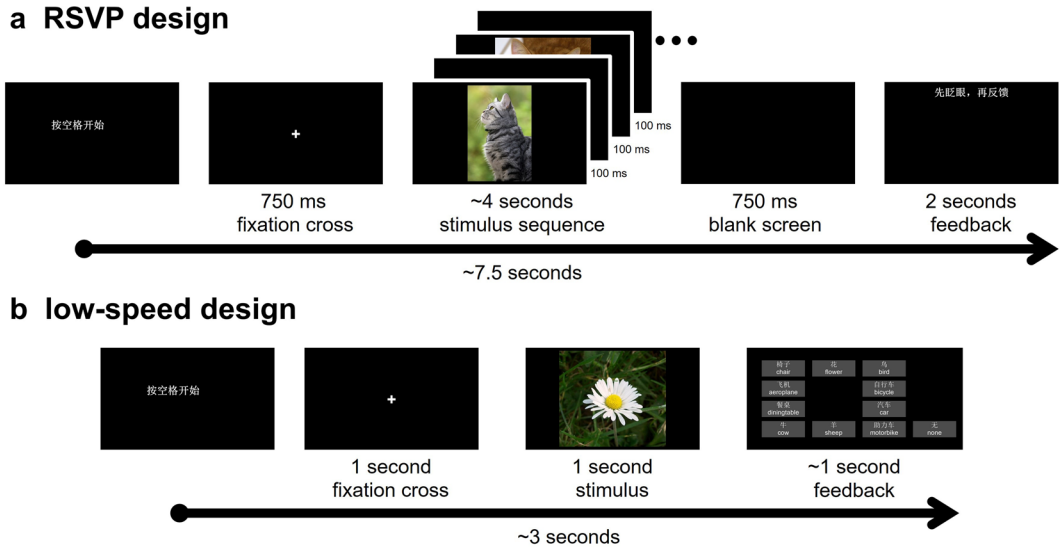

Fig. 1 Experimental procedure of (**a**) RSVP and (**b**) low-speed design.

all five sessions. During this time, participants have the freedom to decide whether or not to continue, and they will still receive payment even if they choose to withdraw partway through.

RSVP experiment. During the RSVP experiment, the participants were shown images at a rate of 5 Hz, and each run consisted of 2,000 trials. The RSVP design utilized brief stimulus durations to reduce eye movements and increase the number of trials. This paradigm has been proven to be effective in reflecting detailed visual information according to studies[29–32]. There were 20 image categories, with 100 images in each category, making up the 2,000 stimuli. The 100 images in each category were further divided into five image sequences, resulting in 100 image sequences per run. Each sequence was composed of 20 images from the same class, and the 100 sequences were presented in a pseudo-random order.

After every 50 sequences, there was a break for the participants to rest, and the recording system was checked by the experimenter. Each rapid serial sequence lasted approximately 7.5 seconds, starting with a 750ms blank screen with a white fixation cross, followed by 20 or 21 images presented at 5 Hz with a 50% duty cycle. The sequence ended with another 750ms blank screen.

After the rapid serial sequence, there was a 2-second interval during which participants were instructed to blink and then report whether a special image appeared in the sequence using a keyboard. During each run, 20 sequences were randomly inserted with additional special images at random positions to enhance the subjects' attention and decrease the chance of eye blinks. The special images are logos for brain-computer interfaces, which are placed in "stimuli/special". The procedure of RSVP experiment is shown in Fig. 1a.

Low-speed experiment. During the low-speed experiment, each run consisted of 100 trials, with 1 second per image for a slower paradigm. This design has been used in previous works[6,18,20] that needed longer recording times than RSVP to collect non-overlapping signals. The 100 stimuli were presented in a pseudo-random order and included 20 image categories, each containing 5 images. A break was given to the participants after every 20 images for them to rest.

Each image was displayed for 1 second and was followed by 11 choice boxes (1 correct class box, 9 random class boxes, and 1 reject box), as shown in Fig. 1b. Participants were required to select the correct class of the displayed image using a mouse to increase their engagement. To reduce search time and prevent participants from remembering the position of choice boxes, only 10 class boxes were displayed. A reject box was also included in the lower right corner of the screen in case participants were absent-minded. After the selection, a white fixation cross was displayed for 1 second in the centre of the screen to remind participants to pay attention to the upcoming task. The procedure of low-speed experiment is shown in Fig. 1b.

Stimuli. The study utilized stimuli from two image databases: PASCAL[28] and ImageNet[22], both renowned for their naturalistic images and extensive use in computer vision research. The PASCAL VOC Challenges 2007 database is publicly available at http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html. The PASCAL VOC Challenges 2012 database is publicly available at http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html. The ImageNet database can be accessed at https://image-net.org/download.php.

Since the PASCAL image library consists of a 20-category dataset, our selection of stimuli was based on these 20 semantic categories. For each category, we manually selected 500 high-quality images that featured clear subjects without any background or watermark interference. However, due to a shortage of high-quality images in the PASCAL dataset, we also sourced images from the ImageNet dataset that matched the PASCAL categories. We carefully screened these related images to complete our stimuli image library of 500 images for each of the

Final categories	Categories of the PASCAL	Categories of the ImageNet
aeroplane	aeroplane	—
bicycle	bicycle	n02835271, n03792782
bird	bird	—
boat	boat	n03095699, n04483307
bottle	bottle	n03062245, n03983396
bus	bus	n04487081
car	car	—
cat	cat	—
chair	chair	n02791124, n03376595, n04099969
cow/bull	cow	n02403003, n02408429, n02410509
diningtable	diningtable	n03201208
dog	dog	—
flower	—	n11939491
horse	horse	n02389026
motorbike	motorbike	n03785016, n03791053
person	person	—
sheep/goat	sheep	n02412080, n02415577, n02417914, n02422106, n02422699, n02423022
sofa	sofa	n04344873
train	train	—
tvmonitor	tvmonitor	n03782006, n04152593

Table 3. Categories of the stimuli.

Event	EventID	Description
Stimulus index	1-20	The index of stimuli in a sequence
Special stimulus	99	The index of the special stimuli
Correct	252	Subject's response is correct
Wrong	254	Subject's response is wrong

Table 4. The event information during the RSVP experiment.

