GA-ANFIS and CRNN for Mental Attention State Classification using Passive BCI-based EEG data

Mian Muhammad Awais[1], Zoraiz Qureshi[2], Syed Ali Hassan Bukhari[3]

[1 2 3] Department of Computer Science, Lahore University of Management Sciences, Lahore, Pakistan, 54000
awais@lums.edu.pk, 21100130@lums.edu.pk, 21100187@lums.edu.pk

Abstract. By using electroencephalography (EEG) based BCI intrinsic or passive activity data self-generated by specified individuals under simulation or obtained live [3], we aim to detect and classify the current mental attention state of an individual into several categories of states, the most prominent being - focused, unfocused (may also be described as lost-in-thought or mind-wandering) and drowsy (or sleep). Our approach uses a hybrid neuro-genetic fuzzy system optimized and trained on a select number of features and channels extracted from the data as well as a shallow convolutional and convolutional recurrent neural network trained on the raw EEG signal data to predict human mental attention state with signals of trial lengths as short as 6 seconds.

Keywords: EEG, Brain-Computer Interface, Mental Attention State Classification, ANFIS, Genetic Algorithm, Convolutional Neural Networks

1 Introduction

Specific frequency bands have been detected at the frontal and centroparietal lobes of the brain have been found to be linked to the human mental attention state [1]. An experiment has shown that the delta, theta and beta activities have been found to increase more in a mental focused state at the frontal lobe than an unfocused one [2]. By using electroencephalography (EEG) based BCI intrinsic or passive activity data self-generated by specified individuals under simulation or obtained live [3], we aim to detect and classify the current mental attention state of an individual into several categories of states, the most prominent being - focused, unfocused (may also be described as lost-in-thought or mind-wandering) and drowsy (or sleep). The classification of an individual's mental attention state has various applications such as monitoring control tasks where high focus and concentration levels have to be maintained ranging, such as the detection of driver fatigue or classroom student attention [4].

However, there are several problems with the recognition of a BCI user's cognitive state that dynamically varies and affects classification performance along with the lack of interpretability and availability of BCI obtained datasets. A complete EEG data processing and classification pipeline will be designed and developed including a neural network-based classifier as well as a hybrid neuro-genetic fuzzy system (HNGFS) for this task.

2 Related Works

A prominent study in the area used 5 different models for performing 2 different tasks: detecting stress and no-stress states; and classifying stress levels into high and low stress [5]. These models were Linear Discriminate Analysis (LDA), K-Nearest Neighbor (KNN), Linear and Cubic Support Vector Machine (SVM) Classifiers, and Random Forest Classifiers. Of these, Cubic SVMs and KNNs performed better than all other metrics in most of the experiments. Another research study targets classification of human mental attention states into focused, unfocused and sleep classes with machine learning methods such as SVM and KNN as well as ANFIS [6]. It establishes a data processing pipeline for 25 hours of EEG-based passive BCI data obtained from 5 different participants. SVM yields the best classification results, achieving a high accuracy (96.7% best, 91.72% avg) after a comparative analysis. One study even showed how LDA, KNN, SVM, Naive Bayes, and Decision Trees could be used together (ensembled bucket of models) to improve classification accuracy of different emotional states [7]. The same paper discussed how dimensionality reduction techniques like Sequential Feature Selection (SFS) could be used to significantly improve classification accuracy by simplifying the feature space.

Recent studies have also achieved optimal performance utilizing deep learning-based classifiers on this domain and its associated applications. One such research work dives into a crucial application of identifying mental attention states for drivers specifically to detect and prevent the adverse consequences of driver fatigue, a major cause of traffic accidents [8]. For this purpose, it trains deep convolutional neural network-based classifiers along with residual learning on 15 channels of EEG signals and physiological signal data obtained from users operating driving simulation software. The study develops and presents two resultant mental state classifiers EEG-Conv suited for inter-subject mental variability and EEG-Conv-R which integrates residual learning for faster training convergence, delivering better performance than prior LSTM, HMM and SVM-based classifiers. Another paper designs and employs an end-to-end CNN based framework using transfer learning for inter-subject classification of continuous attention and diverted attention from EEG BCI data results which are also evaluated against a neurophysiological assessment test [9]. Furthermore, the study also presents an experiment to investigate the effects of movement intent on detection of attention diversion and its corresponding BCI performance. Finally, to counter the adverse impact on BCI performance and restore it, the study presents a GAN model for EEG data augmentation under attention diverse, achieving interpretable results and a more generalizable pipeline.

