BCI Open Source & Dataset Project

Main Project II

2016-2018

Open Source & Dataset Project Summary

KOREA UNIVERSITY
Pattern Recognition & Machine Learning Lab

- In graduate school, I participated in an open source & dataset project as the main contributor.
- In the open-source project, our team (total 4, as part leader) developed an open software package dedicated to the development of Brain-Computer Interfaces with various advanced pattern recognition algorithms.
- In the case of an open dataset project, our team presents a BCI dataset that includes the three major BCI paradigms (Motor imagery, SSVEP, P300) with a 54 subjects (64-channels) over 2 sessions.
- I contributed to building an entirely open-source & dataset project. In particular, I developed the open-source project such as a variety of BCI paradigms and analysis functions. I also managed the experimental process during the open-dataset project.
- Through the open-source project, we produced easy and useful analysis tools for BCI users and various types of BCI paradigms. In addition, the open-data project is not only one of the largest datasets in BCI fields, but also one of the high-quality data.
- Now, this entire project was publicly open via GitHub and GigaScience paper. This project was mainly implemented in Matlab.
- Through this project, I learned the whole basic process of the BCI (how to design BCI environment, analyze, and manage datasets)

Open Source & Dataset Project Summary

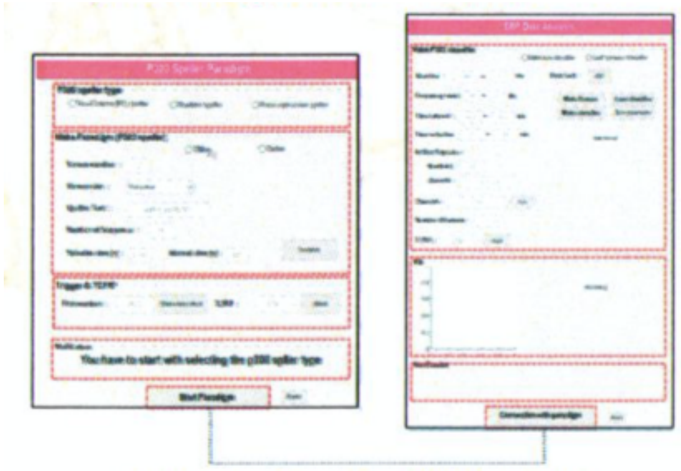
KOREA UNIVERSITY
Pattern Recognition & Machine Learning Lab

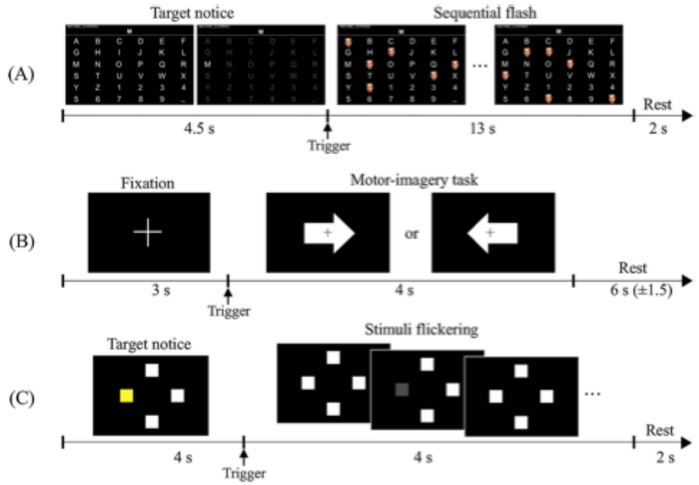
- **Main Contribution**
  - Develop a BCI open-source and dataset
- **Development tools**
  - Matlab
- **Achievement**
  - Paper publication (2nd author paper publication)
  - Release open source codes (https://github.com/PatternRecognition/OpenBMI)
  - Release open data (Link:*'gigadb.org/dataset/100542'*)

