



OPEN

DATA DESCRIPTOR

EEG Dataset for the Recognition of Different Emotions Induced in Voice-User Interaction

Ga-Young Choi^{1,6}, Jong-Gyu Shin^{2,6}, Ji-Yoon Lee^{1,3}, Jun-Seok Lee^{1,3}, In-Seok Heo²,
Ha-Yeong Yoon⁴, Wansu Lim⁵, Jin-Woo Jeong⁴, Sang-Ho Kim²✉ & Han-Jeong Hwang^{1,3}✉

Electroencephalography (EEG)-based open-access datasets are available for emotion recognition studies, where external auditory/visual stimuli are used to artificially evoke pre-defined emotions. In this study, we provide a novel EEG dataset containing the emotional information induced during a realistic human-computer interaction (HCI) using a voice user interface system that mimics natural human-to-human communication. To validate our dataset via neurophysiological investigation and binary emotion classification, we applied a series of signal processing and machine learning methods to the EEG data. The maximum classification accuracy ranged from 43.3% to 90.8% over 38 subjects and classification features could be interpreted neurophysiologically. Our EEG data could be used to develop a reliable HCI system because they were acquired in a natural HCI environment. In addition, auxiliary physiological data measured simultaneously with the EEG data also showed plausible results, i.e., electrocardiogram, photoplethysmogram, galvanic skin response, and facial images, which could be utilized for automatic emotion discrimination independently from, as well as together with the EEG data via the fusion of multi-modal physiological datasets.

Background & Summary

Emotion is an umbrella term referring to various mental states that are brought on by changes to an individual's environment and that can result in behavioral and cognitive changes in that individual. Emotion research has recently attracted considerable attention in many research fields, such as neuroscience, affective computing, ergonomics, medicine, psychology, and so on. In particular, research on emotion recognition has significantly increased because the next-generation HCI (Human-Computer Interaction) applications will be developed as adaptive systems that recognize user emotions^{1,2}.

Many studies have attempted to automatically recognize human emotions, in particular by using various biosignals such as electroencephalography (EEG), electrocardiography (ECG), electromyography (EMG), photoplethysmogram (PPG), and galvanic skin response (GSR)^{3–5}. Some of the studies attempted to distinguish dichotomous emotions, such as happiness vs. sadness^{6–9}, whereas others attempted to quantify emotions, such as by measuring arousal level or valence^{10,11}. Among the various biosignals, EEG has been the most widely used for emotion recognition, and many EEG-based open-access datasets are currently available for emotion recognition studies, such as DEAP^{12,13}, MAHNOB-HCI^{11,14}, SEED^{15,16}, and so on. In previous EEG-based emotion studies^{1–16}, visual or auditory stimuli pre-defined to trigger certain emotions were used to artificially evoke corresponding brain activity; pictures with natural scenes were used to evoke positive emotions, such as happiness and relaxation, whereas pictures with abominable scenes were used to evoke negative emotions, such as sadness and disgust. Although previous EEG-based emotion studies have shown promise for the possibility of decoding human emotions, it is necessary to decode human emotions induced in natural environments,

¹Department of Electronics and Information Engineering, Korea University, Sejong, 30019, Republic of Korea.

²Department of Industrial Engineering, Kumoh National Institute of Technology, Gumi, 39177, Republic of Korea.

³Interdisciplinary Graduate Program for Artificial Intelligence Smart Convergence Technology, Korea University, Sejong, 30019, Republic of Korea. ⁴Department of Data Science, Seoul National University of Science and Technology, Seoul, 01811, Republic of Korea. ⁵School of Electronic and Electrical Engineering, Sungkyunkwan University, Suwon, 16419, Republic of Korea. ⁶These authors contributed equally: Ga-Young Choi, Jong-Gyu Shin.

✉e-mail: kimsh@kumoh.ac.kr; hwanghj@korea.ac.kr

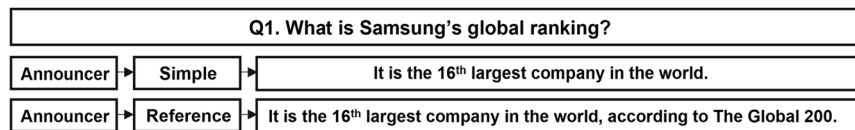


Fig. 1 Example of a simple and a detailed answer in terms of information quantity.

such as human-to-human communication situations, to develop more reliable applications based on emotion recognition.

In this study, we provide a novel EEG dataset containing the emotional information induced during a realistic HCI using a voice user interface (VUI) system that mimics natural human-to-human communication, thereby contributing to the advancement of emotion recognition research. The EEG data were acquired while fifty subjects were interacting with the VUI system. To induce emotional changes during the experiment, we controlled the two answer parameters of the VUI system, i.e., voice type (child/adult) and information quantity (simple/detailed), and a combination of the two parameters was randomly used to answer a user's question. During the experiment, we surveyed users concerning their satisfaction with the answers provided by the VUI system for each trial, and the satisfaction value was used as the ground truth for the induced emotions. We also simultaneously recorded other types of physiological data, such as ECG, PPG, GSR, and facial images, along with EEG data.

For data verification, we performed data analysis using a series of signal processing and machine learning methods that are widely used in EEG-based emotion recognition studies, and a simple detrending was applied only to the other physiological data that could be used for emotion recognition with the EEG data.

Methods

Subject. Fifty undergraduate students were recruited for this study, and they had no history of any psychiatric disease that could affect the results of the study. Prior to the experiment, the subjects were provided with the details of the experimental procedure, and they signed a form consenting to participate in the experiment and for videos of their likenesses to be shared in an open dataset. Among the subjects, 44 subjects (26 males and 18 females; 24.64 ± 2.13 years) agreed to data publication, whereas the others disagreed for privacy reasons. Adequate reimbursement was provided after the experiment. This study was approved by the Institutional Review Board (IRB) of Kumoh National Institute of Technology (202207-HR-003-01) and was conducted in accordance with the principles of the Declaration of Helsinki.

VUI system. We constructed a virtual VUI system with two different answer parameters, i.e., voice type (child/adult) and information quantity (simple/detailed), which were selected based on our previous studies that investigated various answer parameters in a VUI system in terms of emotion changes based on a subjective survey^{17,18}. Additionally, to prevent potential biases from subjects familiar with the experimental concept due to previous participation in stimulus rating, we recruited new subjects without any overlap with previous studies for the current experiment. Because commercial VUI products generally use female adult voice, such as, AWS, Bixby, Naver, and Google, we selected girl and female voices as the first VUI parameter. Moreover, we found in our previous studies^{17,18} that age difference was a significant factor influencing emotions, and thus age difference was selected as the second VUI parameter. In other words, we used a girl's voice and woman's voices for the child and adult voices, respectively. In terms of information quantity, a short, correct answer to a question was provided for the simple answer, whereas a correct answer along with the source of the answer was provided for the detailed answer (Fig. 1). Because it was impossible to prepare all possible answers to subjects' questions, which were unpredictable, we prepared 80 questions and their answers in such a way that all four answer types were identical in number (20 for each type; child/simple, child/detailed, adult/simple, and adult/detailed). These 80 questions are provided in a supplementary file (questions_answers.pdf). For this, we produced audio files for the answers to each of the 80 questions using a text-to-speech service (TYPECAST, <https://typecast.ai>), which is an audio creation tool that uses given voice types. All question-answers were based solely on one factual information without emotional content, ensuring that emotions would be induced solely by the two VUI parameters. Additionally, each question-answer pair was used only once, eliminating any potential learning effect.

