EEG Emotion Recognition Via Ensemble Learning Representation

Paper Review

Signal Processing (3231), Spring 2024, Konkuk University, Seoul, Prof. Changhoon Yim

Wonjun Park, Computer Science and Engineering, Konkuk University, Seoul, South Korea | May. 31st, 2024

Table of Contents

Acronym and Abbreviation

- Electroencephalography (EEG)
- Convolutional Neural Network (CNN)
- Long-Short Term Memory (LSTM)
- Recurrent Neural Network (RNN)
- Scaled Exponential Linear Unit (SELU)
- Database for Emotion Analysis using Physiological Signals (DEAP)
- Fast Fourier Transform (FFT)
- Adaptive moment estimation (Adam)

Paper

Paper
Authors

Bilal Taha[1]

Dae Yon Hwang[2]

Dimitrios Hatzinakos[3]

1,2,3 Electrical and Computer Engineering Department, University of Toronto, Ontario, Canada

Paper
Information

- **Conference**

2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

- **IEEE Citation**

B. Taha, D. Y. Hwang and D. Hatzinakos, "EEG Emotion Recognition Via Ensemble Learning Representations," *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10094939.

Paper

Agenda

As a fusion of spatial and temporal information,

the paper targeted to extract robust EEG features for emotion recognition.

Literature Reviews

Literature Reviews

EEG Applications

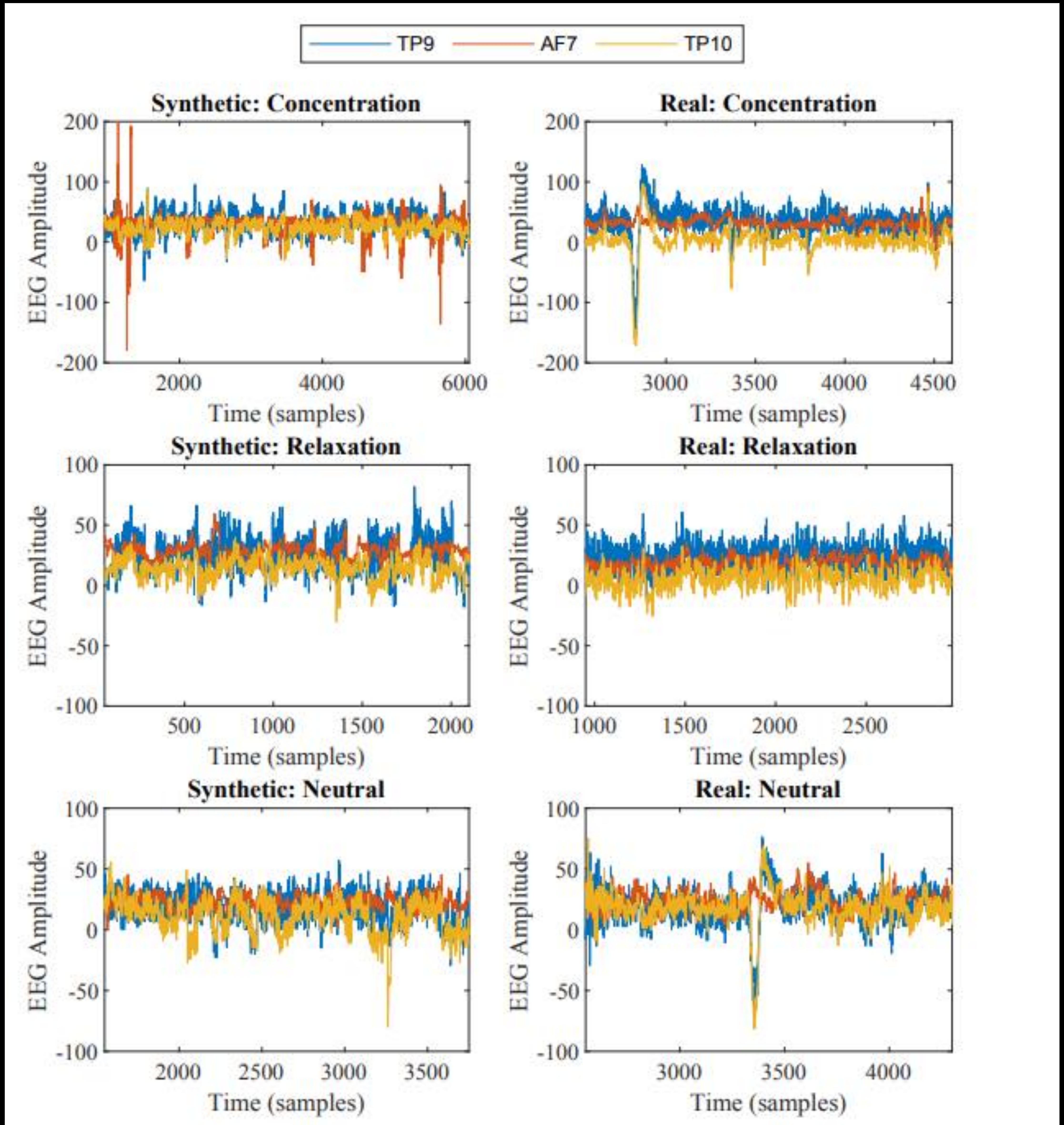
- Brain Disease Diagnoses
- Sleep Disorder and Physiology
- Motor Imagery Classification
- Emotion Analysis
- Vision and Auditory Reconstruction
- etc.