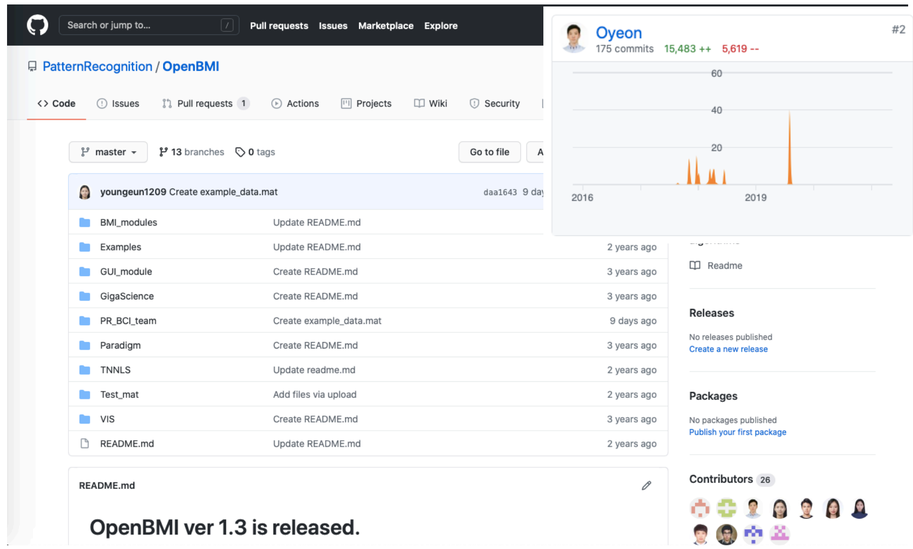
Open Source & Dataset Project Methods

KOREA UNIVERSITY
Pattern Recognition & Machine Learning Lab

- **About Open Source**
  - 1. Various BCI experiments
    - A binary-class MI system, a 36 symbol ERP speller, and a four target frequencies SSVEP system
  - 2. Full open-source package from study design to the outcome







http://openbmi.org/

https://github.com/PatternRecognition/OpenBMI

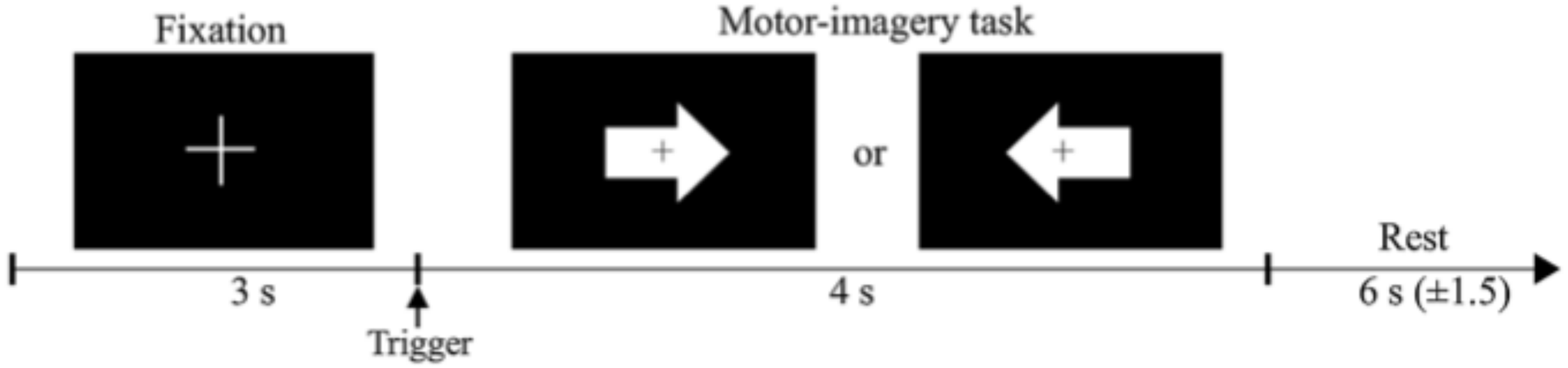
About Open Dataset

- 1. Subject number
  - Fifty-four healthy subjects (ages 24-35; 25 females) participated in the experiment
- 2. Data quantity & quality
  - Multiple ECG 64 channels & multiple sessions
  - 5~6 hours per subject for a single experiment
- 3. Data types
  - Including motor imagery, event-related potential, and steady-state visually evoked potential data

• Motor imagery setup

- After the first 3 s of each trial began, the subject performed the imagery task of grasping with the appropriate hand for 4 s when the right or left (100 trials) arrow appeared as a visual cue.