Experimental paradigm. During the experiment, the subjects sat on a comfortable armchair, in front of a 21-inch monitor at a distance of 1 m, and they were instructed to remain relaxed and not to move. A single-blind experiment was performed to prevent the subjects' anticipation of answer types, as well as to mimic a realistic conversational setting. The questions were presented to the subject in the order written in the supplementary file (questions_answers.pdf). The subject was instructed to ask the VUI system a question by reading a sentence presented on the monitor, without prior knowledge of the sequence of questions. Then, a corresponding answer was provided to the subject from among the four types of answers (child/simple, child/detailed, adult/simple, and adult/detailed). Note that the duration of the answer period varied across trials depending on the question, with a mean duration of 3.01 ± 1.02 s. Additionally, event information was recorded at the time points indicated by green arrows in Fig. 2, and the corresponding temporal information can be found in the mrk file. After listening to the answers, the subjects were given 3 seconds to think about the answers so that they could recognize their emotions and participate in the survey^{19–21}. Then, the subject evaluated their feelings about the response made by the VUI system. In contrast to previous studies that have elicited specific emotions associated with valence-arousal through audio-visual stimuli, we focused on emotions induced in bidirectional interaction situations based on

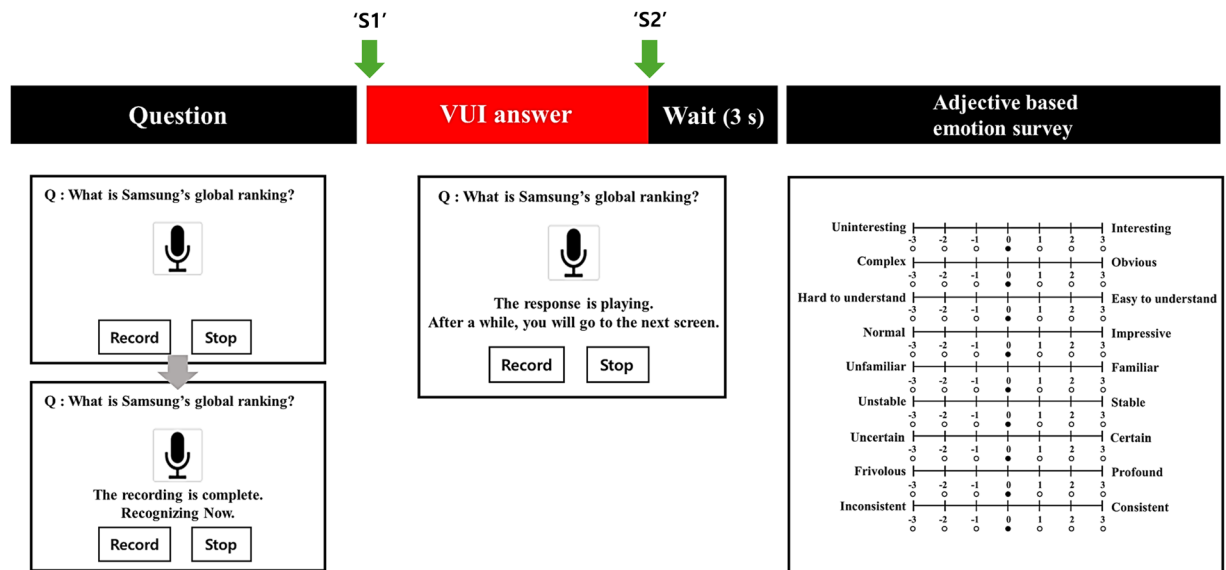


Fig. 2 Experimental paradigm for a single trial. A question is presented to the subject, and the subject asks the question by reading it out loud to the voice user interface (VUI) system. After the VUI system recognizes the question, an answer corresponding to one of the four types of answers (child/simple, child/detailed, adult/simple, and adult/detailed) is provided to the subject, and then a 3 s break is given to the subject. The duration of the answer period varied across trials depending on the question. A questionnaire about the response of the VUI system is presented to the subject, and the subject answers the questionnaire using a 7-point numeric rating scale for 9 contrasting adjective pairs. The single trial is repeated 80 times, with 20 trials for each of the four types of answers. The green arrows indicate the time points at which event information is recorded, and the start and end points of VUI answer are recorded as S1 and S2 in the mrk file, respectively.

individual characteristics. Therefore, we adapted factor analysis based on the Kansei engineering technique to decode individual emotions, which is a commonly used method in the emotional engineering fields to assess user evaluation²². We initially adopted Kansei words related to emotions that can affect individual emotions in the conversation between the VUI system and users²². More specifically, we first empirically extracted 30 pairs of emotion-related adjectives by expert judgment. Then, we finally determined 9 pairs of adjectives using a card sorting method based on the expert and users' opinions who experienced the VUI system²³. We created a questionnaire sheet to determine the user's emotion state in response to the answer of the VUI system using the extracted nine adjectives. After each VUI interaction, the subjects completed a questionnaire based on a 7-point numerical rating scale. Each subject performed 20 trials for each of the four answer types, which resulted in a total of 80 trials. The sequence of the 80 questions and their answers were the same for all subjects. The answers to the questionnaires are provided in a supplementary file (questions_answers.pdf).

Figure 2 shows the time sequence of the mentioned procedure for a single trial. To avoid excessive fatigue, a break was given to each subject whenever the subject wanted during the experiment.

Data recording. While the subjects were interacting with the VUI system, EEG data were measured using an ActiChamp EEG amplifier (Brain Products, GmbH Ltd., Germany) with a sampling rate of 1,000 Hz. Ground and reference electrodes were attached at Fpz and FCz, respectively. We used sixty-three active electrodes mounted on the scalp based on the international 10-10 system to measure the EEG data (FP1, Fz, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, CP1, Pz, P3, P7, O1, Oz, O2, P4, P8, TP10, CP6, CP2, Cz, C4, T8, FT10, FC6, FC2, F4, F8, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, C1, C5, TP7, CP3, P1, P5, PO7, PO3, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, AF8, AF4, and F2).

We simultaneously measured other biosignals along with the EEG data, using PPG, GSR, ECG, and facial expressions. We attached a PPG sensor on the index finger of the left hand, bipolar GSR sensors on the middle and ring finger of the left hand, and three ECG sensors at the lead-I position (Einthoven's triangle). Facial expression was measured between the starting point of an answer period and the waiting period of 3 s using a web camera (Logitech HD C920 Webcam, 1080p/30fps) based on FaceReader 8 (Noldus Information Technology, Wageningen, Netherlands). attached to the monitor for each trial. Various biosignals were also sampled at 1,000 Hz, the same as the EEG data, except facial expression (30 Hz), and all physiological signals were synchronously recorded^{19,23–28}. Figure 3 shows the positions of the physiological sensors (EEG, PPG, GSR, and ECG) and the web camera, respectively. The subjects were instructed to concentrate on the instructions presented on the screen and minimize unnecessary movements, such as eye and muscle movement. No datasets were excluded from the analysis specifically due to motion artifacts.

