**Suggestion for Improvements Upon Existing State-of-The-Art Deep Learning Neural Network for Electroencephalography-Based Emotion Classification**

Arnold Cheskis, Qi Yang, Karen Kuang, and Guadalupe Bernal

## Introduction

Imagining a near future, it's been a rough day for you. You've just had a bitter argument with your partner, and to add insult to injury, your arch-rival at work has snagged the promotion you were gunning for. Feeling defeated, you merge with the couch, inundated by complete apathy. In an attempt to find some distraction, you ask your smart home assistant to play some music. "Sure!" it replies in its usual cheerful yet oblivious tone. The assistant selects a generic, peppy pop track—a stark mismatch to your gloomy mood. The cheerful beats feel like a series of jabs to your already low spirits. You can't help but think, "If only this thing had half a brain, it’d know I'm not in the mood to party."

**Our team is addressing this problem by exploring the potential of deep learning to decode complex brain activity patterns**. Specifically, we are using electroencephalography (EEG) data to identify various emotional states. This research falls under the umbrella of affective computing, a field dedicated to endowing machines with the ability to detect and respond to human emotions. Such capabilities are essential for creating more intuitive, adaptive, and responsive interactions between humans and computers.

Traditional methods of assessing emotions in affective computing have relied on observable factors like facial expressions, speech patterns, textual analysis, and physical gestures [1]. However, these can often be influenced by the social context. EEG signals offer a unique advantage as they provide a direct measurement of brain activity that is less influenced by these external factors, making them a more consistent predictor of emotional states.

The practical applications of accurately detecting emotions from EEG data are varied. They range from assisting individuals with disabilities in expressing their emotions to enhancing learning experiences by identifying and addressing frustration. Furthermore, this technology can contribute to the advancement of social robotics, market research, and entertainment by facilitating a deeper understanding of human emotions. With the ongoing improvement of consumer-grade EEG headsets and the evolution of brain-computer interfaces [2], the process of acquiring EEG data is becoming increasingly accessible. This shift underscores the importance of developing sophisticated algorithms for emotion detection from EEG data.

Specifically within deep learning, we will attempt to implement a CNN with the justification described further in the Background section. We hypothesize that CNN will be able to classify the emotions based on the EEG data.

In the realm of deep learning, our primary focus will involve the refinement of state-of-the-art CNN structures. We will be looking at the structure described in the 2021 paper “Automated accurate emotion recognition system using rhythm-specific deep CNN” [3] that achieved high results on the DREAMER dataset. Moreover, we aim to explore a more experimental avenue by incorporating and training a TSception model, as justified in the Background section. Our hypothesis posits that both the CNN and the TSception structures will not fall behind the state-of-the-art results, and classify the emotions with a high accuracy based on the EEG data.

## Problem Statement and Background

EEG data is inherently multidimensional, characterized by its time-series structure that captures the brain's electrical activity over time. Each electrode placement on the scalp records voltage fluctuations resulting from ionic current within the neurons of the brain, providing a distinct time-varying signal. These signals are typically described in terms of frequency bands, such as delta, theta, alpha, beta, and gamma waves. The data is often noisy and influenced by artifacts that may arise from muscle movements, eye blinks, or external electrical sources, requiring careful preprocessing before analysis.

Emotion recognition from EEG presents a unique deep learning challenge due to its high dimensionality, significant individual variability, and the temporal nature of its signals. The non-stationary aspect of EEG signals, where the statistical properties can change over time, requires models that can adapt to shifting data distributions. Moreover, the presence of noise and artifacts within the EEG signals necessitates robust feature extraction methods, which some deep learning might provide.

To address these problems, there are multiple model structures that can provide an accurate classification of the EEG dataset. The main question we wish to address before we even begin our research is which model structure is the highest-performing one. In this section, we will show the CNN network combined with an SVM is currently the highest-performing model for data classification on the DREAMER dataset.\*