Which ranges of frequency bands are suitable for each task? [5, 6]

[5] Delorme, A. EEG is better left alone. Sci Rep 13, 2372 (2023). https://doi.org/10.1038/s41598-023-27528-0
[6] de Cheveigne, Alain. "Is EEG best left alone?." bioRxiv (2023): 2023-06.

Literature Reviews

EEG



Literature Reviews

EEG Sensors

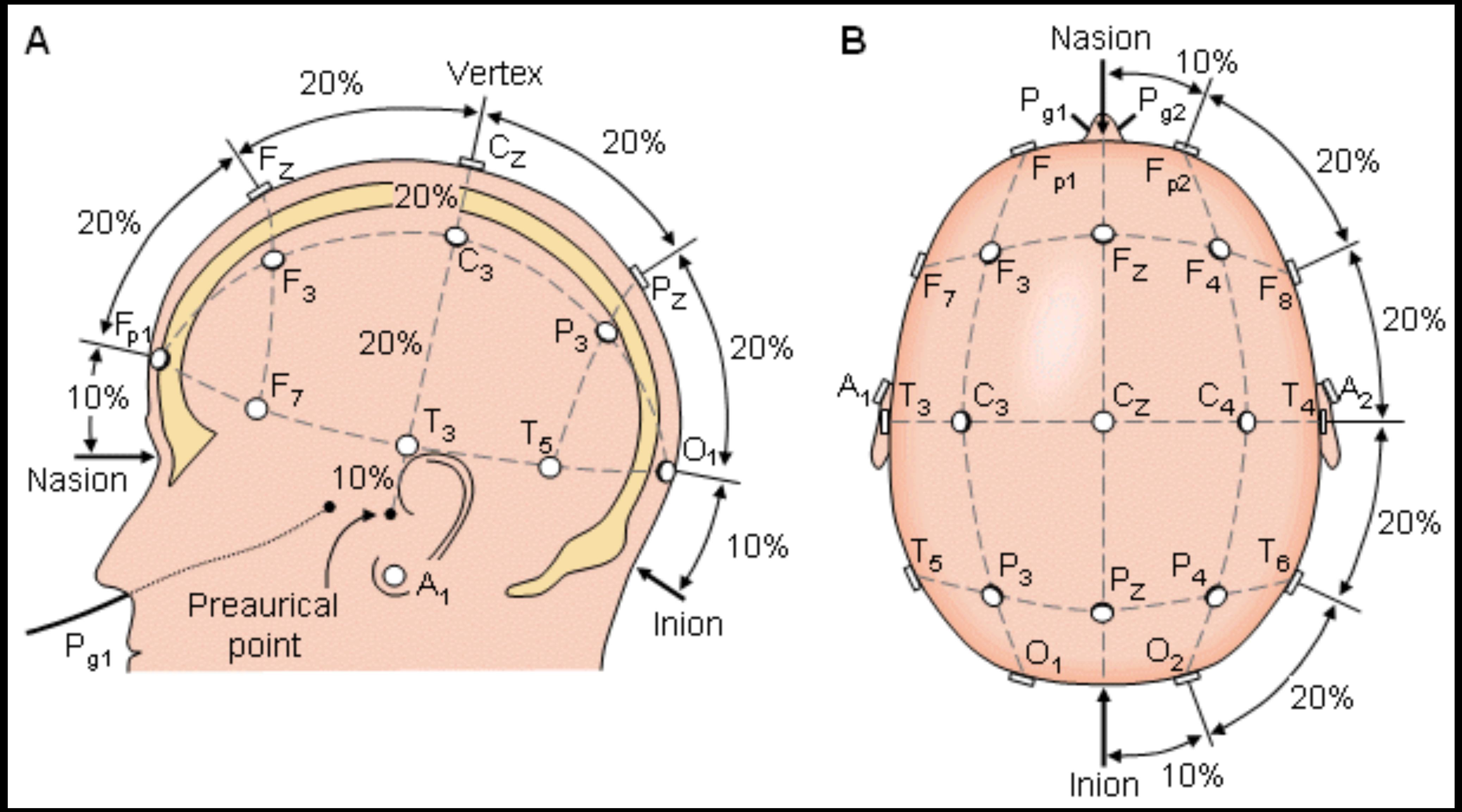


Fig 2. EEG 10-20 Electrode Placement [4]