Factor analysis for questionnaire data. In this study, we used factor analysis to establish a criterion by identifying common factors underlying subjective individual emotions. Factor analysis is a multivariate statistical

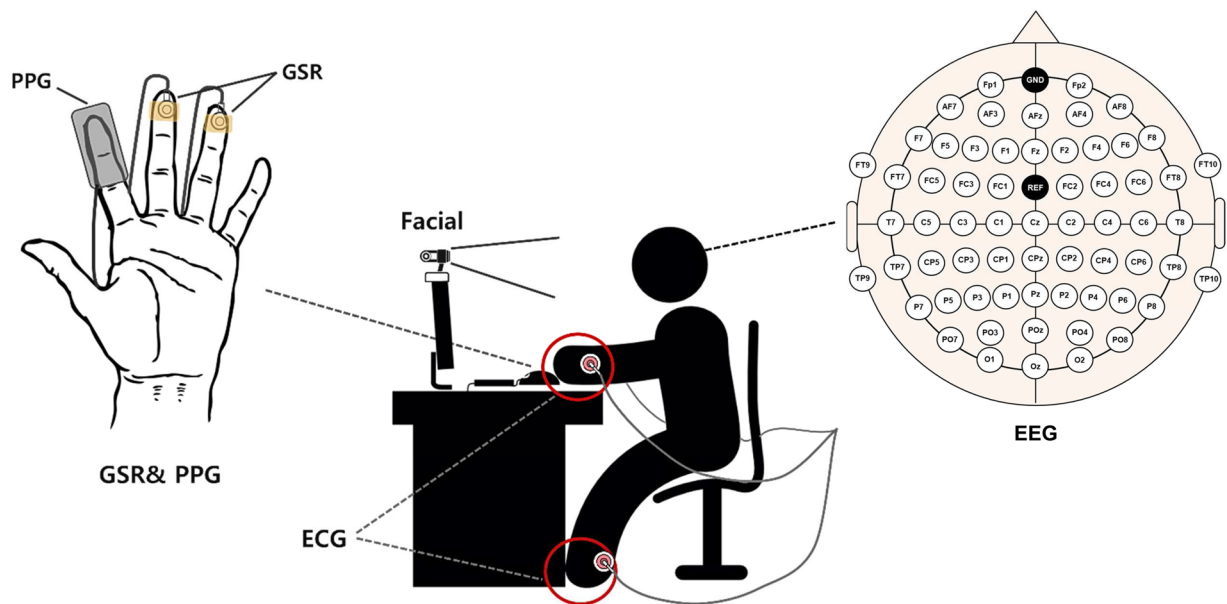


Fig. 3 Schematic diagram of sensor positions for recording EEG, ECG, PPG, GSR, and facial image data.

Adjective pairs	Factor 1	Factor 2
Unstable – Stable	0.820*	0.144
Hard to understand – Easy to understand	0.784*	0.069
Uncertain – Certain	0.742*	0.131
Inconsistent – Consistent	0.742*	0.061
Uninteresting – Interesting	0.067	0.873*
Normal – Impressive	0.032	0.868*
Complex – Obvious	0.707	0.068
Unfamiliar – Familiar	0.644	–0.015
Frivolous – Profound	0.505	–0.110
Variance (%var)	3.563 (0.396)	1.578 (0.175)

Table 1. Results of a factor analysis for 9 contrasting adjective pairs. *Adjective pairs selected by factor analysis. The variances of 3.563 and 1.578 for factors 1 and 2 were rounded to the nearest integer.

method used to examine the correlations between multiple variables to identify underlying common factors^{29,30}. The questionnaire results were statistically analyzed by factor analysis to extract main factors representing emotions using Minitab 20 (Minitab Inc., State College, PA, USA)¹⁸. The factors of adjective pairs were derived based on the correlation between the scores of the 9 adjective pairs. The number of meaningful factors was determined by the eigenvalues in a factor analysis where two factors with eigenvalues greater than 1 were extracted^{29,30}.

Table 1 shows the results of the factor analysis. The major adjective pairs were determined by the variance of each factor. For example, the variance of factor 1 is 3.563, which means that approximately four adjective pairs had a significant influence in constructing the corresponding factor. Therefore, we identified the most influential adjective pairs based on the absolute values of factor loadings, which indicate the correlation between each adjective pair and the respective factors. Consequently, we selected four adjective pairs for factor 1 and two adjective pairs for factor 2, as denoted by asterisks in Table 1. The commonalities of factors 1 and 2 explain approximately 39.6% and 17.5% of the total variance in factor loadings, respectively. We examined the overall meaning of the major adjective pairs (i.e., asterisks in Table 1) for two factors, and we defined factor 1 as “stability” and factor 2 as “favorability” based on the empirical judgment of experts¹⁸. Then, the factor scores for each trial were calculated for the defined two main factors (stability and favorability). The factor scores are defined by the formula:

$$F_i = b_{i1}X_1 + b_{i2}X_2 + \dots + b_{ip}X_p \quad (1)$$

where F_i is the score of the i -th factor, $X_1, X_2, \dots, X_p$ are the variables (pairs of questions and answers), and $b_{i1}, b_{i2}, \dots, b_{ip}$ are the factor loadings between the i -th factor and each variable. In the following section, we performed binary classification by labelling the data based on the factor scores, provided in a supplementary file (Individual_Classification_Accuracy.pdf).

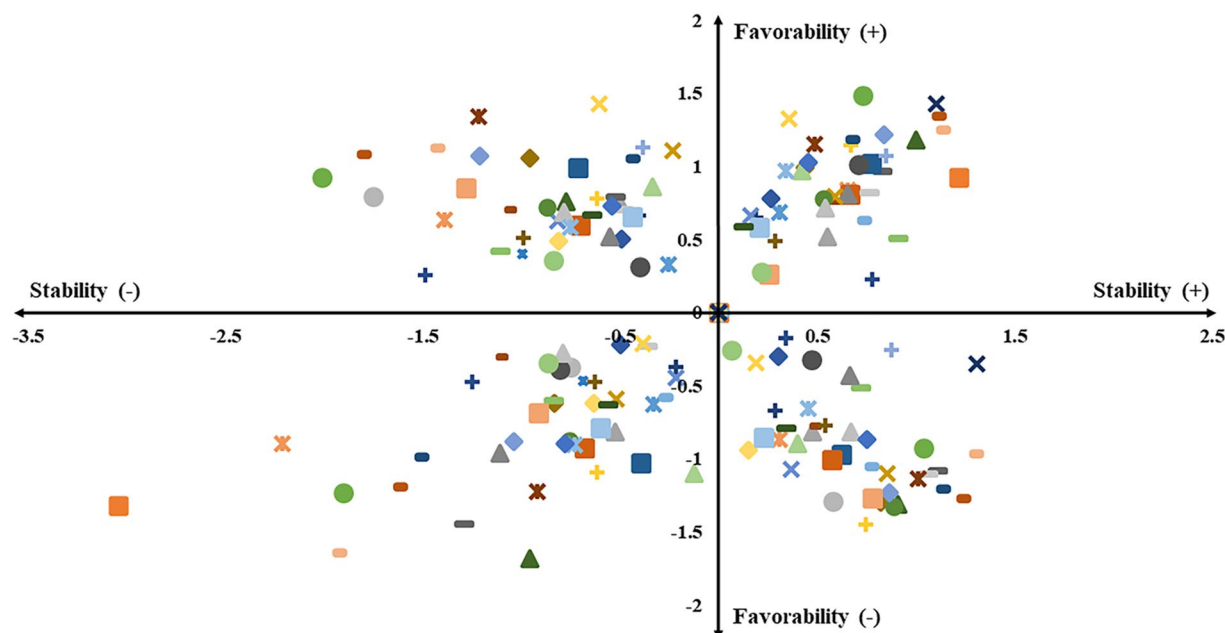


Fig. 4 Grand-averaged emotional distribution for each subject on the 2D factor coordinates (stability-favorability). Each point was estimated by averaging all factor analysis results of 80 trials for each subject in the 2D factor coordinates.