20 semantic categories. Additionally, the original "pottedplant" class in PASCAL was often confused with other categories due to its lack of clarity. Therefore, we used the "n11939491" class from ImageNet as a replacement.

In summary, we compiled a candidate stimuli set from 19 PASCAL categories and 27 ImageNet categories, ultimately establishing the final 20 semantic categories, as shown in Table 3. The final collection consists of 10,000 images, with 500 images per category, selected through a rigorous refinement process. Initially, three experimenters independently chose images, and only those selected by all three were retained. For any categories with insufficient images, five additional experimenters contributed selections. After the initial selection, the chosen images underwent careful review to eliminate any inappropriate ones and to incorporate new images as needed. This review process was conducted five times to ensure that only the most suitable images were included.

Data acquisition. During the EEG recording session, the raw signals were captured using the 128-channel Quik-Cap (Compumedics Neuroscan) for the SynAmps 2/RT and Neuvo amplifier at a sampling rate of 1,000 Hz. A total of 122 wet electrodes, which are Ag/AgCl electrodes with conductive gel, were placed according to the international 10-20 system. Additionally, there were 2 integrated bipolar leads for vertical and horizontal EOG (VEOG, HEOG). Throughout the experiment, the impedance was maintained below 10 kOhm. The EEG amplifier defaulted the reference and ground electrode. To synchronise the EEG signals with the experimental procedure, triggers were sent to the EEG recording system with a parallel port whenever events occurred. The event information during the experiment is available in Tables 4 and 5.

Preprocessing. The EEG signals were pre-processed using the MNE package[33], version 1.3.1, with Python 3.9.16. The codes for data pre-processing were provided along with the dataset. First, the data sampling rate was maintained at 1,000 Hz. Second, a FIR bandpass filter ranging from 0.1 to 100 Hz was applied to the signals using the MNE "filter" functions. Third, a notch filter was set up to eliminate 50-Hz power frequency signals. Finally, epochs were created for each trial ranging from 0 to 500 ms relative to stimulus onset. No further preprocessing or artefact correction methods were applied in technical validation. However, researchers may want to consider widely used preprocessing steps such as baseline correction or eye movement correction. After the preprocessing, each session resulted in two matrices: RSVP EEG data matrix of shape (8,000 image conditions 122 EEG channels 500 EEG time points) and low-speed EEG data matrix of shape (400 image conditions 122 EEG channels 500 EEG time points).

Event	EventID	Description
Stimulus index	1-100	The index of stimuli in a run
Correct	252	Subject selected the correct class
Wrong	254	Subject selected the wrong class

Table 5. The event information during the low-speed experiment.

Data Records

The data collection is available on the OpenNeuro platform at https://openneuro.org/datasets/ds005589 database[34]. The data collection was organized according to the Brain-Imaging-Data-Structure (BIDS) Specification version 1.10.0. As shown in Fig. 2, our data collection includes data description files, raw EEG data collected for each subject in the "sub-*" folders, stimuli images in the "stimuli" folder, code for preprocessing and technical validation in the "code" folder, preprocessed EEG data in the "derivatives" folder. More details about these folders are provided below.

Stimuli Folder. The visual stimuli can be found in the "stimuli" folder (Fig. 2d). Rather than redistributing the public images, we have listed the names of the stimulus images along with their respective folders and sources in the public dataset. Links to access the public dataset are also provided in the section **Stimuli**. Researchers who require these images can obtain them through the links and image names we have provided.

In the "stimuli" folder, "⟨category⟩.csv" includes three columns: "image_file_name", "dataset", and "directory". The "image_file_name" column specifies the filename of the stimuli. The "dataset" column indicates the source datasets of the stimuli. The "directory" column shows the directory of the stimuli within the source dataset.

Raw data folder. Each "sub-⟨subID⟩" folder contains one to five "ses-⟨sesID⟩" subfolders (Fig. 2c). Each "ses-⟨sesID⟩" folder contains raw EEG data and experimental information related to RSVP and low-speed paradigms. Each experimental signal is originally recorded as three files: "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_eeg.cdt", "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_eeg.cdt.ceo", "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_eeg.cdt.dpa". Additionally, the "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_events.tsv" file contains task events of each run. The onset for each trial is provided in the first column of this events file, while the event number for each trial is located in the last column.

The specific stimulus image for each trial can be found in the "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_image.csv" file and the "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_record.csv" file. The "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_image.csv" file contains the file paths of stimuli for each run. The "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_record.csv" file provides a detailed record of each run, including the columns: "imgNO", "img_path", "img_class", "imgshow_time", and "imgshow_localtime_str". The "imgNO", "img_path", and "img_class" denote the stimuli, while the "imgshow_time", and "imgshow_localtime_str" indicate the onset time of the stimuli.

For the low-speed paradigm, the "sub-⟨subID⟩_ses-⟨sesID⟩_task-lowSpeed_run-⟨runID⟩_record.csv" file contains additional information, including: "choice_class_str", "choice_class", "select_time", "select_localtime_str", "flag_correct", and "right_number" columns. The "choice_class_str" and "choice_class" denote the category selected by the participants. The "flag_correct" indicates whether the participant made the correct choice, while the "right_number" denotes the total number of correct selections made by the participants. In the case of the RSVP paradigm, the "sub-⟨subID⟩_ses-⟨sesID⟩_task-rsvp_run-⟨runID⟩_specialImage.csv" file contains the insert position and file paths of special images for each run.

Derivatives folder. The "derivatives" folder contains the preprocessed EEG data for each subject in the "preprocessed_data" subfolder (Fig. 2e). The preprocessed data include the epoched EEG trials sampled at a rate of 1,000 Hz, along with corresponding label files. The "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_1000Hz.npy" file contains the preprocessed data arranged in the format of (number of trials time points number of channels). The "sub-⟨subID⟩_ses-⟨sesID⟩_task-⟨task⟩_run-⟨runID⟩_1000Hz.json" file contains the label of trials that range from 0 to 19.