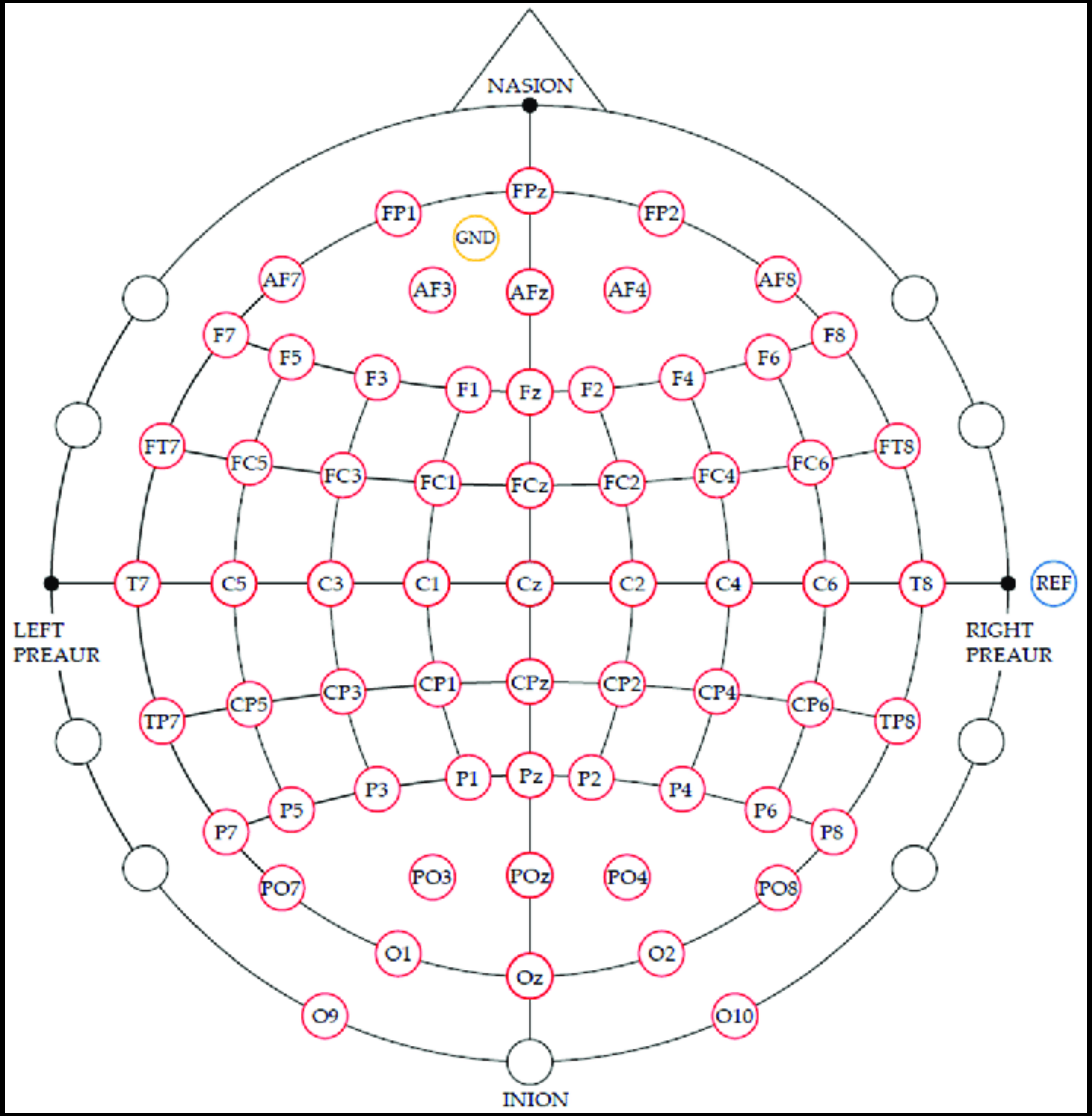


Fig 1. 64 EEG Electrodes Placed on the Scalp [3]

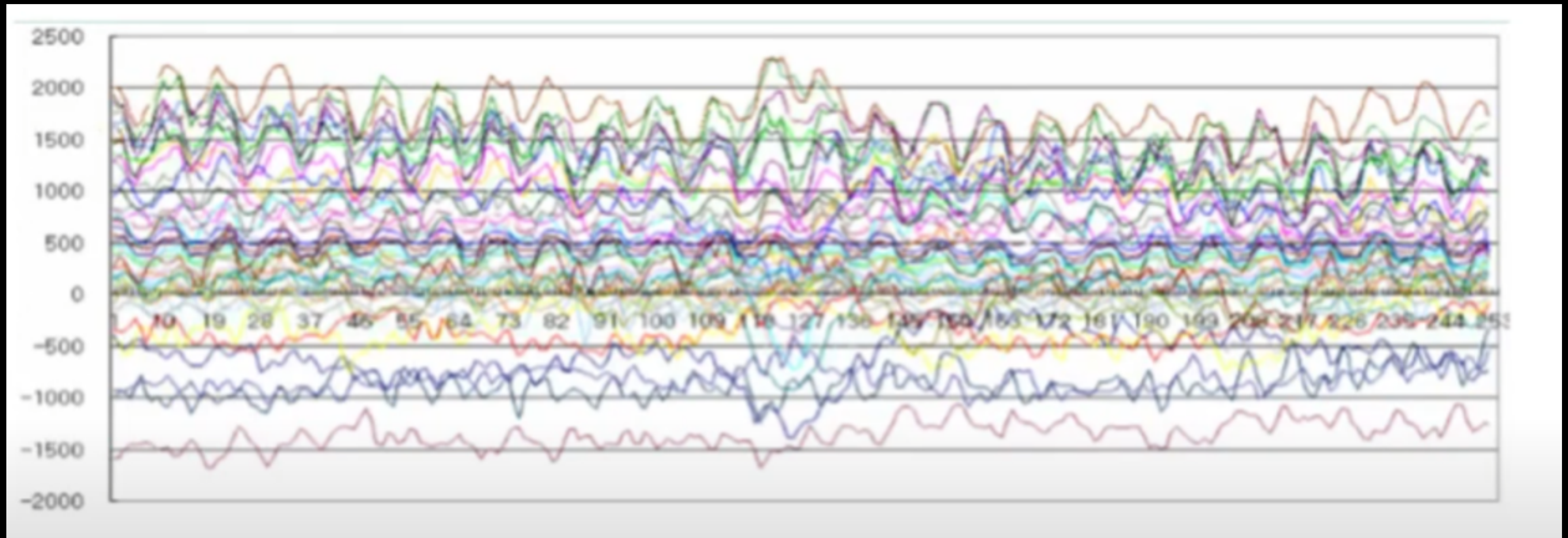
[3] Hayta, Ünal, et al. "Optimizing Motor Imagery Parameters for Robotic Arm Control by Brain-Computer Interface." Brain Sciences 12.7 (2022): 833.
[4]Shriram, Revati, Mahalingam Sundhararajan, and Nivedita Daimiwal. "EEG based cognitive workload assessment for maximum efficiency." Int. Organ. Sci. Res. IOSR 7 (2013): 34-38.