EEG data analysis. The EEG data were pre-processed using the EEGLAB toolbox based on MATLAB 2019b (MathWorks, Natick, MA, USA). The measured EEG data were notch-filtered at 59–61 Hz to remove the DC component. Subsequently, the data were bandpass-filtered at 1–80 Hz for baseline correction and extraction of frequency bands of interest because emotion-related EEG studies have frequently investigated the high-frequency band above 50 Hz at the same time as low-frequency bands^{31,32}. The filtered EEG data were downsampled from 1,000 Hz to 200 Hz to reduce computational load. To investigate emotional changes in response to the answer of VUI, the downsampled EEG data were segmented for the answer period without any overlap because it was expected that user emotions would be mostly induced in this answer period in our question-answer paradigm. The entire data without any segmentation can be found in the raw data folder, and can be flexibly used according to specific purposes. The entire data can be found in the raw data folder. Because the level of EEG amplitudes varied between subjects, we empirically identified bad channels showing exceptionally high activation by visually inspecting spectral topographic maps based on power spectral densities (PSDs), and excluded six bad channels (FT9, T7, TP9, FT10, T8, and TP10) for further analysis (see Supplementary Figure 2). Independent component analysis (ICA) was then applied to the EEG data concatenated by all segmented trials to remove physiological artifacts³³. In addition, we excluded significantly contaminated EEG trials where any channel showed an amplitude greater than $\pm 75 \mu\text{V}$ ^{34,35}, which resulted in an average of 7.39 ± 6.08 trials eliminated over 80 trials across the subjects.

Based on the two adjective factors, “stability” and “favorability”, which were derived by the factor analysis, the factor scores of all single EEG trials were used to categorized and labeled based on the 2D factor coordinates (stability-favorability) to confirm the difference between the emotions felt during the experiment in terms of EEG patterns and to automatically discriminate different emotions based on the factor axis using EEG features, respectively. Figure 4 shows the emotional distribution of all subjects for the factor coordinates. Mean factor scores for each subject within each quadrant were displayed on a two-dimensional factor coordinate system (stability-favorability), instead of displaying all 80 factor scores for each subject for better visibility. Individual detailed factor scores are provided in the open access repository of Figshare (Questionnaire.xlsx)³⁶.

We performed emotion classification for all possible binary combinations of the emotions represented by the factor coordinates; thus six classifications were independently performed for each subject, e.g., high stability and high favorability vs. high stability and low favorability. However, we excluded a binary combination of emotions for classification when the ratio of inter-class trials was higher than 7:3 due to the significantly unbalanced numbers of inter-class trials; six subjects did not have any available cases for binary emotion classification in terms of the ratio of inter-class trials, and thus they were excluded from the emotion classification. Phase-locking values (PLVs) widely used as classification features for emotion recognition were computed between all possible pairs of EEG electrodes for six frequency bands as emotion classification features³⁷: theta (4–8 Hz), alpha (8–13 Hz), low-beta (13–20 Hz), high-beta (20–30 Hz), low-gamma (30–50 Hz), and high-gamma (50–80 Hz)^{13,37–40}. Thus, a total of 9,576 PLV features ($_{57}C_2 \times 6$ frequency bands) were extracted for each subject. To determine optimal features for each subject, the inter-class significance of each feature was assessed using a t-test, aiming to extract the most discriminable feature⁴¹. The derived p-values were sorted in ascending order, and the classification was

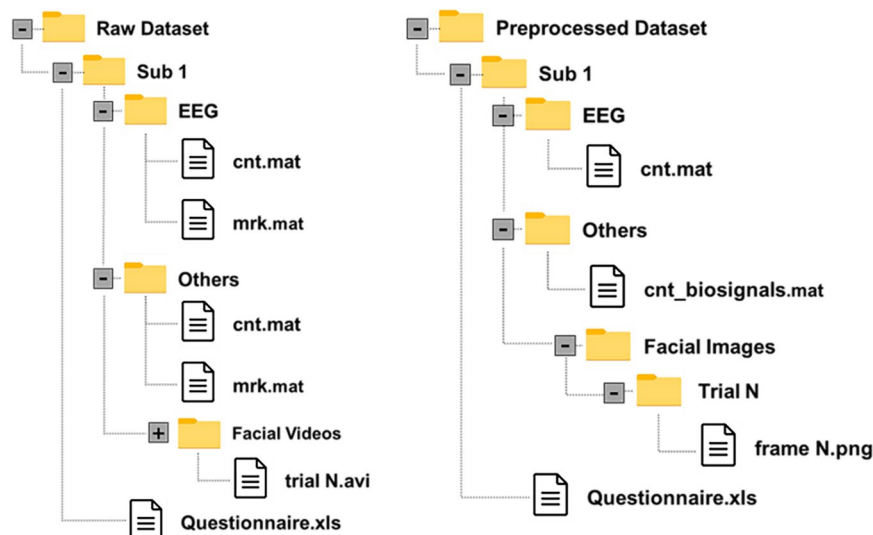


Fig. 5 Folder structures of raw and pre-processed datasets.

then performed independently, gradually increasing the number of features analyzed from rank 1 to 100^{42–44}. Classification accuracy was evaluated using 5-fold cross-validation based on a SVM with a linear kernel, standardized, and uniform prior setting to account for prior probabilities. We evaluated performance using balanced accuracy and F1 score to compensate for the data imbalance issue between classes. The optimal number of features differed for each individual, but maximum classification accuracies were achieved when using 3.95 ± 3.18 features, on average, across subjects. Additionally, we investigated the top four PLV features that were most frequently selected during classification over all subjects for the six combinations of binary emotions represented by the factor coordinates with respect to the frequency band.

$$\text{Balanced accuracy} = \frac{\text{Specificity} + \text{Sensitivity}}{2} \quad (2)$$

$$\text{F1 score} = \frac{2(\text{Sensitivity} \times \text{Precision})}{\text{Sensitivity} + \text{Precision}} \quad (3)$$

Because our main concern was to check whether emotion classification is possible using the EEG data measured in an interaction environment, the other physiological data, including the ECG, PPG, GSR, and facial image data, were simply analyzed to check their reliability. Thus, detrending was only applied to the other physiological data to remove the drift phenomenon, except a facial expression dataset that contained 30 image frames per seconds.

Data Records

The datasets are freely downloadable from the open access repository of Figshare³⁶.

We provide raw and pre-processed data for all biosignals. The raw data refers to the unprocessed initial state of the data, which was measured at a sampling rate of 1,000 Hz. The pre-processed data includes noise rejection and data segmentation for the answer period, based on the description provided in the ‘EEG data analysis’ section. Figure 5 shows the folder structures of the raw and pre-processed data, respectively. The datasets are in the form of MATLAB files (.mat) because we performed the data pre-processing using MATLAB R2013b (MathWorks, Natick, MA, USA).

Each subject has a folder (ex: Sub 1), and the raw data for each subject are stored in two sub-folders, one for the EEG and one for the other biosignal data. Each sub-folder has two MATLAB files (cnt and mrk) in common, which include continuous time-series data for all physiological data (cnt) and event information (mrk). Facial expression data are additionally included in the sub-folder of the other biosignal data as video files (.avi) for each trial. The cnt and mrk files of some subjects were split into several files because data recording was temporarily stopped and resumed whenever the subjects wanted to take a rest during the experiment. The questionnaire results for each trial are also included in the subject-specific folders (Questionnaire.xls). The duration for completing the survey following the VUI’s response varied for each participant, with breaks as needed, but the total recording time did not exceed 1.5 hours. The pre-processed data for each subject were stored in two different sub-folders, for EEG and other biosignal data, respectively, as raw data, but they have only cnt files without mrk files because the information in the mrk files was merged in the cnt files during pre-processing. Facial expression data were image files (.png) (30 frames per second). Figure 5 shows the detailed folder structures of the raw and pre-processed datasets.