The "derivatives" folder also contains the corrupted raw EEG data in the "corrupted_raw_data" subfolder and additional information in the "annotations" subfolder. The "corrupted_raw_data" subfolder contains the raw data file "sub-26_ses-05_task-rsvp_run-03_eeg.cdt". This file includes the EEG data for the entire run; however, its associated file, "sub-26_ses-05_task-rsvp_run-03_eeg.cdt.ceo" was not recorded due to equipment failure. As a result, it cannot be used for supervised classification experiments. We included this run in the release for completeness, and it can still be utilized for unsupervised experiments. The "annotations" subfolder contains dataset information, demos of two paradigms, the schedule for each session, source categories for PASCAL and ImageNet, participants' age and sex, and the meaning of eventID.

Code folder. The "code" folder contains the code for EEG paradigms, preprocessing and technical validation (Fig. 2b). Below is a brief explanation of the stimulus codes used in the EEG paradigms within the dataset.

RSVP Paradigm. Before the start of each run, the functions "rand_sample_multiclass_consecutive" and "rand_special_img_seq_and_pos" are used to randomly select the positions of the stimulus images and the

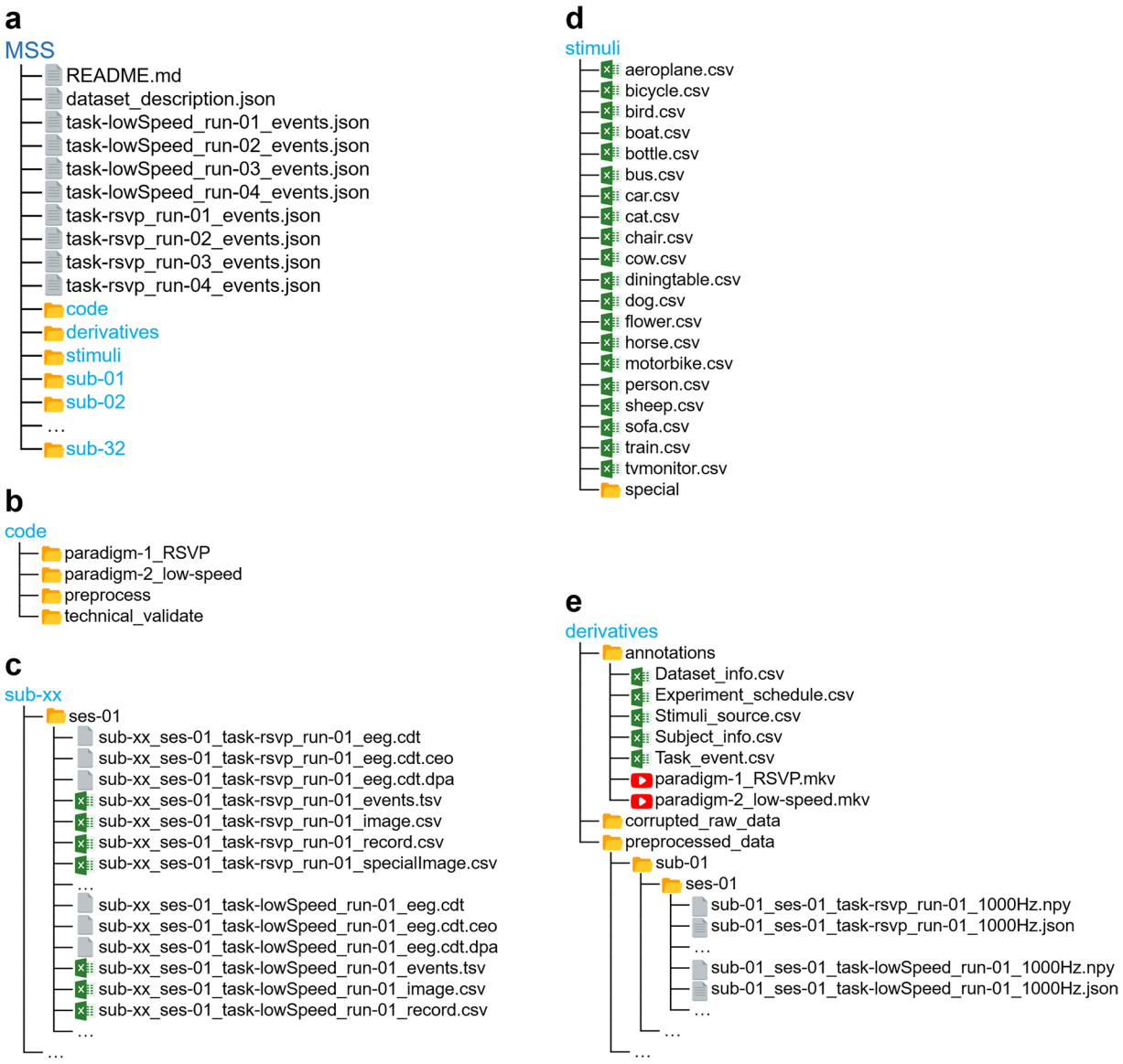

Fig. 2 The file structure of MSS. (**a**) The overall directory structure of MSS. (**b**) Content of code directory. (**c**) Content of the raw data directory. (**d**) Content of stimuli directory. (**e**) Content of the derivatives directory, including annotations, corrupted raw data of a single run, and preprocessed data of 32 subjects.

special images for participant attention monitoring, respectively. The function "save_filepath_per_experiment_consecutive" saves the stimulus images for the experiments to the subfolder "img_info_RSVP", while "save_special_img_seq_and_pos" is responsible for saving the special stimulus images to the subfolder "special_img_info_RSVP".

Once the required stimuli are sequenced, we utilize components from the PsychoPy library to create a display screen for the paradigm. Images are preloaded into "pic_all_win" before each sequence display, and stimuli are shown one by one. The sequence number of the displayed image and the participant's selection in the attention detection task are recorded using the "imgshow_serialwrite" and "select_serialwrite" functions, respectively. Additionally, during the experiment, the function "save_exp_info" logs the experimental records into the "exp_info_RSVP" subfolder.

For further details, please refer to our code.