Literature Reviews

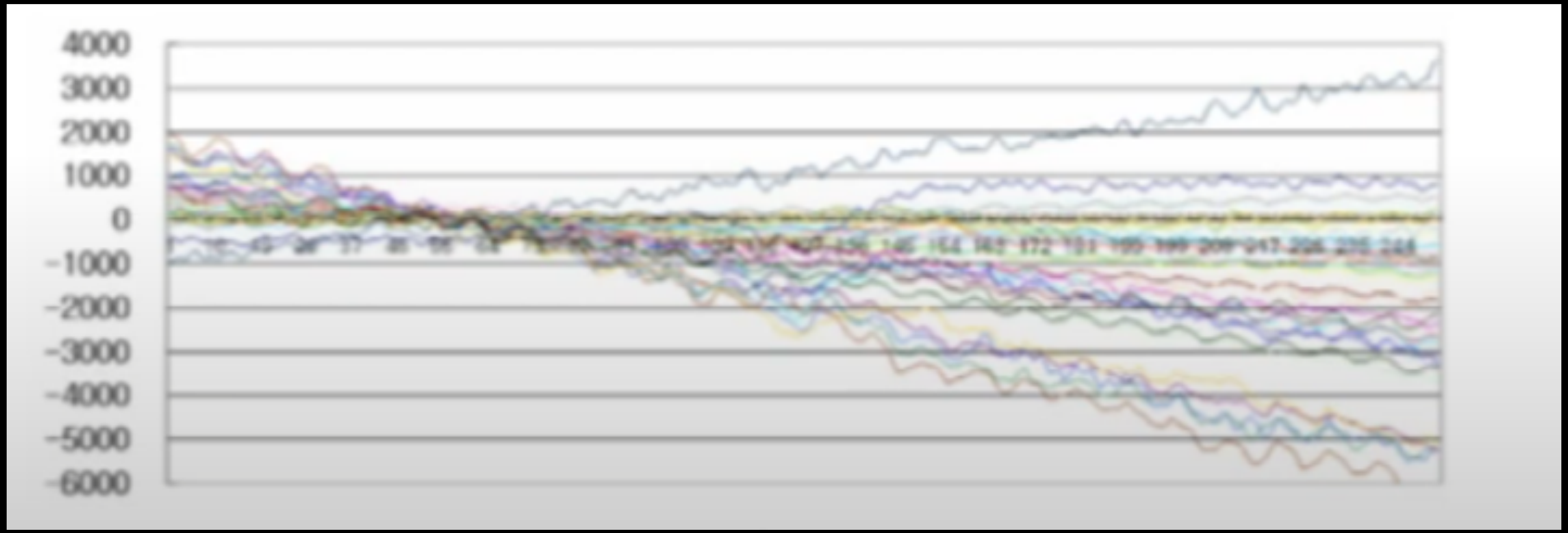
EEG Baseline Corrections

- Complex Brain Activity
- Muscle Tension
- Sweating
- Other Noises

➡ The zero level of each channel becomes different



Required Constant Trend Removal With Baseline Interval Average of Each Channel



Entailed Linear Trend Removal

Literature Reviews
DEAP Dataset [2]

- Composed of 22 participants' video and EEG signal recordings which were recorded while watching 40 music video clips
  - Another 10 participants were only recorded EEG signals.
  - 32 Participants, in Total
- Labeled by participants themselves about Arousal, Valence, Liking, Dominance, and Familiarity
  - Familiarity (Discrete Scale of 1-5)
  - Others (Continuous Scale of 1-9)
- 10-20 system, 32 channels

[2] S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, Jan.-March 2012, doi: 10.1109/T-AFFC.2011.15.