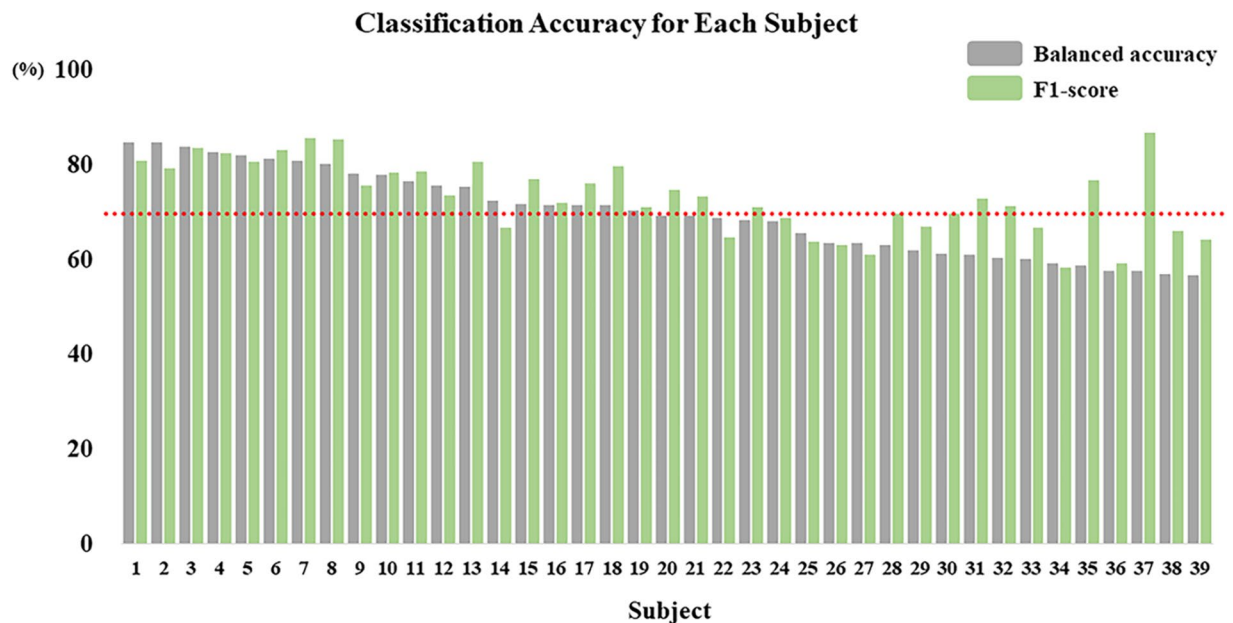


Fig. 6 Classification accuracies for each subject based on balanced accuracy (gray), and F1-score (green). The maximum accuracy of six classification accuracies was used for each subject. The mean balanced accuracy and F1 score were $69.8 \pm 8.8\%$ and $73.3 \pm 7.6\%$, respectively. 28 of 39 subjects showed a classification accuracy greater than 70% for at least one of the performance metrics. Individual classification accuracies for all six test cases are provided in a supplementary file (Individual_Classification_Accuracy.pdf). Note that we excluded certain cases for classification where the ratio between two classes was more extreme than 7:3 in terms of the number of trials due to unbalanced numbers of inter-class trials. Therefore, their classification accuracies were not estimated and are denoted by a dash (—) in the table.

Technical Validation

Figure 6 shows the classification accuracies of each subject in descending order based on balanced accuracy. Since each individual would respond differently to VUI answers, leading to different discriminable emotions, the pair of most discriminable emotions for each individual was identified and its classification accuracy was selected as the representative accuracy for each subject. The mean balanced accuracy and F1 score were $69.8 \pm 8.8\%$ and $73.3 \pm 7.6\%$, respectively. 28 of 39 subjects showed a classification accuracy greater than 70% for at least one of the performance metrics. Individual classification accuracies for all six test cases are provided in a supplementary file (Individual_Classification_Accuracy.pdf). Note that we excluded certain cases for classification where the ratio between two classes was more extreme than 7:3 in terms of the number of trials due to unbalanced numbers of inter-class trials. Therefore, their classification accuracies were not estimated and are denoted by a dash (—) in the table.

Figure 7 shows the PLV features most frequently selected during classification for the six binary combinations of emotions represented by the factor coordinates with respect to the frequency band. The PLV features selected for classification were generally found in the theta band, followed by the alpha and low-beta bands, and most of the EEG channels for the frontal areas were functionally connected with those in the temporal or occipital areas. These neurophysiological phenomena can be explained by the fact that frontal-occipital connections are closely related to emotion changes in the theta, alpha, and low-beta bands, respectively^{45–50}. For instance, in Fig. 7(b), depicting the condition of ‘high stability and high favorability vs. low stability and low favorability’ (i.e., positive vs. negative emotions), significant patterns are observed in the frontal, temporal, and occipital regions of the left hemisphere, consistent with prior investigations on word-based emotion recognition^{51–53}. In other classification scenarios, frontal-occipital connections were also prevalent, albeit with variations in the frequency bands exhibiting significant connections. Nonetheless, consistent with previous research^{45–50}, meaningful connections were consistently observed in the theta, alpha, and low-beta bands.

Figure 8 shows exemplary results of other physiological data that were measured simultaneously with the EEG data. The representative data were extracted from Subject 25 for the 18th trial while the subject received an answer from the VUI system; regular ECG and PPG patterns were observed well with a slow GSR fluctuation, and a sample facial image was also presented because no significant change was observed during the trial due to the conversation experimental environment.

Usage Notes

In this study, we provided a novel EEG dataset that contained emotion-related information acquired during interacting with VUI-based HCI. Because most existing EEG datasets for emotion recognition were acquired by presenting external stimuli that were pre-defined to induce certain emotions, the research results obtained using the existing EEG datasets may not be applicable to the development of real-world HCI applications based on emotion recognition. Because our EEG dataset has induced emotional information, it may be useful for the development of reliable HCI systems based on the classification of induced emotions. Note that few such EEG datasets have been published. Although we conducted a simple verification of the data availability