We will start by surveying the models that are commonly used to classify the EEG DREAMER dataset: SVM, CNN with ELM, and CNN with SVM. Following the 2021 review paper on emotion recognition based on EEG [5] which compared the performance of commonly used models on several datasets, including DREAMER, we found that the best-performing structures, as shown in Table 1, are a CNN and a regional-asymmetric convolution neural network (RACNN). The CNN structure is part of a multi-channel EEG emotion recognition algorithm [3]. For the CNN the accuracy of valence, arousal, and dominance in the DREAMER dataset was 97.17, 96.81, and 97.24%, respectively. The valence and arousal accuracy of the RACNN algorithm on the DREAMER dataset were 95.55 and 97.01%, respectively [5]. We see that RACNN outperformed the CNN model in classifying arousal, and the CNN performed better with valence. The paper researching the performance of the RACNN model [4] did check its accuracy in classifying dominance. Due to this fact, and the insignificant difference in accuracy when measuring arousal, we have elected to implement the CNN model as described in the 2021 paper Automated accurate emotion recognition system using rhythm-specific deep CNN [3].

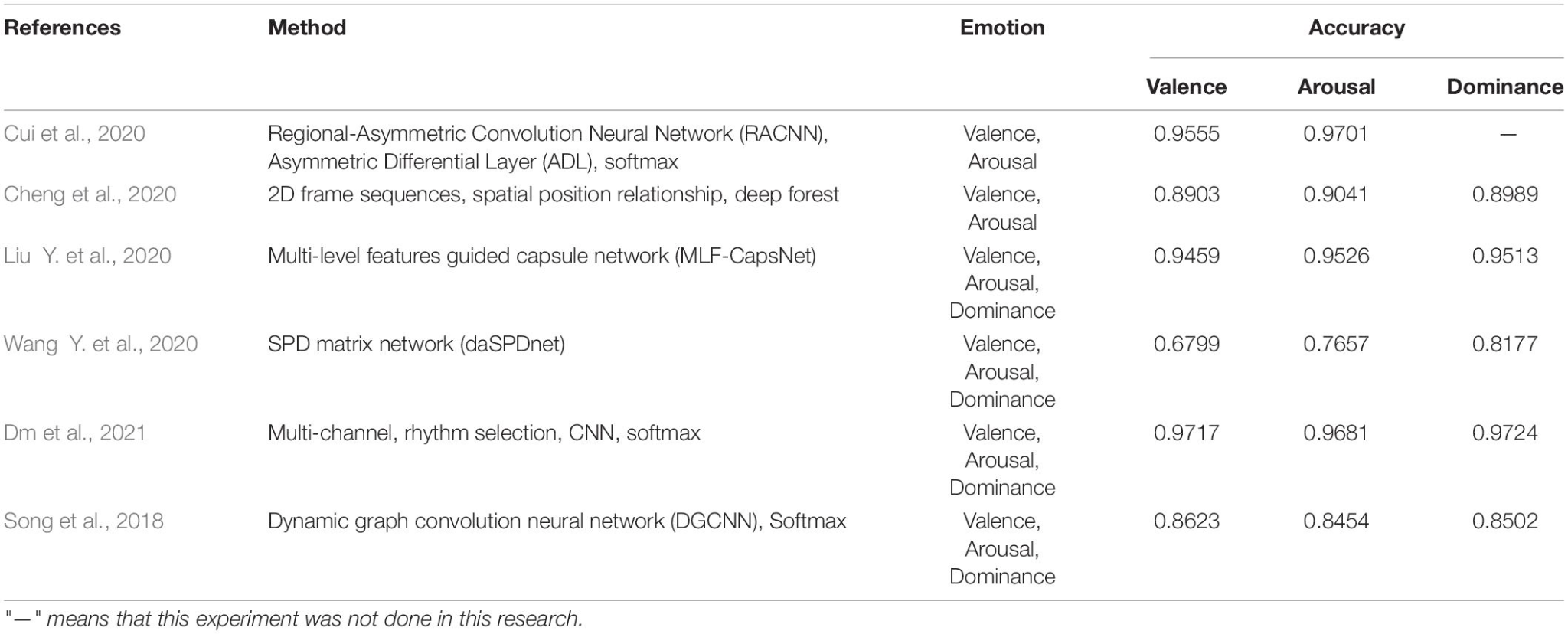


Table 1. Research results based on DREAMER

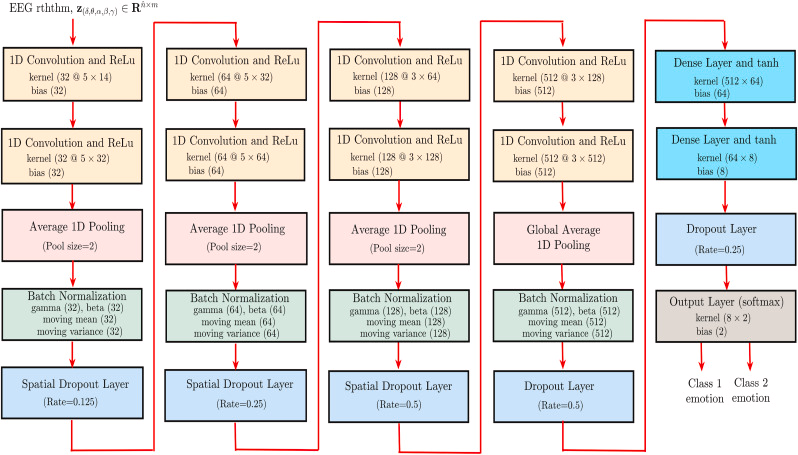


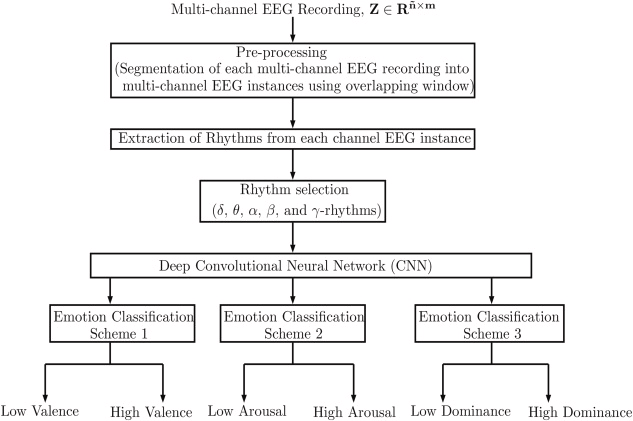
Figure 1. Rhythm-specific Deep CNN architecture for the automated emotion recognition using multi-channel EEG signals from the DREAMER database.