Low-Speed Paradigm. Before each run of the experiment starts, the function "rand_sample_multiclass" randomly selects stimulus images. The function "save_filepath_per_experiment" saves these stimulus images into the "img_info_Class_Select" subfolder.

After the stimuli are arranged, we again use components from the PsychoPy library to build a display screen. Images are preloaded into "pic_all_win", and they are shown one by one. The serial number of the displayed image and the participant's selection in the attention detection task are recorded using the "imgshow_serialwrite" and "select_serialwrite" functions, respectively. The function "rand_card_position" generates ten choices

Subject	Session-1	Session-2	Session-3	Session-4	Session-5	Average
sub-01	11.45	9.80	—	—	—	**10.62 ± 0.83**
sub-02	11.60	13.65	10.10	9.30	14.35	**11.80 ± 1.95**
sub-03	6.95	8.75	7.50	6.30	6.60	**7.22 ± 0.86**
sub-04	12.45	9.85	—	—	—	**11.15 ± 1.30**
sub-05	8.85	11.50	10.20	11.75	10.00	**10.46 ± 1.06**
sub-06	7.80	8.60	—	—	—	**8.20 ± 0.40**
sub-07	13.40	—	—	—	—	**13.40 ± 0.00**
sub-08	7.20	7.80	—	—	—	**7.50 ± 0.30**
sub-09	8.10	9.70	10.95	7.30	8.95	**9.00 ± 1.26**
sub-10	9.60	—	—	—	—	**9.60 ± 0.00**
sub-11	10.85	12.95	11.65	—	—	**11.82 ± 0.87**
sub-12	9.70	—	—	—	—	**9.70 ± 0.00**
sub-13	11.80	—	—	—	—	**11.80 ± 0.00**
sub-14	10.55	11.30	7.95	10.90	10.55	**10.25 ± 1.18**
sub-15	8.75	9.20	8.95	7.90	6.50	**8.26 ± 0.98**
sub-16	7.60	—	—	—	—	**7.60 ± 0.00**
sub-17	10.30	12.70	11.75	10.30	8.95	**10.80 ± 1.30**
sub-18	14.05	11.35	8.10	—	—	**11.17 ± 2.43**
sub-19	13.10	10.80	8.00	8.85	7.35	**9.62 ± 2.09**
sub-20	14.45	13.25	13.96	16.25	8.80	**13.34 ± 2.48**
sub-21	11.10	8.60	6.00	—	—	**8.57 ± 2.08**
sub-22	13.65	9.00	13.50	—	—	**12.05 ± 2.16**
sub-23	14.15	8.60	8.40	10.05	12.35	**10.71 ± 2.22**
sub-24	11.10	11.15	11.20	10.50	11.45	**11.08 ± 0.31**
sub-25	12.40	—	—	—	—	**12.40 ± 0.00**
sub-26	11.75	9.95	7.55	9.85	8.05	**9.43 ± 1.50**
sub-27	7.55	—	—	—	—	**7.55 ± 0.00**
sub-28	16.40	13.95	14.75	16.85	13.70	**15.13 ± 1.28**
sub-29	10.05	9.75	9.60	9.45	10.75	**9.92 ± 0.46**
sub-30	6.45	—	—	—	—	**6.45 ± 0.00**
sub-31	6.65	—	—	—	—	**6.65 ± 0.00**
sub-32	6.50	—	—	—	—	**6.50 ± 0.00**

Table 6. Test accuracies(%) of 100 sessions of RSVP paradigm.

before each selection. Additionally, the "save_exp_info" function logs the experimental records in the "exp_info_Class_Select" subfolder during the experiment.

For more details, please see our code.

Technical Validation

To examine the signal quality, we first assessed whether the subjects completed the attention task. We then used a fully connected (FC) model to verify the signal quality and perform the offline signal semantic classification task. It is important to highlight that this dataset consists of single-trial signals, indicating that there are no multiple repetitions of the same image. Consequently, this dataset is more appropriate for experiments aimed at semantic classification rather than for classifying visually similar objects.

Online analysis. Before each acquisition run, the experimenter checked the EEG signal waveform, ensured the electrode cap was stable, and verified the electrode resistance. During the signal acquisition process, the experimenter monitored the quality of the collected signal and periodically checked whether the labels were marked correctly. Subjects will be reminded to avoid blinking, jaw movements, or excessive lower limb movements during key marker periods. Despite these precautions, it is important to note that there could inevitably be noisy channels and bad trials in the collected signals.

Behavior results. For the RSVP paradigm, we adopted the special image detection task[10] and recorded the detection success rate of subjects. For the low-speech paradigm, we designed a task for picture classification and recorded the success rate of classification. As shown in Table 2, the accuracies of most subjects reach more than 97%. The average accuracy of the RSVP detection task for all subjects was 97.93% with a variance of 1.87%, and the average accuracy of the low-speed classification task for all subjects was 98.38% with a variance of 1.37%. These results indicate that the concentration of subjects during the data collection process was ensured. Additionally, the experiment information files record the success or failure of each trial. Researchers may consider deleting the failed trials during the study.

