Physiological and Eye-tracking Signals

Enobio 32 EEG system

Ergoneers Eye-tracking system

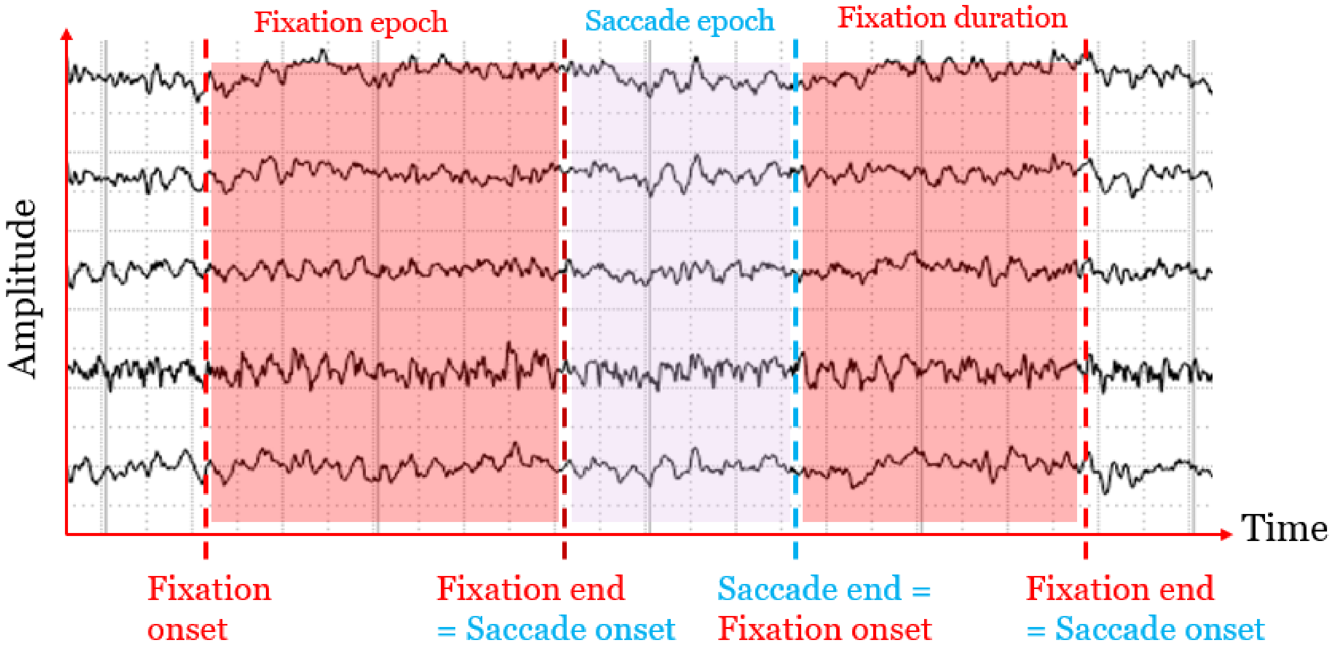
Empatica E4 system

Table: Sampling frequencies of signals used in this study

Signal type	EEG	Eye-tracking	PPG	EDA	Temp.
fs [Hz]	500	60	64	4	4

PPG: Photoplethysmogram / EDA: Electrical Dermal Activity / Temp.: Skin temperature

EEG preprocessing: Artifact removal



EEG signals in fixation and saccade epochs are compared/preprocessed, separately.

- ICA-based: Compare ***the variation ratio*** of EEG signals in ***fixation to*** EEG signals in ***saccade durations.***

- CAR-based: Re-reference EEG signals in ***fixation*** and EEG signals in ***saccade durations, independently.***

- This content is based on the manuscripts: Lee, H., Jiang, N., & Samuel, S. (2024). Detection of Error in Static and Dynamic Visual Stimulation via Electroencephalogram and Eye-tracking Systems. Engineering Applications of Artificial Intelligence. (Under review after the first revision)

EEG preprocessing - Automated Artifact Removal
: Eye-Tracking-based ICA (ET-ICA)

ICA decomposition: Each IC is fed into the selection algorithm

ET-based components selection
- Var. ratio of saccade to fixation
- Removed the ICs (> the ratio 1.1)