3 Methodology

By using electroencephalography (EEG) based BCI intrinsic or passive activity data self-generated by specified individuals under simulation or obtained live [3], we aim to detect and classify the current mental attention state of an individual into several categories of states, the most prominent being - focused, unfocused (may also be described as lost-in-thought or mind-wandering) and drowsy (or sleep). For this purpose, a dataset obtained from previous research study by C.I. Acı et. al. on mental attention state classification using an SVM was used which was also available on Kaggle [1]. The data obtained was from the EMOTIV BCI headset and from 5 different subjects under a stress driving simulation with focused, unfocused and drowsy phases. After habituation data, each of the 5 subjects had 5 experimental trials associated (except for the last subject) so there were 24 trials of 35 to 55 minutes each. Each experiment was conducted with the following labelled durations for the state of the participant subject as per the state of a driving simulation: focused from the 0 to 10 minute mark, unfocused from the 10 to 20 minute mark and drowsy/sleep from 20 minutes to the end of the experiments. The 14 channels which would act as our decision variables featuring BCI EEG based data only were utilized from the 25 channels obtained from the EMOTIV headset given (channel names as follows: 4-'EDAF3' 5-'EDF7' 6-'EDF3' 7-'EDFC5' 8-'EDT7' 9-'EDP7' 10-'EDO1' 11-'EDO2' 12-'EDP8' 13-'EDT8' 14-'EDFC6' 15-'EDF4' 16-'EDF8' 17-'EDAF4'). Given conditionally short trials of at least 6 seconds, using these 14 channels as our training features our goal is to predict in real time which of the 3 mental attention state classes associated with that trial.

3.1 Data Preparation

The sampling frequency was 128 Hz and conditionally, we took samples of fixed splits of 6 seconds each were extracted and used with one-hot encoded categorical labels for the corresponding one of the 3 classes and merged training samples by class instead of on a subject or experiment basis with a training-test split of 80:20 and training-validation split of 80:20 as well. Pre-processing of the raw EEG signals in these trials was done by feature scaling (per-channel normalization based on both mean and standard deviation with a biased estimator using sklearn's Standard Scalar transform) as shown by Figure 3.1.

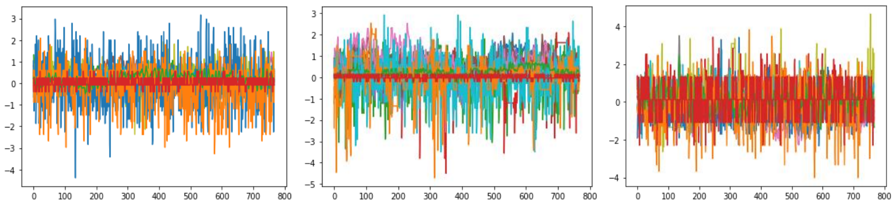

Fig. 3.1. Random feature-scaled ample time-signal plots for the EEG data obtained.

For preprocessing of this data for the ANFIS model which requires a low number of inputs - each of the 14 EEG channels had various time features extracted from it - Number of Zero Crossings, SVD entropy, Perm Entropy, App Entropy, Petrosian's and Higuchi's Fractal Dimension shown in Figure 3.2. These were used as inputs the hybrid ANFIS classifies individually and in combination using feature selection.

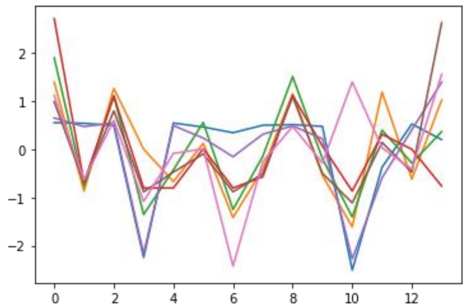

Fig. 3.2. The seven mean-normalized channel features generated for each EEG trial sample.