Fig. 2. Flow-chart for the proposed approach for emotion recognition using multi-channel EEG signal

Before feeding the CNN with the DREAMER dataset we will first pre-process it in order to make the data compatible with the model, and in addition, we will augment it, with the hope of outperforming the current state-of-the-art results. The flowchart is shown in Figure 2.

Another paper published on IJCNN in 2020 [6], suggested a new DL framework called the TSception, inspired by the Inception block of GoogleNet [7]. The EEG data that was used was collected from 18 subjects using immersive virtual reality. They suggested a broader application, not particular to the DREAMER dataset. The proposed method was compared with SVM, EEGNet, and LSTM. TSception achieves a high classification accuracy of 86.03%, which outperforms the prior methods significantly (p<; 0.05). Considering the TSception model is built on top of the Inception block of GoogleNet [7], and is also a CNN, we believe it has the potential to achieve high accuracy on the DREAMER dataset as well. We later found an even newer paper [8], which strengthens these suspicions. In this paper, the performance of the proposed network is compared with a larger variety of models SVM, KNN, FBFgMDM, FBTSC, Unsupervised learning, DeepConvNet, ShallowConvNet, and EEGNet. TSception achieves higher classification accuracies and F1 scores than other methods in most of the experiments.

The structure of the Tsception is depicted in Figure 3.