Reconstruction of selected components

$$Var_{epoch_j}\binom{fixation}{or\ saccade} = \frac{1}{N}\sum_i^N (y_i - EEG_{mean\_j})^2$$

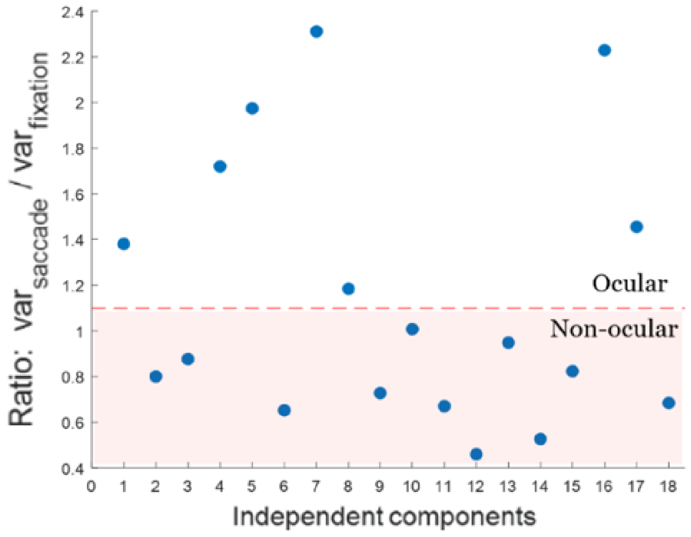
i = i-th signal point in either saccade or fixation epoch
N = the number of the signal points in the epoch
EEG_{mean_j} = mean EEG of j-th saccade or fixation epoch
Var_{epoch_j} = variance of EEG of j-th saccade or fixation epoch

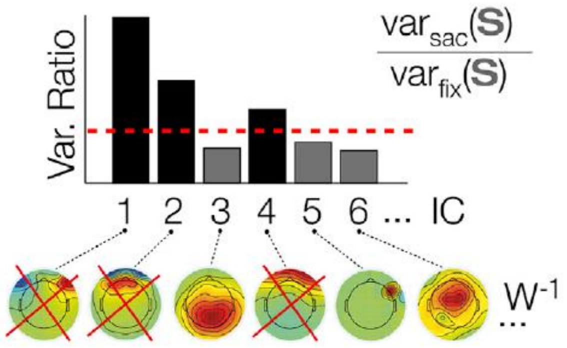
$$Var_{fixation\ or\ saccade} = \frac{1}{M}\sum_j^M Var_{epoch\_j}$$

j = j-th signal point in either saccade or fixation epoch
M = the number of either saccade or fixation epochs
$Var_{fixation\ or\ saccade}$ = variance of EEG of saccade or fixation in a trial

$$Variance\ ratio = \frac{Var_{saccade}}{Var_{fixation}} \quad \text{in a trial}$$

$Variance\ ratio$ = variance in saccades of fixation in fixations



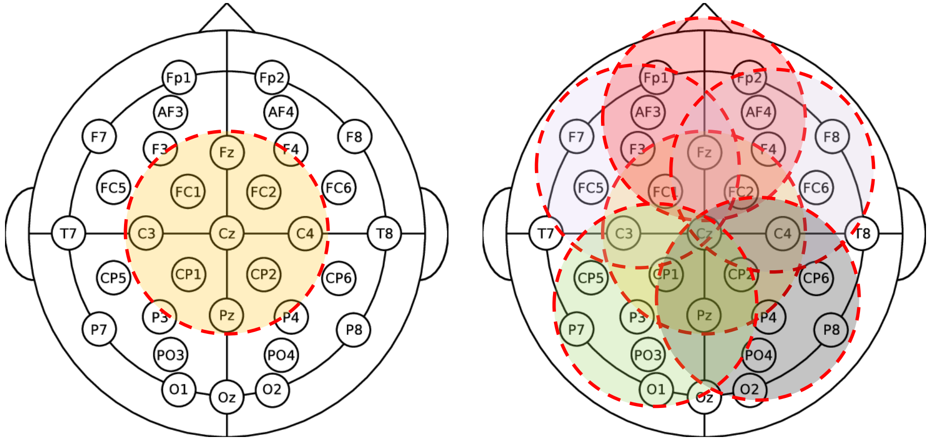


- *This content is based on the manuscripts: Lee, H., Jiang, N., & Samuel, S. (2024). Detection of Error in Static and Dynamic Visual Stimulation via Electroencephalogram and Eye-tracking Systems. Engineering Applications of Artificial Intelligence. (Under review after the first revision)*