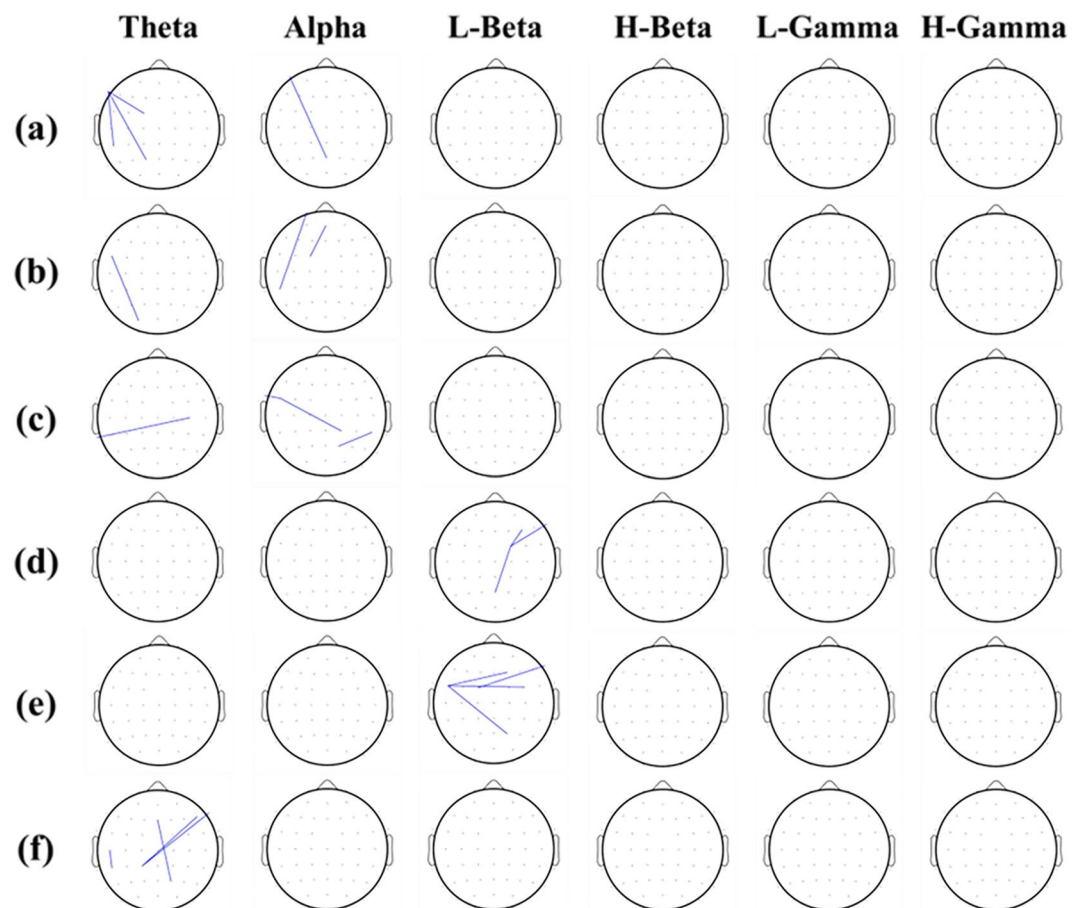


Fig. 7 PLVs most frequently selected for the six combinations of binary emotions for each frequency band: **(a)** high stability and high favorability vs. low stability and high favorability, **(b)** high stability and high favorability vs. low stability and low favorability, **(c)** high stability and high favorability vs. high stability and low favorability, **(d)** low stability and high favorability vs. low stability and low favorability, **(e)** low stability and high favorability vs. high stability and low favorability, **(f)** low stability and low favorability vs. high stability and low favorability.

in this study, our EEG dataset could be used in various ways for emotion recognition studies. For example, a multi-class emotion classification based on the 2D factor coordinates would be possible, e.g., high stability vs. low stability vs. high favorability vs. low favorability, and the estimation of discrete emotional levels is also possible for each of 9 contrasting adjective pairs based on the questionnaire results. Moreover, we tested various machine learning algorithms, including convolutional neural network-based deep learning algorithms, to determine the optimal classification algorithm for our dataset. As a result, the SVM model exhibited the highest mean classification accuracy. In general, a sufficient data is needed to effectively train deep learning models, which is typically larger than the amount required to train traditional machine learning models, such as SVM used in this study. Therefore, this multi-channel physiological dataset for emotion recognition can also be used to enhance the development of advanced emotion recognition classifiers. Furthermore, the effects of answer parameters and gender on induced emotions could be explored because our dataset was acquired using two different answer parameters, i.e., voice type and information quantity, from 26 males and 18 females.

Auxiliary physiological data were simultaneously measured with the EEG data, including ECG, PPG, GSR, and facial images. Even though we only provided an example result for the physiological data because our main concern was the EEG-based emotion recognition, they could be also utilized for emotion recognition studies. In fact, many previous studies have used ECG, PPG, GSR, and facial images for emotion recognition^{3-5,53-58}, and thus the auxiliary physiological data could be used for emotion recognition independently as well as together in the same way as suggested for the EEG dataset. In particular, when using more than two different physiological data, it is expected that the performance of emotion recognition could be improved by appropriately hybridizing them from a machine learning point of view, as shown in hybrid brain-computer interface studies based on the fusion of EEG and near-infrared spectroscopy^{19,25-28,41,48-50,53-79}.

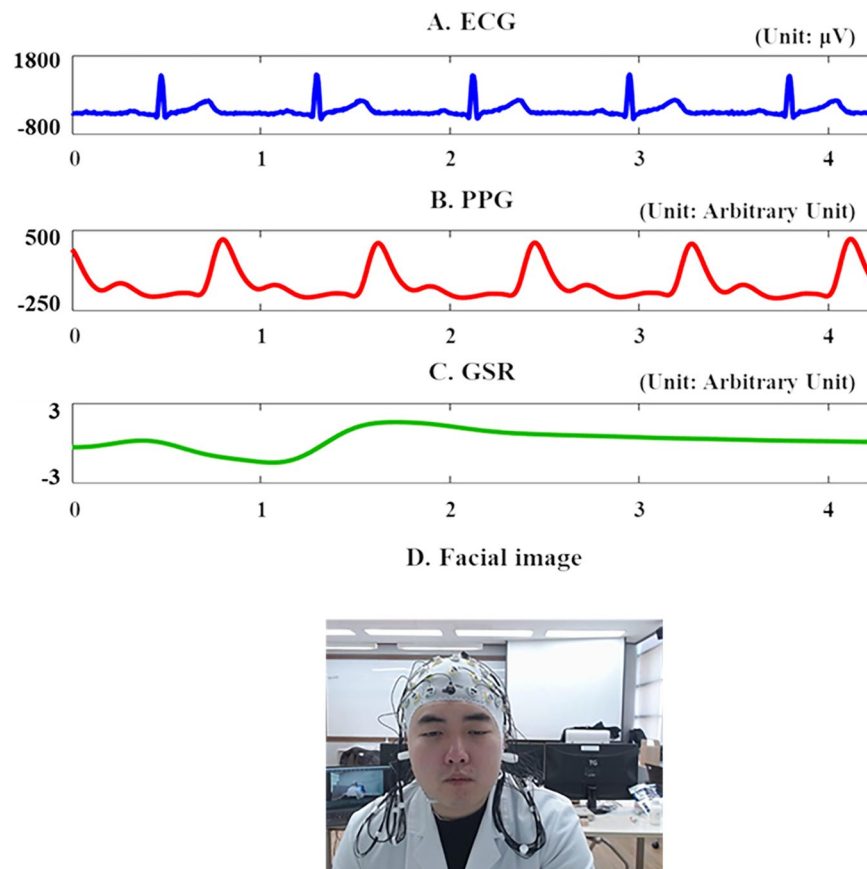


Fig. 8 Examples of four types of physiological data: (A) ECG, (B) PPG, (C) GSR, and (D) facial image. Note that the representative subject gave permission for sharing their facial image.

Code availability

We have provided relevant analysis MATLAB scripts used for data conversion and pre-processing at Figshare³⁶. Other researchers can conduct their own studies using the physiological data we provide, as we provide the necessary codes for data preprocessing. However, this study does not include a detailed analysis of the physiological data; it only provides scripts for basic filtering and data segmentation. Note that no custom code was used in this study.