Subject	Session-1	Session—2	Session—3	Session—4	Session—5	Average
sub-01	8.0	9.0	—	—	—	**8.5 ± 0.5**
sub-02	10.0	8.0	9.0	9.0	16.0	**10.4 ± 2.9**
sub-03	10.0	7.0	11.0	11.0	10.0	**9.8 ± 1.5**
sub-04	14.0	11.0	—	—	—	**12.5 ± 1.5**
sub-05	8.0	12.0	13.0	17.0	14.0	**12.8 ± 2.9**
sub-06	9.0	13.0	—	—	—	**11.0 ± 2.0**
sub-07	11.0	—	—	—	—	**11.0 ± 0.0**
sub-08	6.0	11.0	—	—	—	**8.5 ± 2.5**
sub-09	14.0	8.0	8.0	13.0	10.0	**10.6 ± 2.5**
sub-10	8.0	—	—	—	—	**8.0 ± 0.0**
sub-11	12.0	14.0	12.0	—	—	**12.7 ± 0.9**
sub-12	11.0	—	—	—	—	**11.0 ± 0.0**
sub-13	14.0	—	—	—	—	**14.0 ± 0.0**
sub-14	10.0	6.0	10.0	10.0	14.0	**10.0 ± 2.5**
sub-15	11.0	12.0	9.0	7.0	10.0	**9.8 ± 1.7**
sub-16	10.0	—	—	—	—	**10.0 ± 0.0**
sub-17	10.0	12.0	10.0	12.0	9.0	**10.6 ± 1.2**
sub-18	13.0	10.0	7.0	—	—	**10.0 ± 2.4**
sub-19	13.0	10.0	10.0	9.0	11.0	**10.6 ± 1.4**
sub-20	12.0	9.0	17.0	13.0	12.0	**12.6 ± 2.6**
sub-21	12.0	13.0	12.0	—	—	**12.3 ± 0.5**
sub-22	17.0	10.0	10.0	—	—	**12.3 ± 3.3**
sub-23	11.0	7.0	11.0	11.0	15.0	**11.0 ± 2.5**
sub-24	8.0	10.0	8.0	10.0	9.0	**9.0 ± 0.9**
sub-25	11.0	—	—	—	—	**11.0 ± 0.0**
sub-26	7.0	10.0	14.0	11.0	11.0	**10.6 ± 2.2**
sub-27	10.0	—	—	—	—	**10.0 ± 0.0**
sub-28	14.0	13.0	9.0	12.0	10.0	**11.6 ± 1.9**
sub-29	8.0	9.0	13.0	16.0	10.0	**11.2 ± 2.9**
sub-30	10.0	—	—	—	—	**10.0 ± 0.0**
sub-31	10.0	—	—	—	—	**10.0 ± 0.0**
sub-32	10.0	—	—	—	—	**10.0 ± 0.0**

Table 7. Test accuracies(%) of 100 sessions of low-speed paradigm.

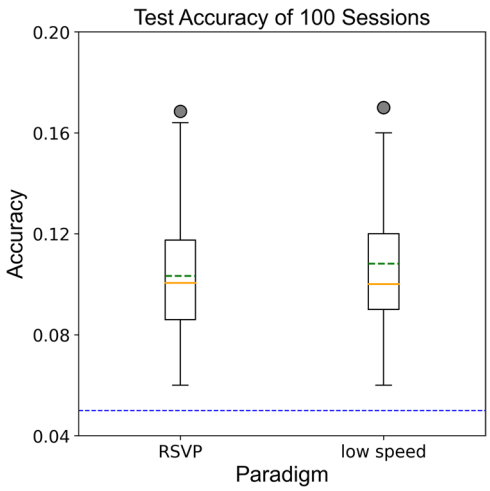

Fig. 3 Classification accuracy of 100 sessions of RSVP and low-speed paradigms. The blue dash-dotted line indicates 5% accuracy chance level. The orange solid line indicates median accuracy. The green dash-dotted line indicates average accuracy.

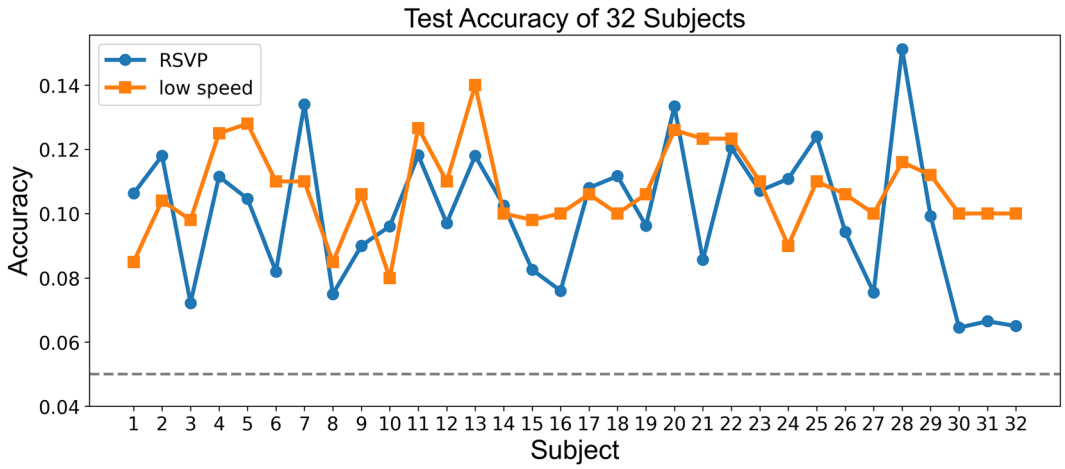

Fig. 4 Classification accuracy of 32 subjects of RSVP and low-speed paradigms. The accuracies were averaged over sessions of each subject. The grey dash-dotted line indicates 5% accuracy chance level.