Literature Reviews
CNN

- **Neocognitron**, 1983 [7],
  LeNet-5, 1998 [8]

- Convolution
- Stride
- (Max) Pooling
- Dense

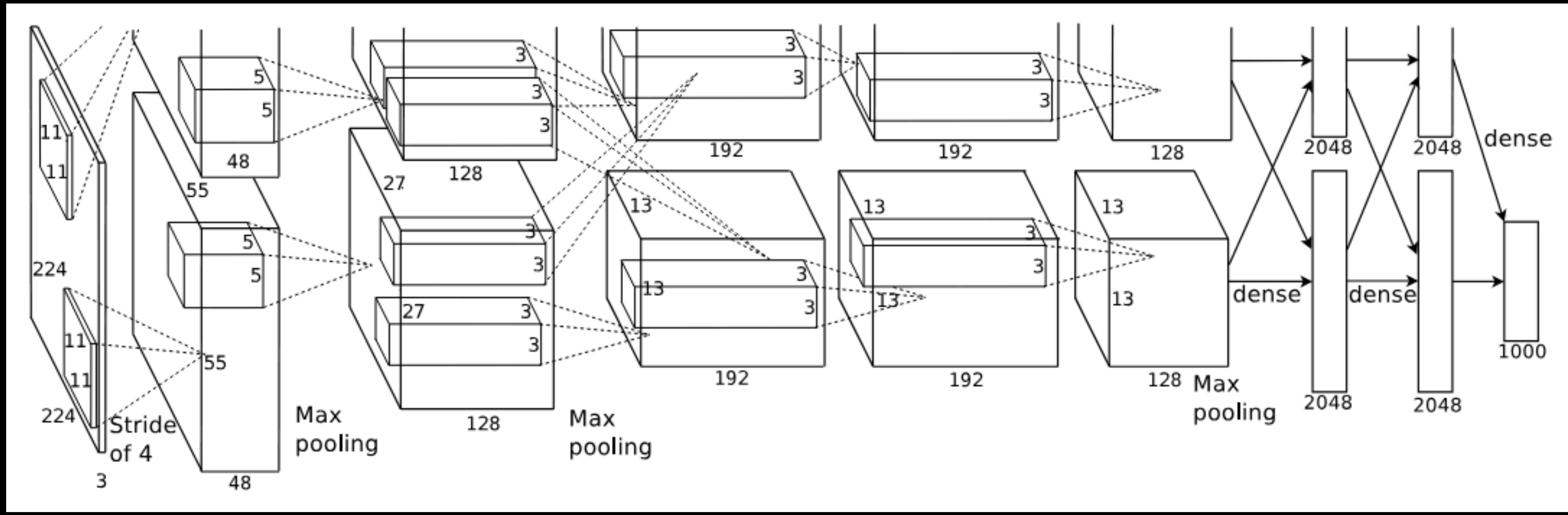


Fig 3. AlexNet Architecture [9]

[7] Fukushima, Kunihiko, Sei Miyake, and Takayuki Ito. "Neocognitron: A neural network model for a mechanism of visual pattern recognition." IEEE transactions on systems, man, and cybernetics 5 (1983): 826-834. [8] LeCun, Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.
[9] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." Communications of the ACM 60.6 (2017): 84-90.

Literature Reviews
LSTM [10]