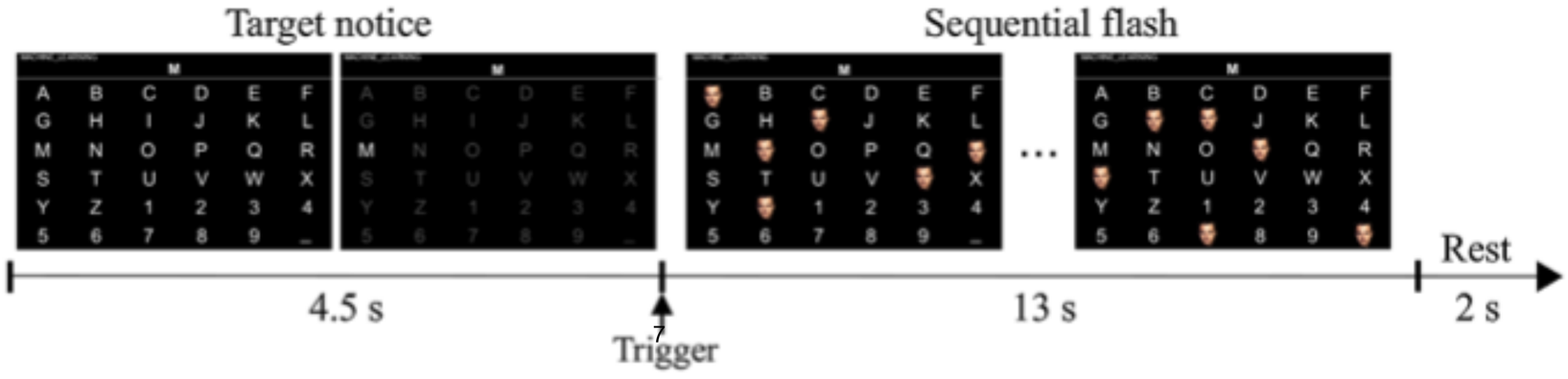
- The continuous EEG data were then segmented from 1,000 to 3,500 ms with respect to stimulus onset. EEG epochs were therefore constituted as 250 (data points) × 20 (electrodes) × 100 (trials).