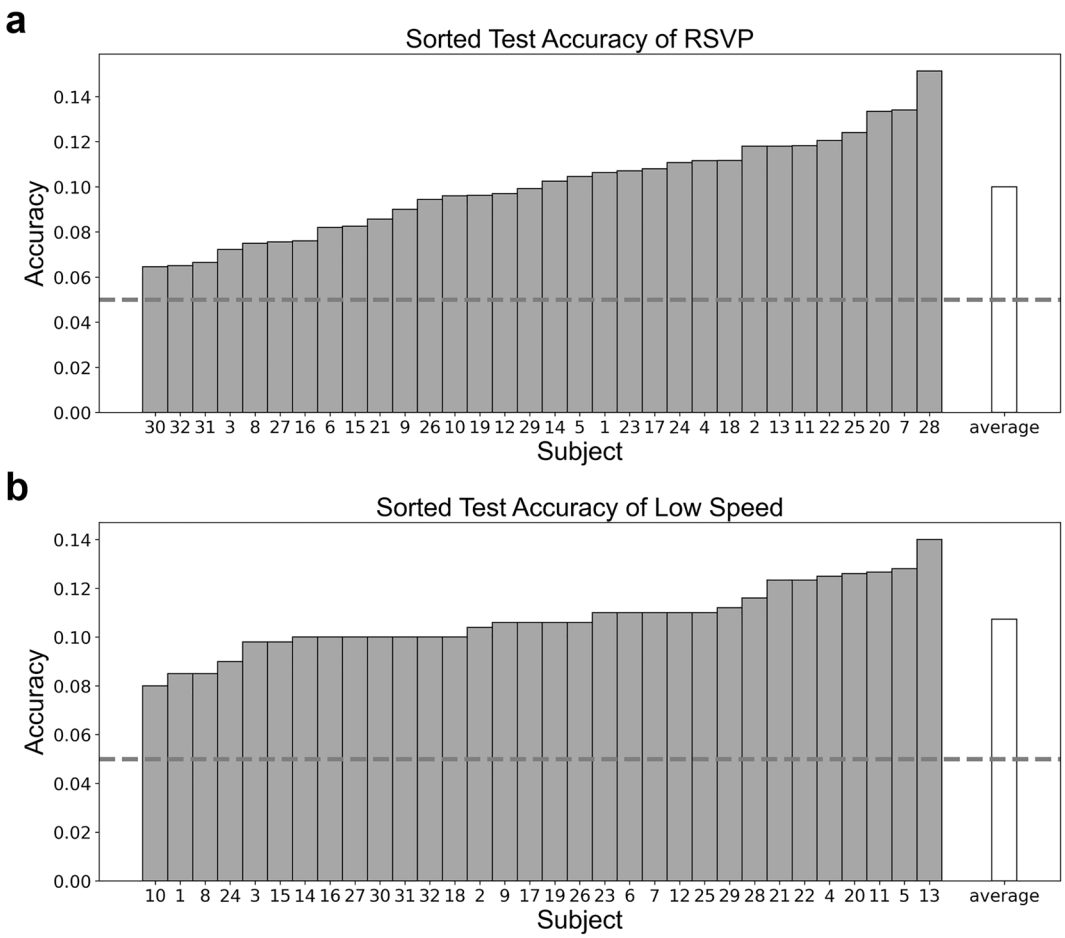

Fig. 5 Sorted classification accuracy of 32 subjects of (**a**) RSVP and (**b**) low-speed paradigms. The accuracies were averaged over sessions of each subject.

Within-session classification. To evaluate whether the dataset can support finer-grained visual recognition, we conducted a within-session classification analysis in three parts. First, we analyzed the semantic classification results from 100 sessions. Next, we assessed the variability in results across 32 subjects. Finally, we assessed the variability in results across different sessions.

Tables 6 and 7, along with Fig. 3, report the classification accuracy in each session of all subjects. We performed a cross-time verification to evaluate performance. For each session, the first three runs were used for

Subject	Session-1	Session-2	Session-3	Session-4	Session-5	Average
sub-02	13.98	15.53	12.21	11.29	12.70	**13.14 ± 1.47**
sub-03	5.90	6.78	6.73	6.02	6.78	**6.44 ± 0.39**
sub-05	9.54	9.61	10.54	10.09	7.14	**9.38 ± 1.18**
sub-09	8.70	10.03	10.16	7.22	8.81	**8.99 ± 1.07**
sub-14	10.24	13.06	9.95	12.80	11.96	**11.60 ± 1.29**
sub-15	10.25	10.78	10.69	6.68	8.30	**9.34 ± 1.61**
sub-17	9.94	13.08	13.31	10.82	8.16	**11.06 ± 1.94**
sub-19	10.51	8.51	7.20	10.66	8.44	**9.06 ± 1.33**
sub-20	16.54	15.57	13.64	16.81	10.93	**14.70 ± 2.19**
sub-23	12.29	9.49	11.76	10.09	11.28	**10.98 ± 1.04**
sub-24	12.91	12.54	13.50	7.81	12.88	**11.93 ± 2.08**
sub-26	10.39	10.80	10.90	8.03	8.95	**9.81 ± 1.13**
sub-28	18.52	12.57	17.84	14.19	16.09	**15.84 ± 2.22**
sub-29	7.91	8.38	8.99	7.85	8.53	**8.33 ± 0.42**

Table 8. Cross-session classification accuracy(%) across five sessions of RSVP paradigm. The table shows the accuracies of fourteen subjects who has finished five sessions.

Subject	Session-1	Session-2	Session-3	Session-4	Session-5	Average
sub-02	10.50	12.50	12.00	11.50	10.75	**11.45 ± 0.75**
sub-03	8.00	9.75	8.75	8.75	8.25	**8.70 ± 0.60**
sub-05	10.75	10.25	12.00	11.00	10.00	**10.80 ± 0.70**
sub-09	10.50	9.50	10.75	9.00	10.25	**10.00 ± 0.65**
sub-14	8.50	11.50	10.50	9.50	11.00	**10.20 ± 1.08**
sub-15	11.25	9.75	11.25	10.00	10.00	**10.45 ± 0.66**
sub-17	13.75	15.00	14.25	10.00	11.50	**12.90 ± 1.86**
sub-19	12.00	9.00	9.25	8.75	10.50	**9.90 ± 1.21**
sub-20	16.25	14.00	12.50	14.75	10.25	**13.55 ± 2.05**
sub-23	13.50	10.50	11.75	10.75	10.50	**11.40 ± 1.15**
sub-24	9.25	9.00	12.00	10.50	12.50	**10.65 ± 1.41**
sub-26	11.50	11.25	10.75	9.00	12.00	**10.90 ± 1.03**
sub-28	17.00	15.75	14.50	12.75	16.75	**15.35 ± 1.57**
sub-29	10.00	10.00	11.00	11.25	10.25	**10.50 ± 0.52**

Table 9. Cross-session classification accuracy(%) across five sessions of low-speed paradigm. The table shows the accuracies of fourteen subjects who has finished five sessions.