- A Sort of Neural Network Proposed to Address Time-series Data

- Introduced a special type of units different from RNN

- Forget / Input / Output Gate

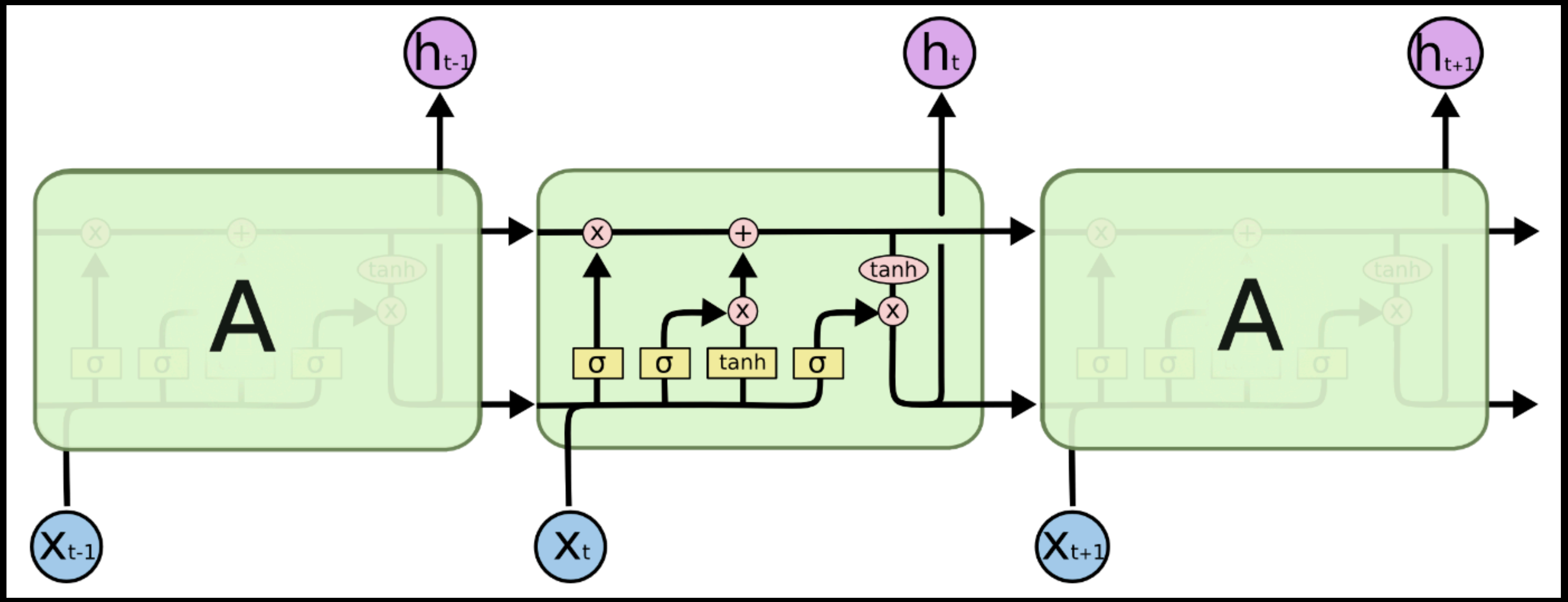


Fig 4. A LSTM Cell structure [11]

[10] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
[11] Christopher Olah. "Understanding LSTM Networks." Aug. 27, 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/. accessed in May. 30, 2024.

Experiments

Experiments

EEG Dataset (DEAP)

- Compared for Arousal and Valence accuracy
  - Followed the standard of related works
    - Labeled “High”, if the value was higher than 5.5
    - Labeled “Low”, if the value was lower than 4.5

- 5-fold cross validation
- Data Shuffling

Experiments

EEG Preprocessing (1)

- X_b: Baseline (Interval)
- X_t: Actual Trial Signal
- K: The Length of X_b (in Sample)
- M: The Length of X_t (in Sample)

- X_p: Preprocessed Signal

- $\bar{X_b} = \frac{\sum_{n=1}^{K} X_b^n}{K}$
- $X_p^j = X_t^j - \bar{X_b}$

➡ Constant Trend Removal

Experiments

EEG Preprocessing (2)

- $\{X_p^j \rightarrow S \mid S = [S_1, S_2, \cdots, S_n]\}$
- S: An Array of **Non-overlapping** Slices
- Each $\{S_k \mid k \in N\}$ had the number of samples which represents **3 seconds**.
  - The state of human emotion persists between 1 and 12 seconds.
- Transforming the EEG signal into the spatial info
  - EEG Signal → Spectrum (FFT) → 2D Image (The Sum of Squared Absolute Value)

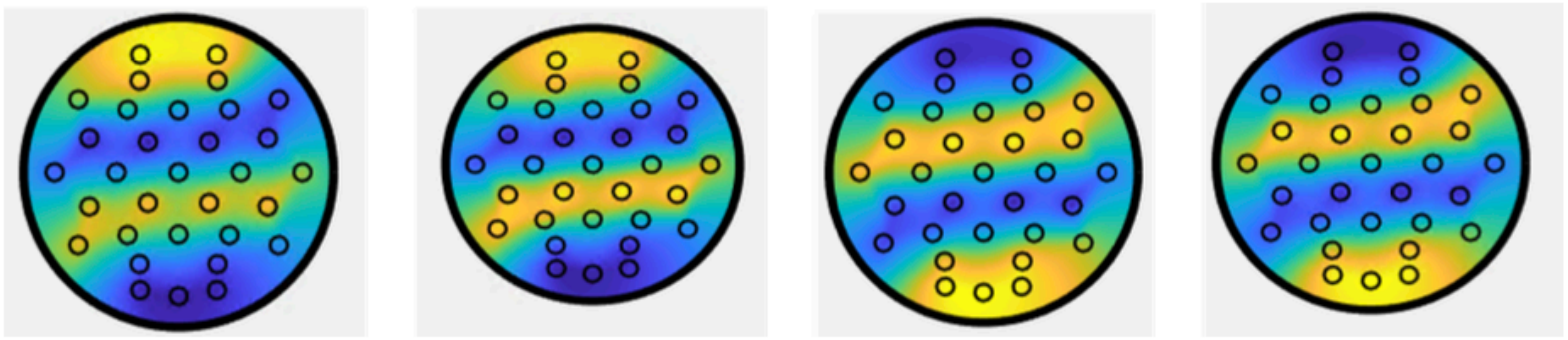


Fig. 2: Sample visualization of the EEG representation for different video effects. The circles inside the images are the locations of the electrodes which are shown only for visualization purposes.

Experiments

Model Architecture (1)

- CNN sub-model
  - Got 2D Images
  - **Five** Convolutional Layers
    - Max-pooling, Batch Normalization, SELU followed the each layer
  - **Two** Fully Connected Layers
  - A Spatial Attention Layer

- LSTM sub-model
  - Got Time-series Signals
  - **Three** LSTM Layers
  - Dropout
  - A Temporal Attention Layer

Experiments

Attention Mechanisms

- **Temporal Attention**
  - $z_t = W^T \sigma(W_1 h + W_2 q + b_1) + b$
    - σ, SELU [1]
    - h: LSTM Feature Vector
    - q: Aligned Pattern Vector from h ?
  - $p_t = \text{softmax}(z_t \cdot h)$
  - $A_t = p_t \cdot h$

- **Spatial Attention [12]**
  - $\hat{F} = \text{MaxPool}(F) \quad (\hat{F} \in R^{C \times H \times W})$
    - $F \in R^{C \times H \times W}$: CNN output
  - $p_s = \text{softmax}(W\hat{F} + b)$
    - W, b: Convolution Parameters
  - $A_s = p_s \times F$

[1] Klambauer, Günter, et al. "Self-normalizing neural networks." Advances in neural information processing systems 30 (2017).
[12] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV). 2018.