EEG preprocessing - Automated Artifact Removal
: Revised Common Average Reference (R-CAR)

Localizing of CAR: selection of neighboring channels

Re-referencing each channel using different neighboring channels



$$Fixation: X_{non-saccade} = X_{raw\ in\ non-saccade} - \frac{\sum X_{i_{non-saccade}}}{N_{all}}$$

$$Saccade/blinks: X_{saccade} = X_{raw\ in\ saccade} - \frac{\sum_{i=1}^{N\ selected} X_{i_{saccade}}}{N_{selected}}$$

Localizing of CAR: Selection of neighboring channels

↓

Revised CAR filter
- CAR filter in non-saccade ranges
- CAR filter in saccade ranges

↓

Signal reconstruction using the two re-referenced signals (non-saccade and saccade ranges)

Preprocessed EEG

Figure 8. EEG amplitudes from raw signals, CAR, R-CAR, ICA and A-ICA.

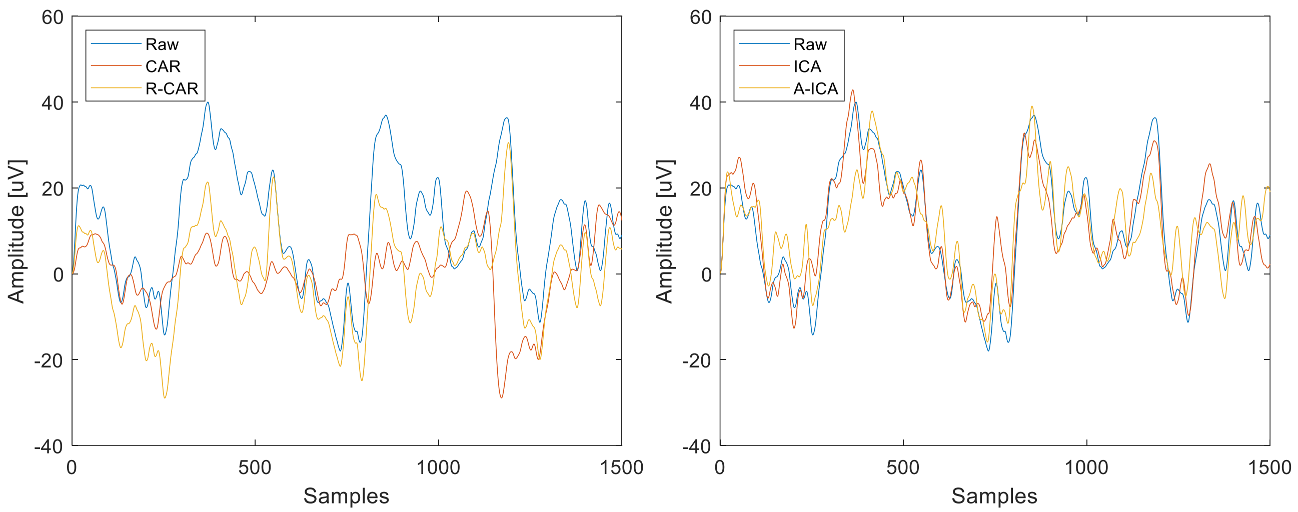
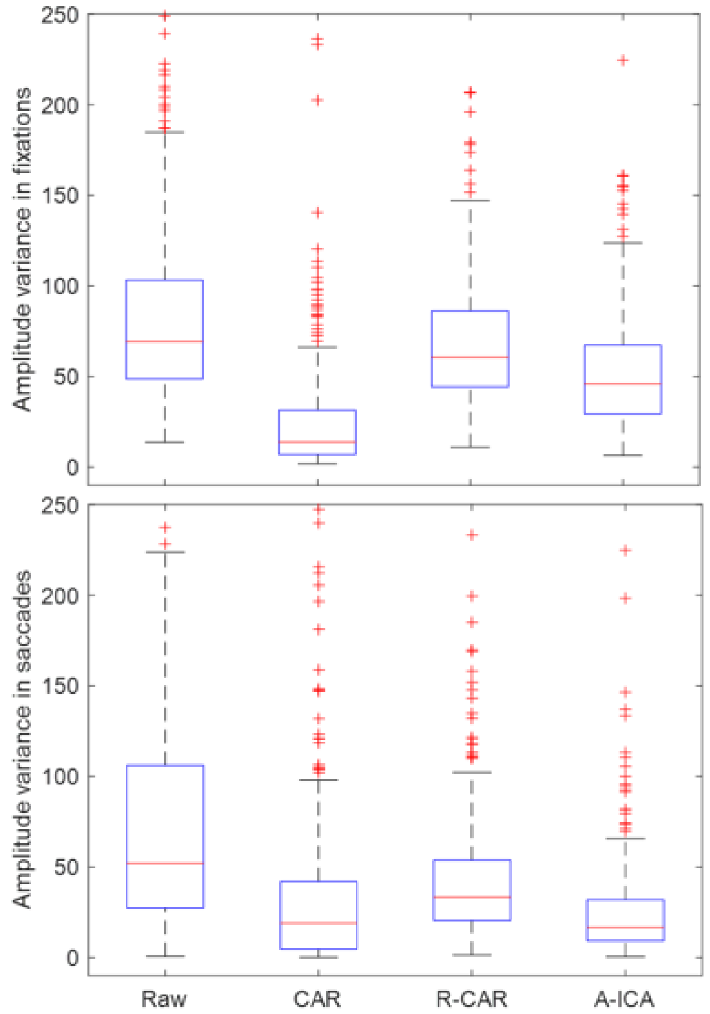