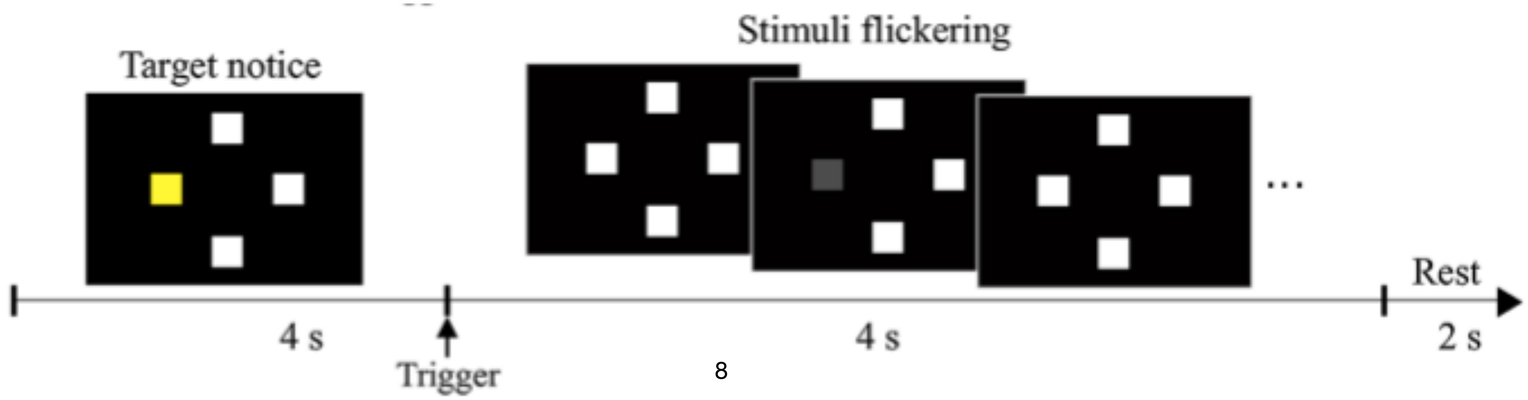
Open Source & Dataset Project Methods

KOREA UNIVERSITY
Pattern Recognition & Machine Learning Lab

- **Event-related potential setup**
  - The six rows and six columns were configured with 36 symbols. To enhance the signal quality, random-set and face stimuli were utilized.
  - The stimulus-time interval was set to 80 ms (e.g., face on), and the inter-stimulus interval (ISI) to 135 ms (btw face on-off).
    - A single iteration of stimulus presentation in all rows and columns was considered a sequence.
      - 1 sequence consisted of 12 stimulus flashes
      - A maximum of five sequences (i.e., 60 flashes) was allotted without prolonged inter-sequence intervals for each target character.
    - After the end of five sequences, 4.5 s were given to the user for identifying the next target character.
  - The continuous EEG data were segmented from –200 to 800 ms with respect to stimulus onset. EEG epochs in the offline phase therefore formed 100 (data points) × 32 (electrodes) × 1,980 (target and non-target trials).