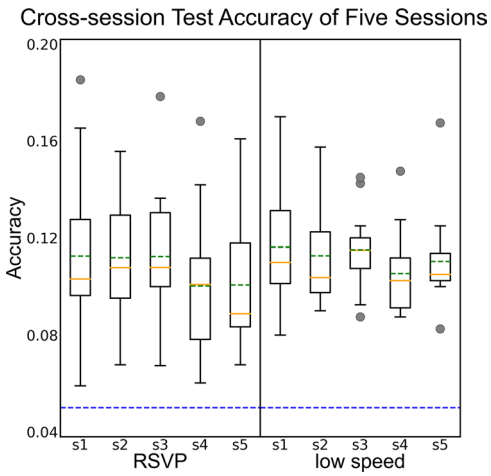

Fig. 6 Cross-session classification accuracy across five sessions of RSVP and low-speed paradigms. The blue dash-dotted line indicates 5% accuracy chance level. The orange solid line indicates median accuracy. The green dash-dotted line indicates average accuracy.

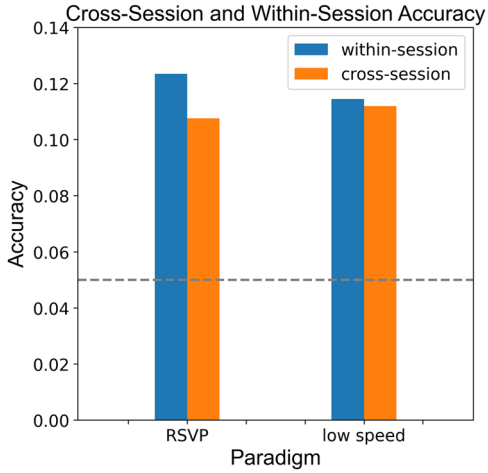

Fig. 7 Cross-session classification accuracy compared with within-session classification accuracy. The accuracies were averaged over 70 sessions of fourteen subjects who has finished five sessions. The grey dash-dotted line indicates 5% accuracy chance level.

Subject	RSVP	Low Speed
sub-01	9.78	11.88
sub-02	9.28	9.95
sub-03	6.99	9.40
sub-04	10.21	14.75
sub-05	8.75	11.15
sub-06	7.76	10.62
sub-07	9.54	12.75
sub-08	8.07	10.38
sub-09	8.07	11.40
sub-10	8.55	10.25
sub-11	9.74	12.42
sub-12	10.26	14.00
sub-13	9.46	11.25
sub-14	9.69	9.75
sub-15	7.07	10.50
sub-16	8.05	8.75
sub-17	8.85	12.00
sub-18	9.52	9.42
sub-19	8.08	9.40
sub-20	9.71	12.15
sub-21	8.78	10.33
sub-22	10.11	12.00
sub-23	9.03	10.05
sub-24	8.99	11.05
sub-25	9.32	13.50
sub-26	8.29	10.35
sub-27	7.64	11.25
sub-28	9.42	13.95
sub-29	6.98	8.95
sub-30	5.76	6.75
sub-31	6.33	7.75
sub-32	6.54	9.25

Table 10. Cross-subject classification accuracy(%) across 32 subjects of RSVP and low-speed paradigms.

training, and the last run was used for validation. The results indicated that most sessions achieved classification accuracy rates of approximately 10%, significantly higher than the 5% accuracy of random classification.

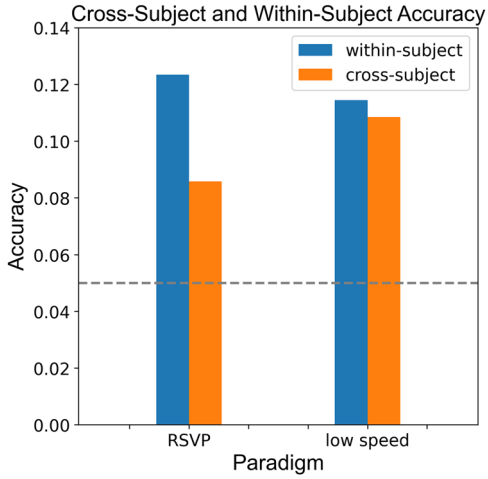

Fig. 8 Cross-subject classification accuracy compared with within-subject classification accuracy. The accuracies were averaged over fourteen subjects who has finished five sessions. The grey dash-dotted line indicates 5% accuracy chance level.

However, there was variability in accuracy rates across different sessions, with some achieving rates over 16%, while others were close to random classification levels.

Figure 4 illustrates the average classification accuracies averaged over sessions for each subject. The performance varied widely among subjects, indicating that some subjects demonstrated better adaptability to EEG-based BCIs, while others exhibited potential BCI illiteracy[35] in this visual-stimulation paradigm. Furthermore, Fig. 5 displays the sorted accuracies of the 32 subjects arranged in order of increasing accuracy. Notably, Subject 28 performed best in the RSVP paradigm, while Subject 13 excelled in the low-speed paradigm.

From Tables 6 and 7, it can be observed that some subjects showed consistency across sessions, whereas others exhibited significant variability. For instance, Subjects 20 and 23 displayed considerable variability in both the RSVP and low-speed paradigms, while Subject 24 demonstrated more stable performance.

Cross-session classification. Tables 8 and 9, along with Figs. 6 and 7, present the cross-session classification accuracy across five sessions. It can be observed that the cross-session accuracy is significantly lower than the within-session accuracy observed in the RSVP paradigm. This indicates that the variability of signals over time presents challenges for cross-session classification.

Cross-subject classification. Table 10 and Fig. 8 present the cross-subject classification accuracy. It can be observed that the cross-subject accuracy is significantly lower than the within-subject accuracy observed in the RSVP paradigm. This indicates that the variability of signals among different subjects creates challenges for cross-subject classification.

In summary, the behavior results showed that participants maintained their attention throughout the experiment. The classification analysis indicated that object categories could be effectively identified from EEG signals using the deep-learning model. Furthermore, there was variability in classification accuracy among participants and across different sessions. Our dataset is well-designed for validating the generalization capabilities of classification methods. Overall, the technical validation results support the dataset's validity and reliability.

Code availability

All codes for the experimental design, preprocessing, and technical validation are available at https://github.com/xuesn/EEGDataset. Preprocessing was performed using MNE version 1.3.1 (https://mne.tools/stable/index.html).