Experiments

Model Architecture (2)

- Fusion Model
  - Two vectors from both CNN and LSTM sub-model are fused (concat? or point-wise?)
  - Two Fully Connected Layers
    - Softmax Function

- Reported Hyperparameters and Configurations
  - Cross-entropy Loss Function
  - Adam Optimizer
    - L2 Regularization
  - Learning Rate 1e-4
  - 80 Epochs

Results

Results

Summary

- Discriminative EEG features were extracted, leading the classification performance to the highest among the listed models in the paper for the DEAP dataset
- A Fusion with Attention was critical

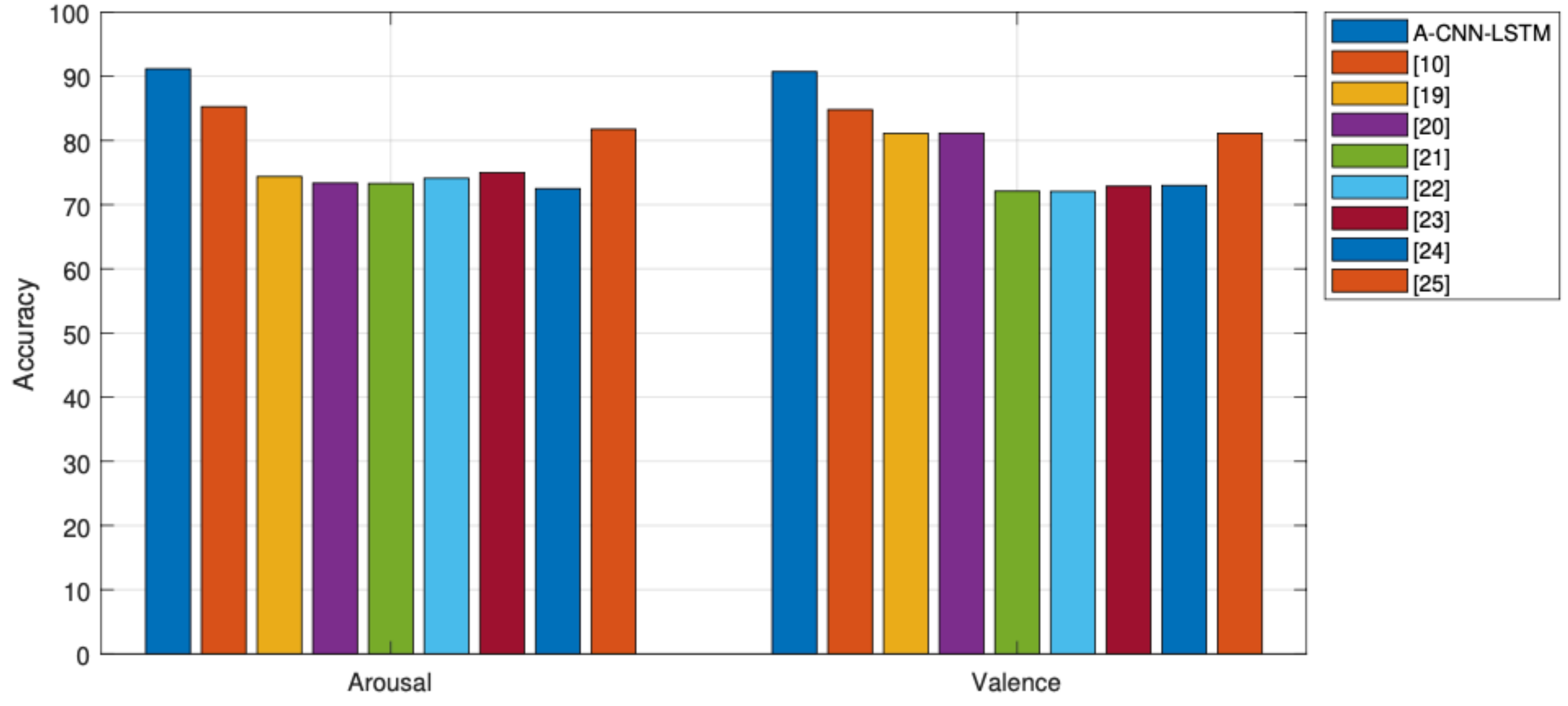


Fig. 3: Accuracy results for the SOTA compared to the proposed approach.