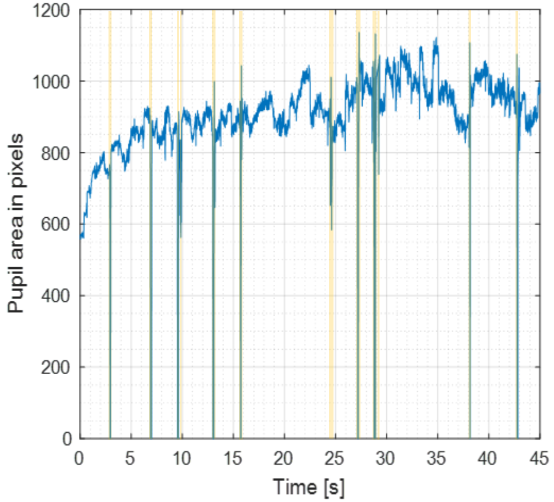
Table 9. Average number of selected ocular components by ICA and A-ICA in 20 trials.

Algorithm type	Single stimulus experiment		Multi stimuli experiment	
	avg. N	SD. N	avg. N	SD. N
ICA (manual)	3.5	1.1	4.4	1.9
A-ICA (automatic)	6.5	2.8	6.8	3.1

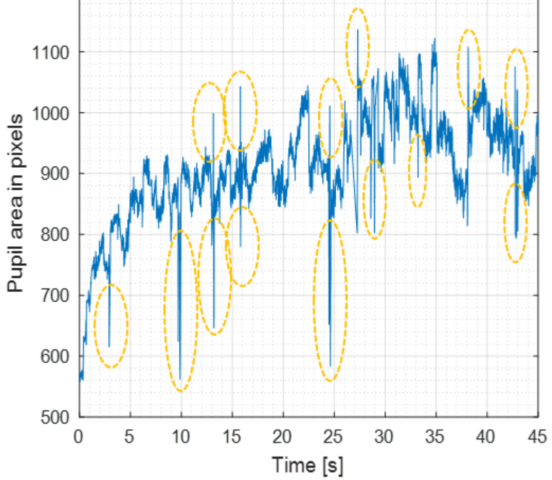
Figure 9. Filtered EEG amplitude's variances from raw, CAR, R-CAR and A-ICA algorithms with a bandpass filter over 300 trials: Left (variances in fixations) and Right boxplot (variances in saccades)