Received: 18 July 2022; Accepted: 17 September 2024;

Published online: 03 October 2024

References

- Dadebayev, D., Goh, W. W. & Tan, E. X. EEG-based emotion recognition: Review of commercial EEG devices and machine learning techniques. *J. King. Saud. Univ. – Comput. Inf. Sci.* **1**, 29–40 (2021).
- Suhaimi, N. S., Mountstephens, J. & Teo, J. EEG-based emotion recognition: a state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* **8875426**, 1–19 (2020).
- Chunawale, A. & Bedekar, D. Human Emotion Recognition using Physiological Signals: A Survey. *Proceeding of 2nd International Conference on Communication & Information Processing (ICICIP)*. 3645402 (2020).
- Sarkar, P. & Etemad, A. Self-supervised learning for ECG-based emotion recognition. *Proceeding of ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. **1**, 3217–3221 (2020).
- Raheel, A., Majid, M., Alnowami, M. & Anwar, S. M. Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia. *Sensors*. **20**, 4037 (2020).
- Liu, Y. J. *et al.* Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* **9**, 550–562 (2017).
- Li, M. & Lu, B. L. Emotion classification based on gamma-band EEG. *Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. **1**, 1223–1226 (2009).
- Hassan, M. M. *et al.* Human emotion recognition using deep belief network architecture. *Inf. Fusion*. **51**, 10–18 (2019).
- Wiem, M. B. H. & Lachiri, Z. Emotion classification in arousal valence model using MAHNOB-HCI database. *Int. J. Adv. Comput.* **8**, 1–6 (2017).
- Reuderink, B., Mühl, C. & Poel, M. Valence arousal and dominance in the EEG during game play. *Int. J. Auton. Adapt.* **6**, 45–62 (2013).
- Chen, J. X. *et al.* Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks. *IEEE Access*. **7**, 44317–44328 (2019).
- Li, X. *et al.* Exploring EEG features in cross-subject emotion recognition. *Front. Neurosci.* **12**, 00162 (2018).
- Zheng, W. L., Zhu, J. Y. & Lu, B. L. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Trans. Affect. Comput.* **10**, 417–429 (2017).

14. Huang, X. Multi-modal emotion analysis from facial expressions and electroencephalogram. *Comput. Vis. Image. Underst.* **147**, 114–124 (2016).
15. Chai, X. Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition. *Comput. Biol. Med.* **79**, 205–214 (2016).
16. Zheng, W. L. & Lu, B. L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental. Develop.* **7**, 162–175 (2015).
17. Shin, J. G., Kim, J. B. & Kim, S. H. A framework to identify critical design parameters for enhancing user's satisfaction in human-AI interactions. *J. Phys. Conf. Ser.* **1284**, 237–243 (2019).
18. Shin, J. G., Choi, G. Y., Hwang, H. J. & Kim, S. H. Evaluation of emotional satisfaction using questionnaires in voice-based human-AI interaction. *Appl. Sci.* **11**, 1920 (2021).
19. Choi, S. I., Lee, J. Y., Lim, K. M. & Hwang, H. J. Evaluation of Real-Time Endogenous Brain-Computer Interface Developed Using Ear-Electroencephalography. *Front. Neurosci.* **16**, 842635 (2022).
20. Shim, M. *et al.* Altered functional networks of alpha and low-beta bands during upper limb movement and association with motor impairment in chronic stroke. *Brain Connect.* (2021).
21. Choi, S. I. & Hwang, H. J. Effects of different re-referencing methods on spontaneously generated ear-EEG. *Front. Neurosci.* **13**, 822 (2019).
22. Nagamachi, M. kansei engineering: a new ergonomic consumer-oriented technology for product development. *Int. J. Ind. Ergon.* **15**, 3–11 (1995).
23. Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. The measurement of meaning, 1st ed. *University of Illinois Press.* **1**, 18–30 (1957).
24. Choi, G. Y., Han, C. H., Jung, Y. J. & Hwang, H. J. A multi-day and multi-band dataset for a steady-state visual-evoked potential-based brain-computer interface. *Gigascience* **8**, giz133 (2019).
25. Han, C. H., Choi, G. Y. & Hwang, H. J. deep convolutional neural network based eye states classification using ear-EEG. *Expert. Syst. Appl.* **192**, 116443 (2022).
26. Han, C. H., Müller, K. R. & Hwang, H. J. Brain-switches for asynchronous brain-computer interfaces: A systematic review. *Electronics* **9**, 422 (2020).
27. Choi, S. I. & Hwang, H. J. Effects of different re-referencing methods on spontaneously generated Ear-EEG. *Front. Neurosci.* **1**, 822 (2016).
28. Choi, S. I. *et al.* On the feasibility of using an ear-EEG to develop an endogenous brain-computer interface. *Sensors* **18**, 2856 (2018).
29. Kaiser, H. F. An index of factorial simplicity. *Psychometrika* **39**, 31–36 (1974).
30. Watson, J. C. Establishing evidence for internal structure using exploratory factor analysis. *Meas. Eval. Couns. Dev.* **50**, 232–238 (2017).
31. Yang, K. *et al.* High gamma band EEG closely related to emotion: evidence from functional network. *Front. Hum. Neurosci.* **14**, 89 (2020).
32. Gao, Z., Cui, X., Wan, W. & Gu, Z. Recognition of emotional states using multiscale information analysis of high frequency EEG oscillations. *Entropy* **21**, 609 (2019).
33. Viola, F. C., Debener, S., Thorne, J. & Schneider, T. R. Using ICA for the analysis of multi-channel EEG data. Simultaneous EEG and fMRI: recording, analysis, and application. *Oxford University.* **1**, 121–133 (2010).
34. Brown, K. W., Goodman, R. J. & Inzlicht, M. Dispositional mindfulness and the attenuation of neural responses to emotional stimuli. *Soc. Cogn. Affect. Neurosci.* **8**, 93–99 (2013).
35. Arruda, J. E., Walker, K. A., Weiler, M. D. & Valentino, D. A. Validation of a right hemisphere vigilance system as measured by principal component and factor analyzed quantitative electroencephalogram. *Int. J. Psychophysiol.* **32**, 119–128 (1999).
36. Choi, G-Y. *et al.* EEG Dataset for the Recognition of Different Emotions Naturally Induced in Voice-User Interaction, *figshare*, <https://doi.org/10.6084/m9.figshare.c.7207839> (2024).
37. Wang, Z., Tong, Y. & Heng, X. Phase-locking value based graph convolutional neural networks for emotion recognition. *IEEE Access.* **7**, 93711–93722 (2019).
38. Gonuguntla, V., Shafiq, G., Wang, Y. & Veluvolu, K. C. EEG classification of emotions using emotion-specific brain functional network. *Proceeding of 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* **1**, 2896–2899 (2015).
39. Sammler, D., Grigutsch, M., Fritz, T. & Koelsch, S. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* **44**, 293–304 (2007).
40. Balasubramanian, G., Kanagasabai, A., Mohan, J. & Seshadri, N. G. Music induced emotion using wavelet packet decomposition—An EEG study. *Biomed. Signal. Process. Control.* **42**, 115–128 (2018).
41. Hwang, H. J., Kwon, K. & Im, C. H. Neurofeedback-based motor imagery training for brain-computer interface (BCI). *J. Neurosci. Methods.* **179**, 150–156 (2009).
42. Deligani, R. J., Borgheai, S. B., McLinden, J. & Shahriari, Y. Multimodal fusion of EEG-fNIRS: a mutual information-based hybrid classification framework. *Biomed. Opt. Express* **12**, 1635 (2021).
43. Aghajani, H. & Omurtag, A. Assessment of mental workload by EEG+ fNIRS. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* **1**, 3773–3776 (2016).
44. Shim, M., Lee, S. H. & Hwang, H. J. Inflated prediction accuracy of neuropsychiatric biomarkers caused by data leakage in feature selection. *Sci. Rep.* **11**, 7980 (2021).
45. Dasdemir, Y., Yildirim, E. & Yildirim, S. Analysis of functional brain connections for positive-negative emotions using phase locking value. *Cogn. Neurodyn.* **11**, 487–500 (2017).
46. Uusberg, A., Thiruchselvam, R. & Gross, J. J. Using distraction to regulate emotion: Insights from EEG theta dynamics. *J. Psychophysiol.* **91**, 254–260 (2014).
47. Tandle, A., Jog, N., Dharmadhikari, A. & Jaiswal, S. Estimation of valence of emotion from musically stimulated EEG using frontal theta asymmetry. *Proceeding of 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery.* **1**, 63–68 (2016).
48. Schubring, D. & Schupp, H. T. Emotion and brain oscillations: high arousal is associated with decreases in alpha-and lower beta-band power. *Cereb. Cortex.* **31**, 597–1608 (2021).
49. Schubring, D. & Schupp, H. T. Affective picture processing: Alpha-and lower beta-band desynchronization reflects emotional arousal. *Psychophysiology* **56**, e13386 (2019).
50. Ray, W. J. & Cole, H. W. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science.* **228**, 750–752 (1985).
51. Reisch, L. M., Węgrzyn, M., Woermann, F. G., Bien, C. G. & Kissler, J. Negative content enhances stimulus-specific cerebral activity during free viewing of pictures, faces, and words. *Hum. Brain Mapp.* **41**, 4332–4354 (2020).
52. Herbert, C. *et al.* Amygdala activation during reading of emotional adjectives—an advantage for pleasant content. *Soc. Cogn. Affect. Neurosci.* **4**, 35–49 (2009).
53. Holtgraves, T. & Felton, A. Hemispheric asymmetry in the processing of negative and positive words: A divided field study. *Cogn. Emot.* **25**, 691–699 (2011).
54. Goshvarpour, A., Abbasi, A. & Goshvarpour, A. An accurate emotion recognition system using ECG and GSR signals and matching pursuit method. *Biomed J.* **40**, 355–368 (2017).