3.2 CNN & CRNN Model Architecture

Due to the non-linearity of features present in EEG brain signal data which are difficult to perceive and process directly without advanced pre-processing techniques, we created neural network classifiers with different architectural variants to be trained and tested for this classification problem. Both shallow convolutional neural networks (CNN) and convolutional recurrent neural network (CRNN) architectures tested with reference from a previous research study on shallow deep learning networks on BCI based MI-EEG data by Hauke Dose et. al. [10]. Furthermore, we created different deeper CNN and CRNN model variants of the mentioned architectures and also tested them after tuning various parameters. L1 regularization was added as well for optional usage. Finally, we added dropout layers added for regularization after convolutional block and major dense layers, to prevent overfitting.

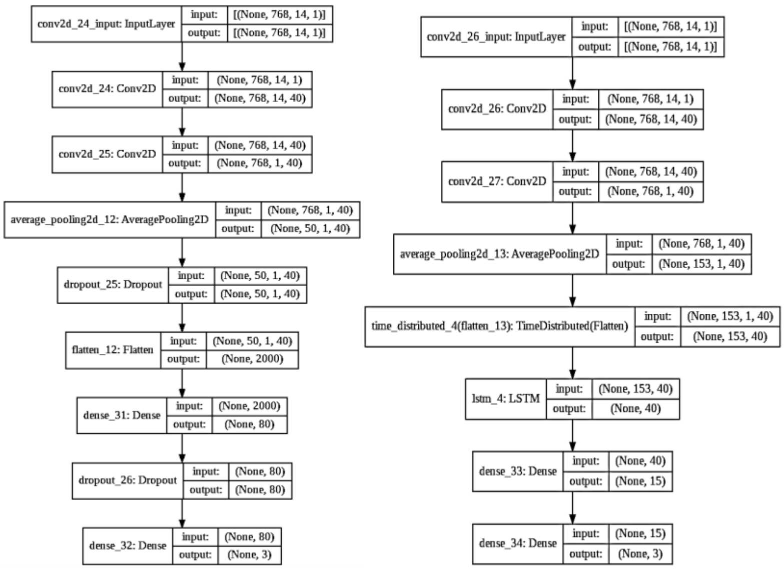

Fig. 3.3. Shallow CNN & CRNN architectures

The CNN architecture uses the first 2D convolution layer to convolve around the time axis across all 14 channels to get temporal features as shown in Figure 3.3. Then it performs vertical spatial feature wise convolutions with the same number of filters as before now preserving both temporal and spatial features which are extracted then by a global average pooling layer which we discovered worked better compared to a max pooling layer. Before the final 3-class output classification dense layer another denser neural network layer was added to learn more complex features from the latent encoded space of this shallow CNN. For optimizers we used Adam with learning rates in the range of 1e-4 to 1e-5.

Similar to the CNN architecture, the CRNN has the same 2D convolutional layers for temporal and spatial feature extraction followed by a time distribution conversion of the flattened latent feature space which is then used by an LSTM layer for temporal sequence classification before the dense layers. This increases the number of parameters so training and convergence time was hypothetically assumed to be more.

3.3 Hybrid ANFIS Design

Along with the convolutional networks, a hybrid fuzzy neuro-genetic network was also designed using the Takagi-Sugeno model and Gaussian membership functions. It was trained on pre-processed forms of the feature-scaled signals per channel. Each of the 14 channels had various time features extracted from it - number of zero crossings, SVD entropy, Perm entropy, App entropy, Petrosian's and Higuchi's fractal dimension. ANFIS classifiers were trained separately on each feature type (14 inputs) as a result to determine the predictability contribution of each of these inputs. To determine the number of rules required for this ANFIS fuzzy inference system rule base and its sensitivity to training performance, a genetic algorithm was applied with the objective to minimize the final training loss given a range of the number of rules (14-1000) to find the optimum value with a fixed number of epochs and learning rate. The ANFIS model was also specifically trained for binary classification between the focused and unfocused mental attention classes, unlike the 3-class CNN/CRNN classifiers due to its architectural limitations.

4 Results

Initially, we tested deeper CNN and CRNN network variants such as ones with 3 dense layers at the end with dropout layers for regularization and 64 convolutional filters instead of 40. This led to early overfitting on training data even after a higher number of epochs and lower learning rates such as 1e-4 but yielding the equivalent or same minimum average loss as the shallow variant. We further found that pool strides of (15,1) of lesser width for the global average pooling layer led to faster convergence and earlier stopping on 700 epochs compared to a pool stride of (30,1). Low learning rates (1e-5) and high number of epochs (700-1000) worked perfectly to achieve the highest accuracy and least loss possible given the training time constraints. A smaller batch size of

16 for a regularizing effect to prevent overfitting. Model checkpointing and early stopping callbacks monitoring minimum validation loss were used to conditionally stop the network from overfitting at the optimum number of epochs with relative patience factors. Furthermore, another callback was used to reduce the initial learning rate on plateaus when validation loss decrease becomes stagnant to a minimum 1e-6 by a factor of 0.2 using a tuned number of patience epochs. From further hyper tuning, a high number of epochs (1000 for the CNN and 200 for the CRNN architecture maximum) were determined to be used along with lower initial learning rates in the 1e-4 to the 1e-5 range were used with the Adam optimizer.