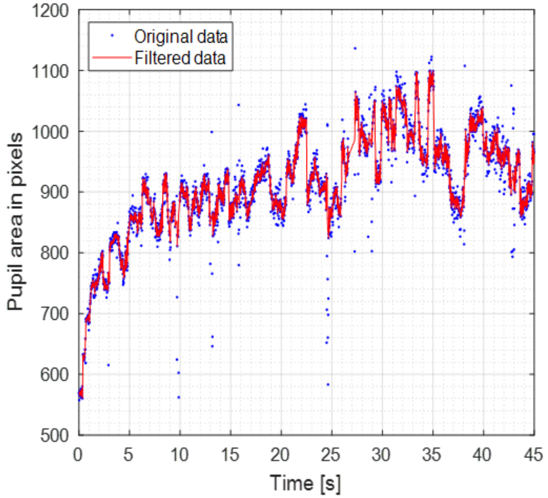
Eye-tracking signal preprocessing



Raw pupil area



Interpolated of pupil area (Outliers detection)



Processed pupil area (Outliers removed)

Interpolation between the blinks/saccades (detected as saccade)

↓

Hampel identifier (outlier detector)
-Median-based outlier detector
- Removal of detected outliers

↓

Processed pupil signals

- This content is based on the manuscripts: Lee, H., Jiang, N., & Samuel, S. (2024). Detection of Error in Static and Dynamic Visual Stimulation via Electroencephalogram and Eye-tracking Systems. Engineering Applications of Artificial Intelligence. (Under review after the first revision)

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Preprocessing for other signals

PPG signal
(Segmentation within 1 min)

↓

Filtering using Chebyshev II and passband of freq. ranges

↓

Peak detection / abnormality check
(e.g., peak-to-peak<0.5 or >1.2 s)

↓

Accepted segments

EDA (GSR) signal

↓

Filtering using Chebyshev II and passband of freq. ranges

↓

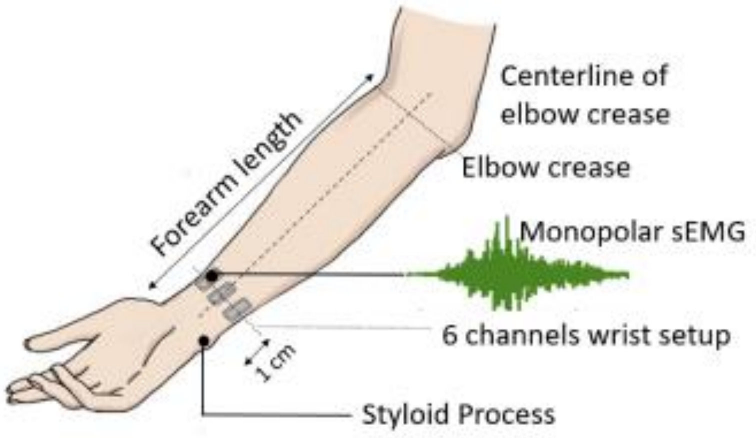
Peak detection / abnormality check
(peak-to-peak< 500 or > 1200 ms)

↓

Accepted segments

- This content is based on the manuscripts: Campanella, S., Altaleb, A., Belli, A., Pierleoni, P., & Palma, L. (2024). PPG and EDA dataset collected with Empatica E4 for stress assessment. Data in Brief, 53, 110102.

EMG signal preprocessing



- 10-500 Hz Butterworth bandpass
- 60 Hz notch filter to remove power-line noise

↓

- Monopolar sEMG: CAR
- Bipolar channel: Differential

↓

Segmentation/windowing

↓

(Feature extraction) Frequency Division Technique (FDT)
- FFT magnitude summation across sub-bands (log-transformed)

Discrete Fourier Transform

$$X[k] = \sum_{n=0}^{N-1} x_w[n]\, e^{-j2\pi \frac{kn}{N}}, \quad k = 0, 1, \ldots, N-1$$

Power sum

$$P_B = \sum_{k \in B} |X[k]|^2$$

Magnitude sum

$$M_B = \sum_{k \in B} |X[k]|$$

Feature Extraction from EEG, PPG, EDA and eye-tracking signals

Features from EEG

1) Energy & Singal Magnitude
2) Energy & Signal Distribution
3) Frequency Band Relationships