Received: 11 November 2024; Accepted: 17 March 2025;
Published online: 19 April 2025

References

1. Cichy, R. M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and time. *Nature neuroscience* **17**, 455–462 (2014).
2. Grill-Spector, K. The neural basis of object perception. *Current opinion in neurobiology* **13**, 159–166 (2003).
3. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *nature* **381**, 520–522 (1996).
4. Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
5. Felsen, G. & Dan, Y. A natural approach to studying vision. *Nature neuroscience* **8**, 1643–1646 (2005).
6. Kaneshiro, B., Perreau Guimaraes, M., Kim, H.-S., Norcia, A. M. & Suppes, P. A representational similarity analysis of the dynamics of object processing using single-trial eeg classification. *Plos one* **10**, e0135697 (2015).
7. Brandman, T. & Peelen, M. V. Interaction between scene and object processing revealed by human fmri and meg decoding. *Journal of Neuroscience* **37**, 7700–7710 (2017).
8. Gong, Z. *et al*. A large-scale fmri dataset for the visual processing of naturalistic scenes. *Scientific Data* **10**, 559 (2023).

9. Telesford, Q. K. *et al*. An open-access dataset of naturalistic viewing using simultaneous eeg-fmri. *Scientific Data* **10**, 554 (2023).
10. Gifford, A. T., Dwivedi, K., Roig, G. & Cichy, R. M. A large and rich eeg dataset for modeling human visual object recognition. *NeuroImage* **264**, 119754 (2022).
11. Grootswagers, T., Zhou, I., Robinson, A. K., Hebart, M. N. & Carlson, T. A. Human eeg recordings for 1,854 concepts presented in rapid serial visual presentation streams. *Scientific Data* **9**, 3 (2022).
12. Won, K., Kwon, M., Ahn, M. & Jun, S. C. Eeg dataset for rsvp and p300 speller brain-computer interfaces. *Scientific Data* **9**, 388 (2022).
13. Xu, M., Han, J., Wang, Y., Jung, T.-P. & Ming, D. Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent p300 and ssvep features. *IEEE Transactions on Biomedical Engineering* **67**, 3073–3082 (2020).
14. Murphy, B. *et al*. Eeg decoding of semantic category reveals distributed representations for single concepts. *Brain and language* **117**, 12–22 (2011).
15. Ma, L., Marshall, P. J. & Wright, W. G. The impact of external and internal focus of attention on visual dependence and eeg alpha oscillations during postural control. *Journal of NeuroEngineering and Rehabilitation* **19**, 81 (2022).
16. Chai, M. T. *et al*. Exploring eeg effective connectivity network in estimating influence of color on emotion and memory. *Frontiers in neuroinformatics* **13**, 66 (2019).
17. Kalantari, S. *et al*. Evaluating the impacts of color, graphics, and architectural features on wayfinding in healthcare settings using eeg data and virtual response testing. *Journal of Environmental Psychology* **79**, 101744 (2022).
18. Ahmed, H., Wilbur, R. B., Bharadwaj, H. M. & Siskind, J. M. Object classification from randomized eeg trials. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3845–3854 (2021).
19. Zhu, S., Ye, Z., Ai, Q. & Liu, Y. Eeg-imagenet: An electroencephalogram dataset and benchmarks with image visual stimuli of multi-granularity labels. *arXiv preprint arXiv:2406.07151* (2024).
20. Spampinato, C. *et al* Deep learning human mind for automated visual classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6809–6817 (2017).
21. Palazzo, S. *et al*. Decoding brain representations by multimodal learning of neural activity and visual features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**, 3833–3849 (2020).
22. Deng, J. *et al* Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255 (2009).
23. Hebart, M. N. *et al*. Things: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PloS one* **14**, e0223792 (2019).
24. Ye, Z., Yao, L., Zhang, Y. & Gustin, S. Self-supervised cross-modal visual retrieval from brain activities. *Pattern Recognition* **145**, 109915 (2024).
25. Du, C., Fu, K., Li, J. & He, H. Decoding visual neural representations by multimodal learning of brain-visual-linguistic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).
26. Song, Y. *et al* Decoding natural images from eeg for object recognition. *The Twelfth International Conference on Learning Representations* (2024).
27. Lu, Z. & Golomb, J. D. Generate your neural signals from mine: individual-to-individual eeg converters. *arXiv preprint arXiv:2304.10736* (2023).
28. Everingham, M., Van Gool, L., Williams, C. K., Winn, J. & Zisserman, A. The pascal visual object classes (voc) challenge. *International journal of computer vision* **88**, 303–338 (2010).
29. Grootswagers, T., Robinson, A. K. & Carlson, T. A. The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage* **188**, 668–679 (2019).
30. Robinson, A. K., Grootswagers, T. & Carlson, T. A. The influence of image masking on object representations during rapid serial visual presentation. *NeuroImage* **197**, 224–231 (2019).
31. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. Untangling featural and conceptual object representations. *NeuroImage* **202**, 116083 (2019).
32. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. The neural dynamics underlying prioritisation of task-relevant information. *Neurons, Behavior, Data analysis, and Theory* **5** (2021).
33. Gramfort, A. *et al*. Mne software for processing meg and eeg data. *neuroimage* **86**, 446–460 (2014).
34. Xue, S. *et al*. a multi-subject and multi-session eeg dataset for modelling human visual object recognition (2025).
35. Vidaurre, C. & Blankertz, B. Towards a cure for bci illiteracy. *Brain topography* **23**, 194–198 (2010).

Acknowledgements

We would like to thank Jingjing Dong, Handong Li, Zikang Liu, Jiahui Sun, Qianyi Liu, Erdong Hu, and Tongtian Yue for their help in data collection. This research is supported by Artificial Intelligence-National Science and Technology Major Project (2023ZD0121200) and the National Natural Science Foundation of China (U21B2043, 6243000159, 62102416), and the Key Research and Development Program of Jiangsu Province under Grant BE2023016-3. Data hosting is supported by the OpenNeuro project.

Author contributions

S.X. designed the study, collected EEG data and wrote the manuscript. B.J. helped in collecting EEG data. J.J. and L.G. revised the manuscript. J.L. supervised the study. J.Z. and C.W. provided resources and equipments. All the authors have reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-04843-x.

Correspondence and requests for materials should be addressed to J.L.