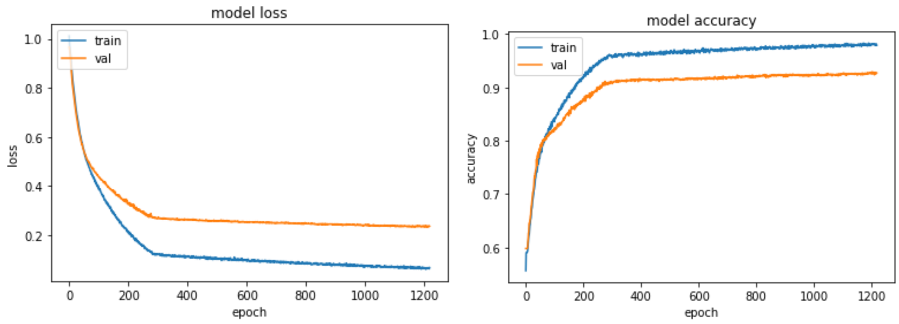

Fig. 4.1. The CNN model achieved a minimum validation loss of 0.235 and test classification accuracy of 92.4% after ~1200 epochs.

The shallow CNN architecture introduced previously achieved a minimum validation loss of 0.235 and test classification accuracy of 92.4% for these 3 classes after 1219 epochs on early stopping as shown by Figure 4.1. It was determined that the neural net was continuously confused between the unfocused and drowsy mental attention states initially with a class balance sample of approximately ~430 training samples each of 6 second long 14-channel EEG preprocessed signal data without any data augmentation. Therefore, an additional of ~860 samples generated from the data present for drowsy data until the end of each experiment were extracted and incorporated, introducing some class imbalance as shown by the confusion matrix in Figure 4.2, however, greatly improving the model's classification accuracy and training performance to the current 92.4% from an 87.1%.

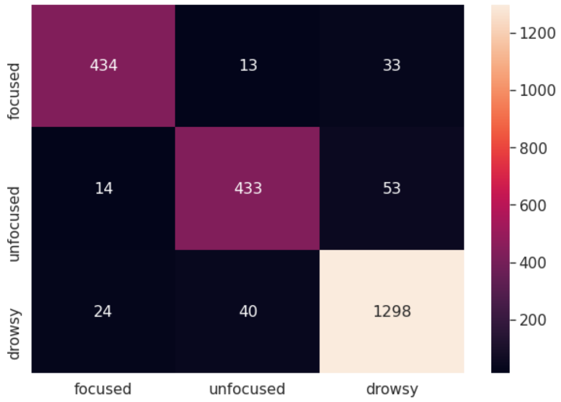

Fig. 4.2. Confusion matrix for the training data EEG trial samples with the 'focused', 'unfocused' and 'drowsy' mental attention states for the CNN & CRNN training with class imbalance.

The CRNN architecture achieved even better performance - a minimum validation loss of 0.124 and test classification accuracy of 95.7% for these 3 classes after 12 epochs (ones with longer duration due to the LSTM layer) on early stopping as shown by Figure 4.3.

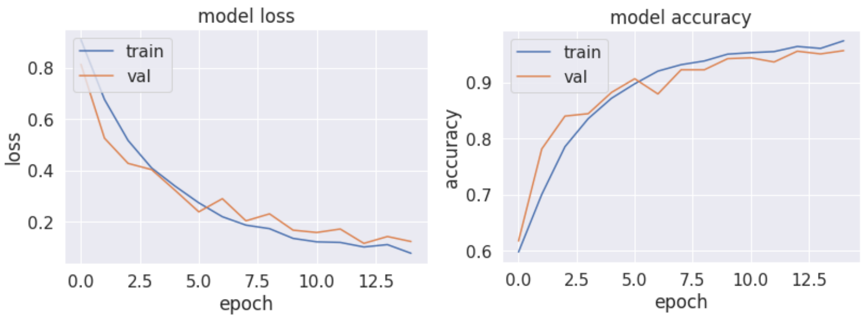

Fig. 4.3. The CRNN model achieved a minimum validation loss of 0.124 and test classification accuracy of 95.7% after 12.5 epochs.