55. Xu, Y. & Liu, G. Y. A method of emotion recognition based on ECG signal. *Proceeding in 2009 International Conference on Computational Intelligence and Natural Computing*. **1**, 202–205 (2009).
56. Udovičić, G., Đerek, J., Russo, M. & Sikora, M. Wearable emotion recognition system based on GSR and PPG signals. *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*. **1**, 53–59 (2017).
57. Sepúlveda, A., Castillo, F., Palma, C. & Rodríguez-Fernández, M. Emotion recognition from ECG signals using wavelet scattering and machine learning. *Appl Sci*. **11**, 4945 (2021).
58. Kartali, A., Roglić, M., Barjaktarović, M., Đurić-Jovičić, M. & Janković, M. M. Real-time algorithms for facial emotion recognition: A comparison of different approaches. *Proceeding in 2018 14th Symposium on Neural Networks and Applications (NEUREL)*. **1**, 1–4 (2018).
59. Han, C. H., Müller, K. R. & Hwang, H. J. Enhanced performance of a brain switch by simultaneous use of EEG and NIRS data for asynchronous brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 2102–2112 (2020).
60. Shin, J., Kim, D. W., Müller, K. R. & Hwang, H. J. Improvement of information transfer rates using a hybrid EEG-NIRS brain-computer interface with a short trial length: Offline and pseudo-online analyses. *Sensors*. **18**, 1827 (2018).
61. Shin, J., Müller, K. R. & Hwang, H. J. Eyes-closed hybrid brain-computer interface employing frontal brain activation. *PLoS. One*. **13**, e0196359 (2018).
62. Han, C. H., Hwang, H. J., Lim, J. H. & Im, C. H. Assessment of user voluntary engagement during neurorehabilitation using functional near-infrared spectroscopy: a preliminary study. *J. Neuroeng. Rehabil.* **15**, 1–10 (2018).
63. Won, D. O., Hwang, H. J., Kim, D. M., Müller, K. R. & Lee, S. W. Motion-based rapid serial visual presentation for gaze-independent brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**, 334–343 (2017).
64. Lim, J. H. *et al.* An emergency call system for patients in locked-in state using an SSVEP-based brain switch. *Psychophysiology*. **54**, 1632–1643 (2017).
65. Shin, J., Müller, K. R., Schmitz, C. H., Kim, D. W. & Hwang, H. J. Evaluation of a compact hybrid brain-computer interface system. *Biomed. Res. Int.* **1**, 6820482 (2017).
66. Hwang, H. J. *et al.* Clinical feasibility of brain-computer interface based on steady-state visual evoked potential in patients with locked-in syndrome: Case studies. *Psychophysiology*. **54**, 444–451 (2017).
67. Shin, J., Müller, K. R. & Hwang, H. J. Near-infrared spectroscopy (NIRS)-based eyes-closed brain-computer interface (BCI) using prefrontal cortex activation due to mental arithmetic. *Sci. Rep.* **6**, 1–11 (2016).
68. Hwang, H. J. *et al.* Toward more intuitive brain-computer interfacing: classification of binary covert intentions using functional near-infrared spectroscopy. *J. Biomed. Opt.* **21**, 091303 (2016).
69. 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).
70. Hwang, H. J. *et al.* A gaze independent brain-computer interface based on visual stimulation through closed eyelids. *Sci. Rep.* **5**, 1–11 (2015).
71. Lim, J. H., Lee, J. H., Hwang, H. J., Kim, D. H. & Im, C. H. Development of a hybrid mental spelling system combining SSVEP-based brain-computer interface and webcam-based eye tracking. *Biomed. Signal. Process. Control*. **21**, 99–104 (2015).
72. Hwang, H. J., Lim, J. H., Kim, D. W. & Im, C. H. Evaluation of various mental task combinations for near-infrared spectroscopy-based brain-computer interfaces. *J. Biomed. Opt.* **19**, 077005 (2014).
73. Han, C. H., Hwang, H. J. & Im, C. H. Classification of visual stimuli with different spatial patterns for single-frequency, multi-class SSVEP BCI. *Electron. Lett.* **49**, 1374–1376 (2013).
74. Hwang, H. J., Kim, S., Choi, S. & Im, C. H. EEG-based brain-computer interfaces: a thorough literature survey. *Int. J. Hum-Comput. Interact.* **29**, 814–826 (2013).
75. Hwang, H. J., Kim, D. H., Han, C. H. & Im, C. H. A new dual-frequency stimulation method to increase the number of visual stimuli for multi-class SSVEP-based brain-computer interface (BCI). *Brain. Res.* **1515**, 66–77 (2013).
76. Park, S. *et al.* Evaluation of feature extraction methods for EEG-based brain-computer interfaces in terms of robustness to slight changes in electrode locations. *Med. Biol. Eng. Comput.* **51**, 571–579 (2013).
77. Lim, J. H., Hwang, H. J., Han, C. H., Jung, K. Y. & Im, C. H. Classification of binary intentions for individuals with impaired oculomotor function: ‘eyes-closed’ SSVEP-based brain-computer interface (BCI). *J. Neural. Eng.* **10**, 026021 (2013).
78. Hwang, H. J. *et al.* Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. *J. Neurosci. Methods*. **208**, 59–65 (2012).
79. Im, C. H., Hwang, H. J., Che, H. & Lee, S. An EEG-based real-time cortical rhythmic activity monitoring system. *Physiol. Meas.* **28**, 1101 (2007).

Acknowledgements

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2023-00258971) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. RS-2024-00455484).

Author contributions

G.-Y. Choi, J.-G. Shin, W. Lim, J.-W. Jeong, S.-H. Kim, and H.-J. Hwang designed this study. G.-Y. Choi, J.-G. Shin, J.-Y. Lee, J.-S. Lee, I.-S. Heo, H.-Y. Yoon conducted the preliminary and main experiments. G.-Y. Choi, J.-Y. Lee, and J.-S. Lee conducted the data analysis. G.-Y. Choi, J.-G. Shin, and H.-J. Hwang wrote the manuscript. All authors reviewed the manuscript and approved the final 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/s41597-024-03887-9>.

Correspondence and requests for materials should be addressed to S.-H.K. or H.-J.H.

Reprints and permissions information is available at 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024