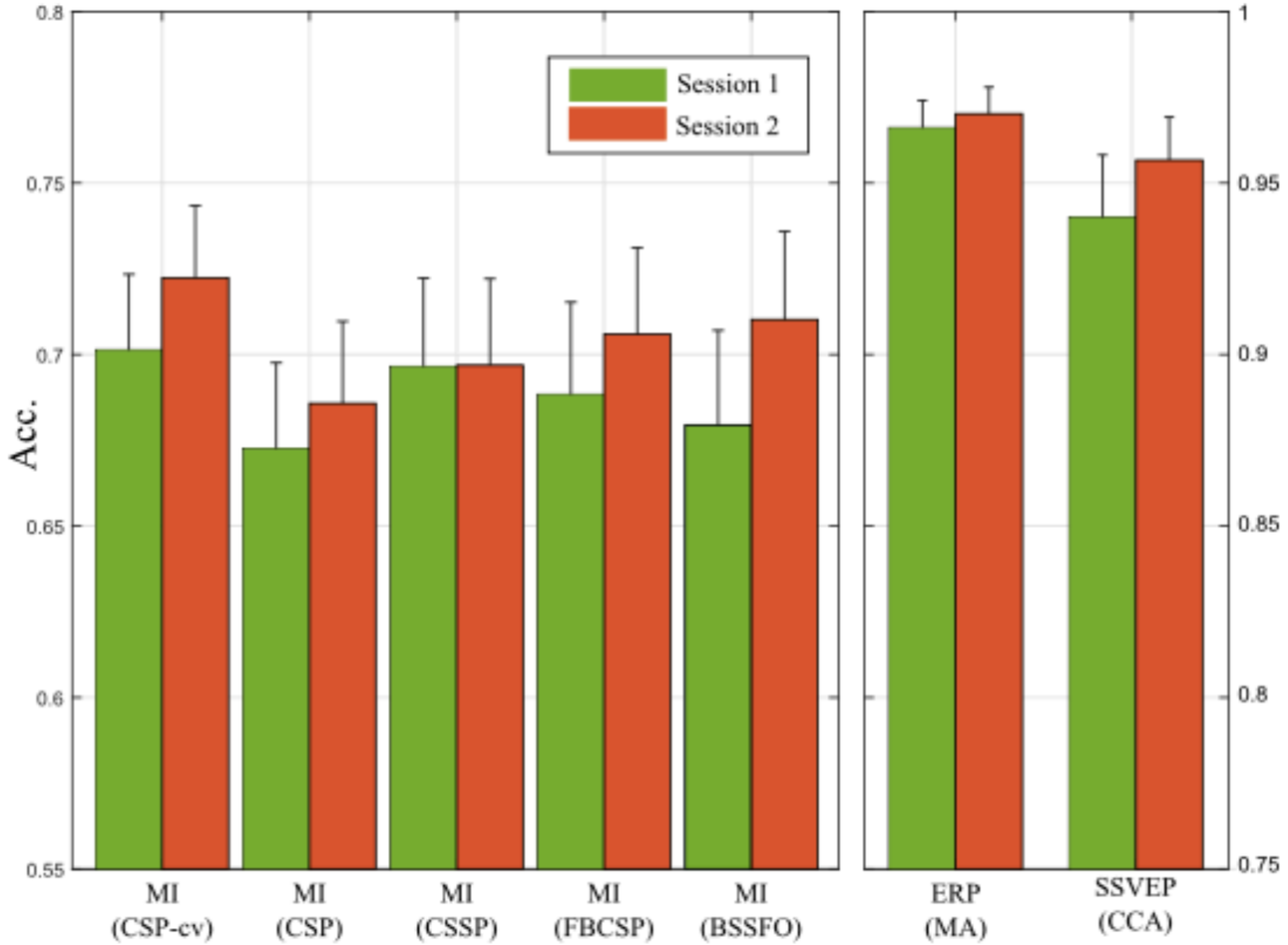


- **Steady-state visually evoked potential setup**
  - Four target SSVEP stimuli were designed to flicker at 5.45, 6.67, 8.57, and 12 Hz and were presented in four positions (down, right, left, and up, respectively) on a monitor.
  - Participants were asked to fixate the center of a black screen and then to gaze in the direction where the target stimulus was highlighted in a different color. Each SSVEP stimulus was presented for 4 s with an ISI of 6 s.
  - The continuous EEG data were segmented from 0 to 4,000 ms with respect to stimulus onset. Therefore, EEG epochs were 400 (data points) × 10 (electrodes) × 100 (trials).



Open Source & Dataset Project Results

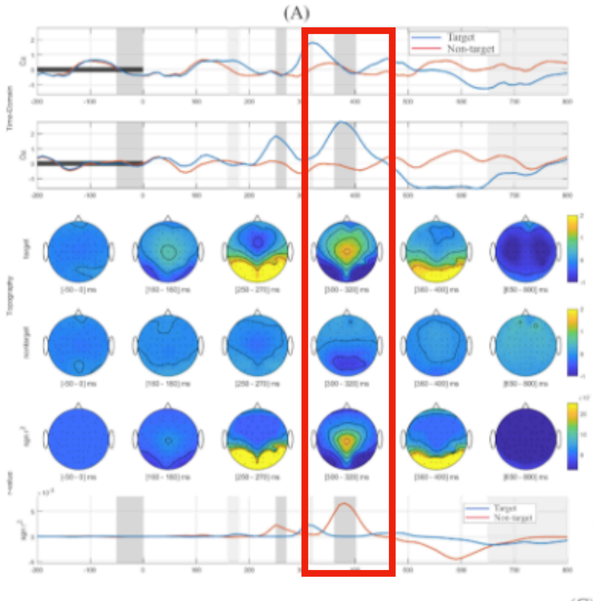
- **Evaluation**
  - We implemented various comparative algorithms to evaluate our dataset and showed that the dataset's performance was reasonable.
  - Moreover, we confirmed that our dataset showed a similar neuro-physiological result that lined with previous knowledge (next page).