Finally, the hybrid ANFIS model was trained as a binary classifier between focused and unfocused states, and tested on each of the different features extracted per EEG channel from the feature-scaled signal. The number of rules defined for the layers to generate was set to 100 initially. The number of zero crossings as inputs to the model yielded the highest accuracy achieved of 86.1% (training and prediction shown in Figure 4.4) among all other features, followed by Petrosian's Fractal Dimension with 84.1% and Perm Entropy with 83.5% test accuracy in Figure 4.5.

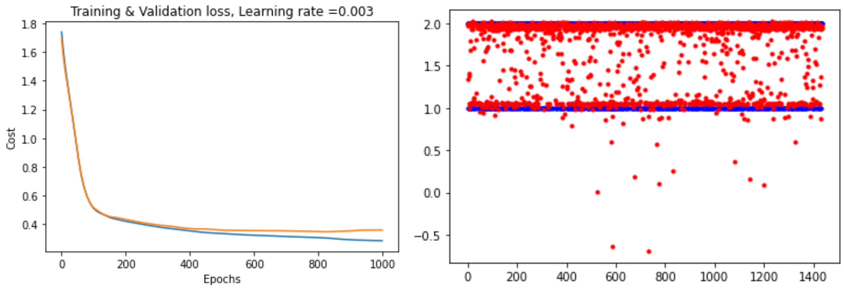

Fig. 4.4. ANFIS training and predictions with 86.1% test accuracy on the per-channel Number of Zero-Crossings after 1000 epochs.

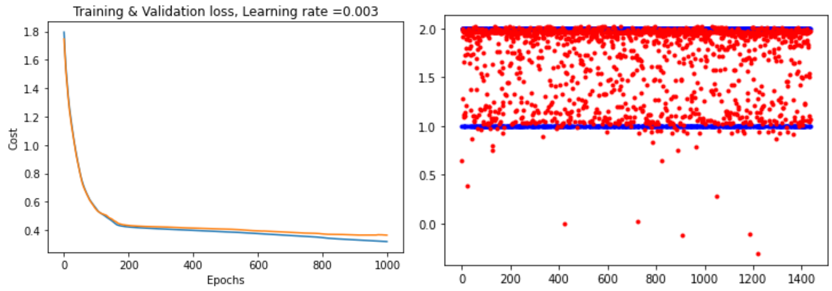

Fig. 4.5. ANFIS training and predictions with 84.1% test accuracy on the per-channel Petrosian's Fractal Dimension.

After testing combinations of these best-performing input features in pairs (28 inputs) and 200 rules (obtained by genetic algorithm optimization on 100 epochs each, aimed to minimize the final validation loss), we found that a combined input of the mean-normalized Number of Zero-Crossings and Petrosian's Fractal Dimension gave us an improved 90.3% test accuracy on 1000 epochs for the HNGFS system. The combination of the Number of Zero-Crossings and Perm Entropy input features performed even better yielding a test accuracy of 91.8% as shown in Figure 4.6.

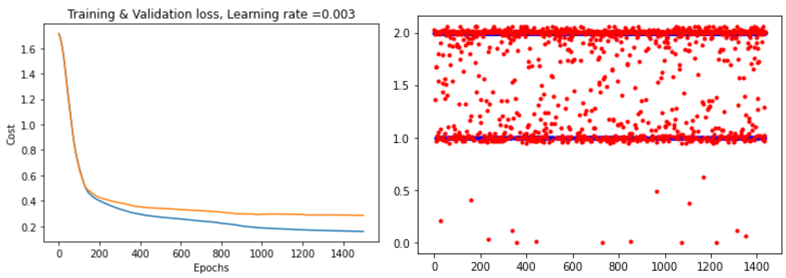

Fig. 4.6. ANFIS training and predictions with 91.8% test accuracy using a combination of 28 inputs from Number of Zero Crossings and Perm Entropy.