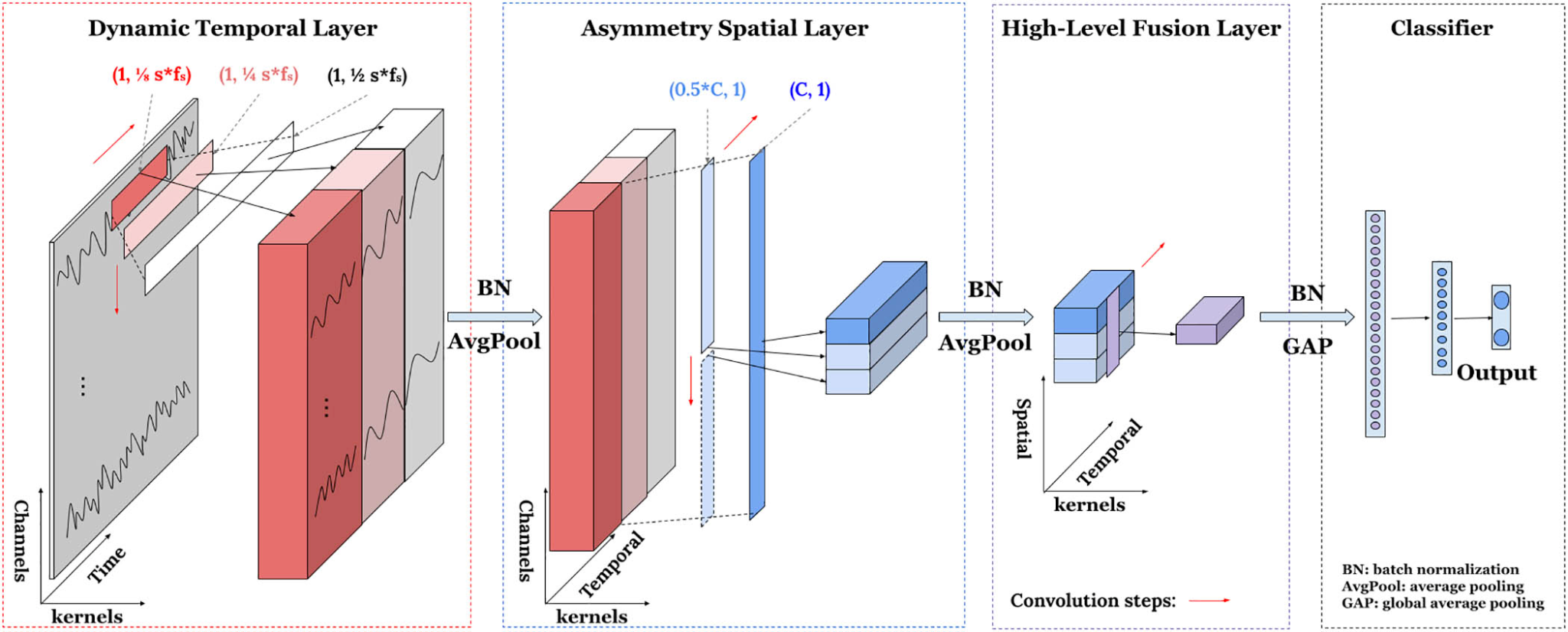


Figure 3: Tsception Model based on DREAMER

In conclusion, our hypothesis is that the 2 proposed CNNs would not fall far behind, and hopefully surpass, the performance of currently commonly used models.

## Dataset Source

In the current literature, some popular datasets included: DEAP, SEED, SEED-V, AMIGO, MPED. The DREAMER dataset, which we selected for this project, contains both raw and preprocessed EEG data [9]. We selected it due to its wide use in the scientific community, overall acceptance of its accuracy during collection, and its open source nature. The DREAMER dataset is a collection of EEG data from 23 individuals who were observed while watching 18 different film clips. These clips were specifically chosen because they are known to evoke a broad spectrum of emotional responses. The data was gathered using portable, affordable, and readily available EEG equipment. EEG signals are essentially time series data. Within this dataset, each sample is structured in a 2D array with the dimensions N x 42, where 'N' denotes the total number of samples and '42' represents the EEG features. 42 is arrived at by multiplying the 14 EEG channels by the 3 band powers (theta, alpha, and beta) of each channel.

As for the labeling of the data, it employs a dimensional approach to categorizing emotions. This method maps emotions across three axes: arousal, valence, and dominance. Arousal assesses the level of alertness, valence the degree of pleasantness, and dominance the extent of control felt by an individual. For instance, happiness is typically characterized by high valence and dominance, with arousal levels that can vary from high to low. In this dataset, emotional ratings across these three dimensions are assigned on a scale from 1 to 5. They are further classified as 'Low' or 'High' based on whether the rating given by a participant falls below or above the midpoint of 3 on this scale.

We requested access to this dataset from zenodo.com, where the authors of the work published it. We were approved and now have permission to use the dataset for this class project. We have imported the dataset to google drive and have begun working with it in a google collab document.