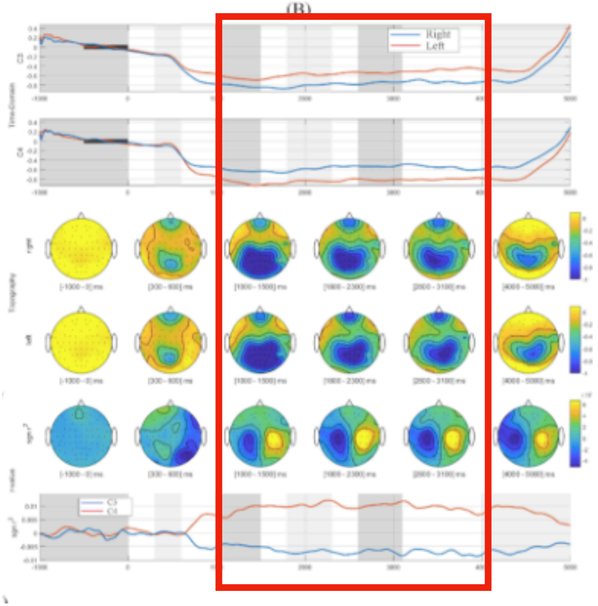
Open Source & Dataset Project Results

KOREA UNIVERSITY
Pattern Recognition & Machine Learning Lab

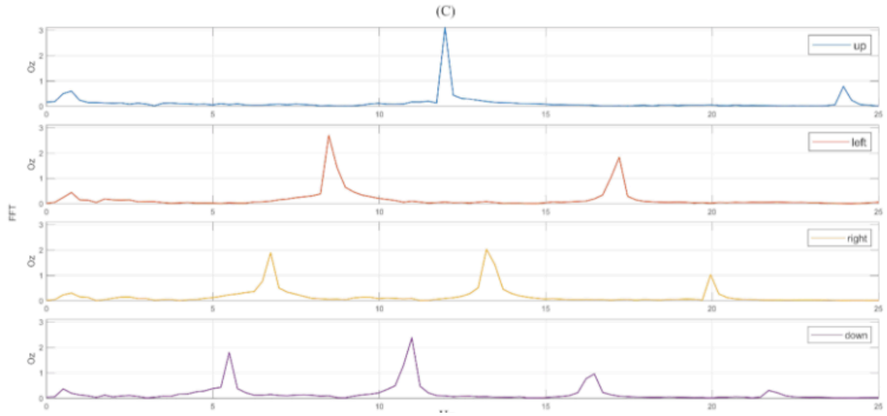
- **P300**
  - ERP response (i.e., P300) were observed at Cz and Oz electrodes



- **MI**
  - ERD/ERS pattern induced by left- or right-hand imagery tasks were observed at C3 and C4 electrodes



- **SSVEP**
  - Given frequency stimulus was observed at Oz electrode



Open Source & Dataset Project Conclusion

KOREA UNIVERSITY

Pattern Recognition & Machine Learning Lab

- This data is not only the largest datasets in BCI fields, but also one of the high-quality data.
- All methods for the data analysis in this study are supported with fully open-source scripts that can aid in every step of BCI technology.
- This experience made me a deep understanding of designing an entire BCI framework.

OXFORD (GIGA)n SCIENCE

GigaScience, 8, 2019, 1–16

doi: 10.1093/gigascience/giz002
Advance Access Publication Date: 30 January 2019
Research

RESEARCH

EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy

Min-Ho Lee [1], O-Yeon Kwon [1], Yong-Jeong Kim [1], Hong-Kyung Kim [1], Young-Eun Lee [1], John Williamson [1], Siamac Fazli [2] and Seong-Whan Lee [1,*]

[1]Department of Brain and Cognitive Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea and [2]Department of Computer Science, Nazarbayev University, Qabanbay Batyr Ave 53, Astana 010000, Kazakhstan

*Correspondence address. Seong-Whan Lee, Department of Brain and Cognitive Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea. Tel: +82-2-3290-3197; Fax: +82-2-3290-3583; E-mail: sw.lee@korea.ac.kr https://orcid.org/0000-0002-6249-4996

GigaScience, IF:7.267

164 citations (on 2022.11.14), 2nd author