5 Discussion & Conclusion

For 3-class classification, C.I. Acı et. al. achieved 96.70% (best) and 91.72% (avg.) experimental classification accuracy using an SVM based classifier on this same 3-class dataset which was generated by them [6]. However, it is important to note that they utilized a much more complex pre-processing and cross validation pipeline comparatively apart from just feature scaling for the raw EEG signal data collected, e.g. the application of the short-time Fourier transform (STFT) and denoising to achieve this high accuracy as EEG signal classification is very sensitive to preprocessing techniques. Moreover, the trial length used by them was longer (15 seconds) compared to our 6 seconds long trials for the same average classification accuracy. A classifier trained on shorter time length trials is able to predict and respond more quickly which is important for the domain of real time mental attention state classification in high hazard scenarios and activities such as driving.

For further implications data augmentation using random time-based splits using a 5-fold cross validation and denoising filters such as high band pass ones can also be applied and conversely, the addition of random noise distributions to the input EEG signals to regularize training further. There is also more work required in feature selection of the channels being passed in and specifying which EEG channels are contributing the most to the predictions at hand, with L1 regularization and optimization of input data by natural-inspired algorithms such as artificial bee colony (ABC). Furthermore, more different normalization techniques including min-max normalization, robust outlier removal normalization and max absolute normalization can also be tested for the signal channels and input features extracted as well as different membership functions than Gaussian for the HNGFS system.

To conclude, our approach used both a hybrid neuro-genetic fuzzy system optimized and trained on a select number of features and channels extracted from the data yielding 91.8% test accuracy for binary classification between the focused and unfocused states. Furthermore, a shallow convolutional and convolutional recurrent neural network trained on the raw EEG signal data to predict human mental attention state with signals of trial lengths as short as 6 seconds yielding 95.7% on 3-class classification with focused, unfocused and drowsy states.

References

1. Groppe, D. M., Bickel, S., Keller, C. J., Jain, S. K., Hwang, S. T., Harden, C., & Mehta, A. D. 2013. Dominant frequencies of resting human brain activity as measured by the electrocorticogram. NeuroImage, 79, 223–233. https://doi.org/10.1016/j.neuroimage.2013.04.044
2. Ko, Li-Wei, Chikara, Rupesh K., Lee, Yi-Chieh, Lin, Wen-Chieh. 2020. Exploration of User's Mental State Changes during Performing Brain–Computer Interface. Sensors 20, no. 11: 3169. https://doi.org/10.3390/s20113169
3. EEG data for Mental Attention State Detection. 2021. Retrieved 15 March 2021, from https://www.kaggle.com/inancigdem/eeg-data-for-mental-attention-state-detection.
4. Gong, Yan & Xu, Samuel. 2019. Mental State Detection in Classroom Based on EEG Brain Signals. Natural Science. 11. 315-322. 10.4236/ns.2019.1111034.

5. Attallah, Omneya. 2020. An Effective Mental Stress State Detection and Evaluation System Using Minimum Number of Frontal Brain Electrodes. Diagnostics (Basel, Switzerland) vol. 10,5 292. doi:10.3390/diagnostics10050292
6. Çiğdem İnan Acı, Murat Kaya, Yuriy Mishchenko. 2019. Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods, Expert Systems with Applications, Volume 134, Pages 153-166, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2019.05.057.
7. Bozhkov, Lachezar & Georgieva, Petia & Trifonov, Roumen. (2014). Brain Neural Data Analysis Using Machine Learning Feature Selection and Classification Methods. Communications in Computer and Information Science. 459. 123-132. 10.1007/978-3-319-11071-4_12.
8. Zeng, H., Yang, C., Dai, G., Qin, F., Zhang, J., Kong, W. 2018. EEG classification of driver mental states by deep learning. Cognitive neurodynamics, 12(6), 597–606. https://doi.org/10.1007/s11571-018-9496-y
9. Fahimi F, Zhang Z, Goh WB, Lee TS, Ang KK, Guan C. 2019. Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI. J Neural Eng. doi: 10.1088/1741-2552/aaf3f6. Epub 2018 Nov 26. PMID: 30524056.
10. Hauke Dose, Jakob S. Møller, Helle K. Iversen, Sadasivan Puthusserypady, “An end-to-end deep learning approach to MI-EEG signal classification for BCIs”, Expert Systems with Applications, Volume 114, 2018, Pages 532-542, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2018.08.031.