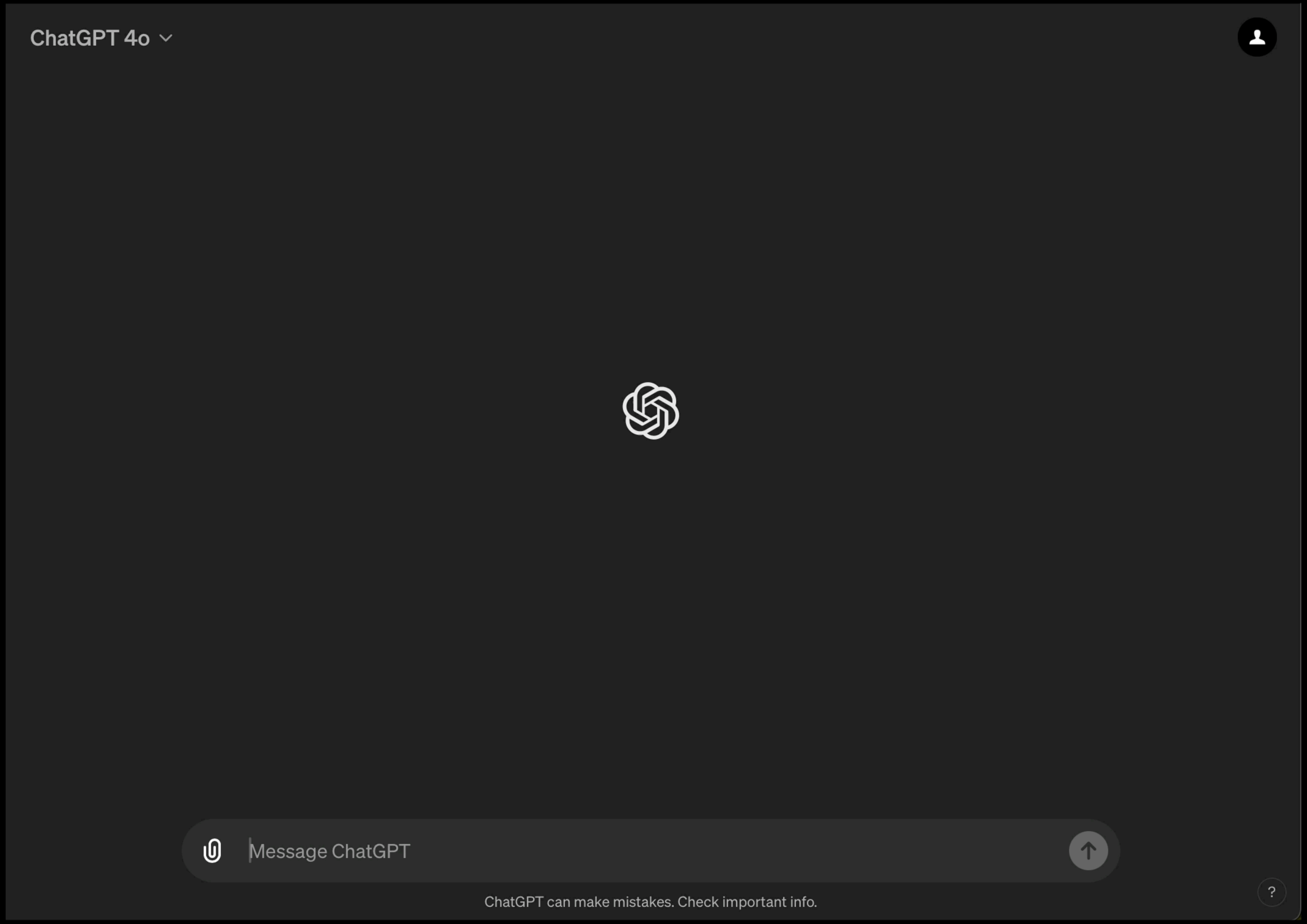
Table 1: Accuracy results for the proposed model with different variations. The 'A' in the first row stands for Attention.

Model	A-CNN-LSTM	A-CNN	A-LSTM	CNN-LSTM
Arousal Accuracy	91.17	85.24	81.93	80.74
Valence Accuracy	90.73	83.88	80.36	80.25

Q&A

Wonjun Park
kuwjjgjk@konkuk.ac.kr

Homepages
github.com/dev-onejun
linkedin.com/in/dev-onejun



References
MLA and IEEE Formats

1. Klambauer, Günter, et al. "Self-normalizing neural networks." Advances in neural information processing systems 30 (2017).
2. S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, Jan.-March 2012, doi: 10.1109/T-AFFC.2011.15.
3. Hayta, Ünal, et al. "Optimizing Motor Imagery Parameters for Robotic Arm Control by Brain-Computer Interface." Brain Sciences 12.7 (2022): 833.
4. Shriram, Revati, Mahalingam Sundhararajan, and Nivedita Daimiwal. "EEG based cognitive workload assessment for maximum efficiency." Int. Organ. Sci. Res. IOSR 7 (2013): 34-38.
5. Delorme, A. EEG is better left alone. Sci Rep 13, 2372 (2023). https://doi.org/10.1038/s41598-023-27528-0
6. de Cheveigne, Alain. "Is EEG best left alone?." bioRxiv (2023): 2023-06.
7. Fukushima, Kunihiko, Sei Miyake, and Takayuki Ito. "Neocognitron: A neural network model for a mechanism of visual pattern recognition." IEEE transactions on systems, man, and cybernetics 5 (1983): 826-834.
8. LeCun, Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.
9. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." Communications of the ACM 60.6 (2017): 84-90.
10. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
11. Christopher Olah. “Understanding LSTM Networks.” Aug. 27, 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/. accessed in May. 30, 2024.
12. Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV). 2018.