Table 4. Extracted Features of EEG, PPG, EDA, Skin temperature, eye-tracking

EEG features	Pupil and other physiological features (PPG, EDA and Skin temperature)
Mean energy	Power
Log energy entropy	Relative power
Mean	Mean energy
Mean (Relative signal)	Log energy entropy
Ratios of two bands	Shannon entropy
Ratio 1: A to B	Mean
Ratio 2: A to C	Variance
Ratio 3: A to D	Skewness
	Kurtosis
	Hjorth activity
The features above are extracted from each EEG channel (Cz, Fz and Pz)	Hjorth mobility
	Hjorth complexity
	Normalized 1st difference
	Normalized 2nd difference

- A is one of the four frequency bands, and others (B, C, D) are the remaining.

Features from others

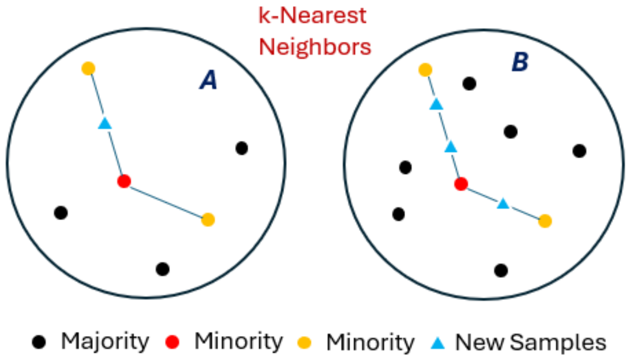
1) Energy & Power Magnitude
2) Complexity/Variability and Entropy

Class imbalance mitigation: Adaptive synthetic sampling (ADASYN)

Oversampling technique: ADASYN

Find all classes having a smaller # of instances than the majority class → Classify the classes as the minority classes → Generate synthetic data for minority classes

e.g., There are four classes.	Classification	Instance addition (Fully correct imbalance case)
1) Class 1: 80 instances	1) Class 1: Majority class	1) Class 1: 0 instance
2) Class 2: 40 instances	2) Class 2: Minority class	2) Class 2: 40 instances
3) Class 3: 50 instances	3) Class 3: Minority class	3) Class 3: 30 instances
4) Class 4: 60 instances	4) Class 4: Minority class	4) Class 4: 50 instances



B cluster (based on k-Nearest Neighbours) has a larger density of the majority class than A.
This corresponds to addition of more instances in B cluster.

1. **Identify Imbalanced Classes:** Detect classes with fewer instances than the majority class.

2. **Classify as Minority Classes:** Label classes with smaller sample sizes as minority classes.

3. **Generate Synthetic Data:** Use k-NN to generate synthetic instances focused on hard-to-classify areas near decision boundaries.
   - Example: More synthetic data is added to regions with higher density of the majority class (e.g., Cluster B in the figure).

4. **Balance Instance Distribution:** Synthetic samples ensure balance by correcting instance numbers across all classes while prioritizing challenging regions.

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Publication list related to the topics above

[1] **H. Lee.**, O. Lim., A. Singh., S. Samuel., "Collision Risk Perception Models Using Physiological and Eye-Tracking Signals", IEEE Access (2025)

[2] **H. Lee.**, N. Jiang., S. Samuel., "Detection of Error in Static and Dynamic Visual Stimulation via EEG and Eye-tracking Systems", Engineering Applications of Artificial Intelligence (2025)

[3] (Under Review) **H. Lee**, O. Lim, and S. Samuel, " Modeling Approach for Adaptive Workload Level Estimation using Physiological Features", AI for Transportation (2025)

[4] (Under Review) **H. Lee** and S. Samuel, " What the Body Tells the Car: Physiological and Eye-Tracking Cues for AV Preference Modeling", IEEE Transactions on Intelligent Vehicles (2025)

[5] A. Pradhan, J. He, **H. Lee**, and N. Jiang, "Multi-day Analysis of Wrist Electromyogram based Biometrics for Authentication and Personal Identification", *IEEE Transactions on Biometrics, Behavior, and Identity Science* (2023)