## Methods and Tools:

### Data-Preprocessing and Feature Selection

Our denoised data come in 14 channels: alpha, beta, gamma, theta waves, etc. In order to record the spatial information of the EEG signals, we will adopt algorithms to generate a holographic features map (HOLO-FM) to represent our data at a higher level [10]. To construct a HOLO-FM, the characteristics of every EEG channel sub-band will be calculated and represented in a three-dimensional space. After that, the three-dimensional representation will the mapped to a two-dimensional plane showing as images. Because the DREAMER dataset has 14 channels, it is expected to produce a lot of noise in the final image. We will adopt the method by Lobaz (2018) to mitigate the noise with an exponential function to smooth the rips and to get more precise values of the feature [11]. After this, we will apply a color scheme to generate the final image. All of the above procedure is expected to be achieved with Numpy and Seaborn packages.

### Model Construction

Our model will consist of three parts: extraction, fusion, and classification. (Topic and Russo, 2021) We will apply the Convolutional Neural Network (CNN) to extract features from HOLO-FM per characteristics (per channel) of the EEG signal. By fusing the extracted features, we will contract a features matrix that sums up features learned by the neural network. Utilizing these feature matrices, we will then be able to perform emotion classification using SVM.

We are choosing CNN because it is a widely used model for learning and extracting features for classification tasks in various fields. We will use CNN with a small number of layers to reduce complexity and runtime, which has also been proven effective in EEG classification by Topic and Russo (2021). The color images will be fed to the network to learn high-level features in the hidden layers. The first few convolutional layers will learn simple features and the later ones, including convolutional and max-pooling layers, will learn high-level combinations of features learned in previous layers. The extracted features will be fused together to be the input for SVM classification in the final stage. Different learning rate, regularization, and batch size will the experimented while we optimize the model. We expect to split the training and testing dataset in the ratio of 4:1 and use cross-validation to evaluate model performance at the end. The above model will be implemented using PyTorch.

## 

## References

[1] Wang, Yan, Wei Song, Wei Tao, Antonio Liotta, Dawei Yang, Xinlei Li, Shuyong Gao et al. "A systematic review on affective computing: Emotion models, databases, and recent advances." *Information Fusion* 83 (2022): 19-52.

[2] Vasiljevic, Gabriel Alves Mendes, and Leonardo Cunha De Miranda. "Brain–computer interface games based on consumer-grade EEG Devices: A systematic literature review." *International Journal of Human–Computer Interaction* 36, no. 2 (2020): 105-142.

[3] Maheshwari, Daksh, Samit Kumar Ghosh, R. K. Tripathy, Manish Sharma, and U. Rajendra Acharya. "Automated accurate emotion recognition system using rhythm-specific deep convolutional neural network technique with multi-channel EEG signals." *Computers in Biology and Medicine* 134 (2021): 104428.

[4] Cui, Heng, Aiping Liu, Xu Zhang, Xiang Chen, Kongqiao Wang, and Xun Chen. "EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network." *Knowledge-Based Systems* 205 (2020): 106243.

[5] Liu, Haoran, Ying Zhang, Yujun Li, and Xiangyi Kong. "Review on emotion recognition based on electroencephalography." *Frontiers in Computational Neuroscience* 15 (2021): 84.

[6] Ding, Yi, Neethu Robinson, Qiuhao Zeng, Duo Chen, Aung Aung Phyo Wai, Tih-Shih Lee, and Cuntai Guan. "Tsception: a deep learning framework for emotion detection using EEG." In *2020 international joint conference on neural networks (IJCNN)*, pp. 1-7. IEEE, 2020.

[7] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[8] Ding, Yi, Neethu Robinson, Su Zhang, Qiuhao Zeng, and Cuntai Guan. "Tsception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition." *IEEE Transactions on Affective Computing* (2022).

[9] Katsigiannis, Stamos, and Naeem Ramzan. "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices." *IEEE journal of biomedical and health informatics* 22, no. 1 (2017): 98-107.

[10] Topic, Ante, and Mladen Russo. "Emotion recognition based on EEG feature maps through deep learning network." *Engineering Science and Technology, an International Journal* 24, no. 6 (2021): 1442-1454.

[11] Lobaz, P. (2018). CGDH Tools: Getting started in computer generated display holography. In *Proceedings of the 11th International Symposium on Display Holography—ISDH* (pp. 159-165).