Rank-1 CNN for mental workload classification from EEG

Sara Sedlar[1]*, Johann Benerradi[2]*, Côme Le Breton[1]*, Rachid Deriche[1], Théodore Papadopoulo[1], Max L. Wilson[2]

[1]Inria, Université Côte d'Azur, France

[2]School of Computer Science, University of Nottingham, United Kingdom

*authors contributed equally

Brain-computer interfaces (BCIs) can be separated into two main types: active and passive BCIs (Clerc et al. 2016). A BCI can be qualified of passive when the system uses signals involuntarily generated by the user. More specifically, this type of BCI is often used with the aim to assess the mental workload of users performing various task with different levels of mental demand, especially with electroencephalography (EEG) (Wang et al. 2015, Appriou et al. 2018, Shalchy et al. 2020). In most cases those systems are built with a classifier that classifies brain signals into different categories. This relies on having collected labelled data beforehand. However, those systems are often developed in laboratory settings, where both the train and test set have known labels. The "Grand Challenge: Passive BCI Hackathon" organised for the Neuroergonomics 2021 conference enables to challenge researchers with a real-life scenario of a passive BCI: classifying data from unseen sessions, with labels concealed for them, preventing any kind of fine tuning on the test set. The dataset provided for this challenge (Hinss et al. 2021) was composed of EEG recordings of 15 participants performing in 3 distinct sessions the Multi-Attribute Task Battery-II (MATB-II) developed by the NASA. Each session is decomposed in blocks of different difficulties: easy, medium and difficult. The data provided consists in epochs of 2 seconds (with a sampling frequency of 250 Hz) from those blocks for a total 447 epochs for each session and each participant. Difficulty labels were provided only for the 2 first sessions.

EEG signals measured on scalp can be represented as a linear combination of signals produced by sources situated in different cortical regions. Depending on function and position of the sources, signals are characterized by different temporal and spatial patterns. Assuming that the head can be modelled by a sphere, spatial patterns can be represented as a linear combination of spherical harmonics. In order to perform the classification, we propose a deep learning model based on a convolutional neural network (CNN) with rank-1 constraint (Dupré la Tour et al. 2018, Kim et al. 2018) with spherical spatial patterns, which significantly reduces number of trainable parameters. We denote input EEG signals as $X \in \mathbb{R}^{N \times T}$ where ***N*** is the number of sensors (number of sampling points over the sphere), and ***T*** is the number of sampling points over time. Firstly, spatial signals are expressed in terms of spherical harmonic (SH) basis for each time point as $\hat{S} = Y^{inv}X$, $\hat{S} \in \mathbb{R}^{L \times T}$, where ***L*** is the number of SH basis elements and $Y^{inv} \in \mathbb{R}^{L \times N}$ is the matrix containing inverted SH basis. This step is performed in order to reduce inter-subject and inter-session variability due to differences in electrode positions. In addition, under the assumption that EEG signals do not contain very high spatial frequency components, this step allows us to reduce the dimensionality of input data from ***N*** × ***T*** to 16 × ***T*.** Further, we assumed that relevant frequency components are below 20 Hz, which requires sampling rate of at least 40 Hz, so the signals $\hat{S}$ are downsampled by factor 6 over time, denoted as $\hat{S}_{ds}$. The architecture of the model is illustrated in Figure 1. It is composed of 3 convolutional layers, each followed by a max-pooling layer and ReLU. In the first one, convolutions are performed with rank-1 kernels which are outer products of spatial and temporal weights, where spatial weights $\{\hat{w}_i^s\}$ are represented in terms of SH coefficients and temporal weights $\{\hat{w}_i^t\}$ in terms of discrete cosine coefficients, which are transformed to signal domain as $w_i^t = D^T \hat{w}_i^t$, where D^T contains discrete cosine basis. The number of kernels is 5, each containing 16 trainable weights for both spatial and temporal weights and 5 bias terms (165 trainable parameters). In the two following layers, convolutions are performed with standard and shorter convolutional filters, 3 kernels of size 5 × 3 and 3 kernels of size 3 × 3 respectively, and 3 bias terms each (48 + 30 = 78 trainable parameters). The convolutional layers are followed by 3 fully connected layers with ReLu activations for the 2 first and softmax for the last one. The sizes of fully connected layers are 15 × 4, 4 × 4 and 4 × 3 respectively, with 4, 4 and 3 bias terms (64 + 20 + 15 = 99 trainable parameters). Total number of trainable parameters is 342. The model uses a cross-entropy loss and Adam optimiser. The learning rate was set to 0.001 and the training was stopped after 25 epochs to avoid overfitting.

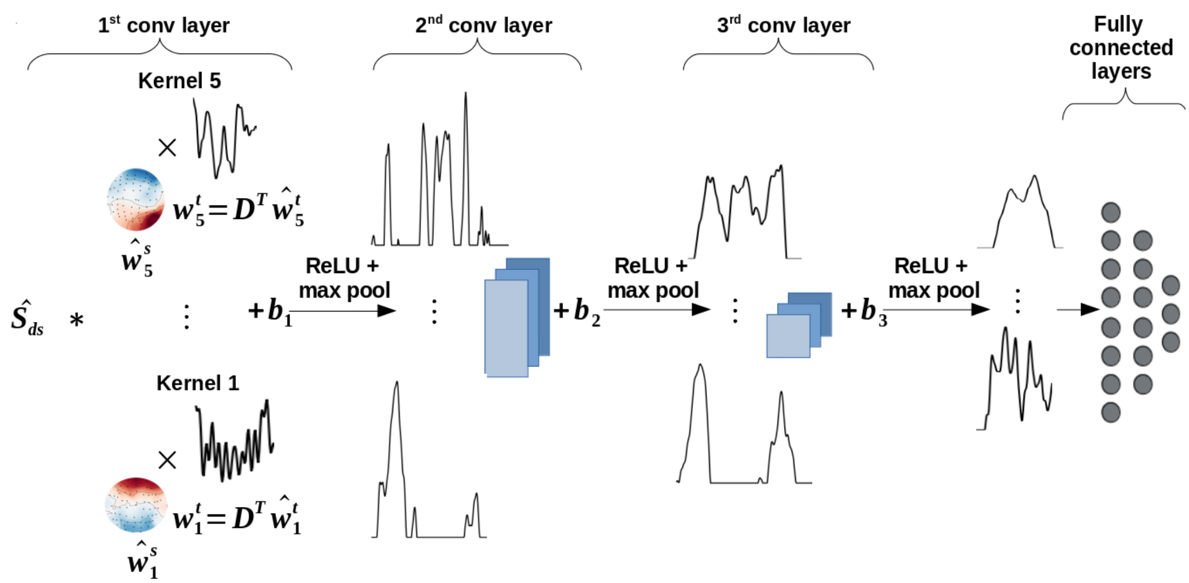

Figure 1. Rank 1 convolutional neural network architecture for passive BCI signal classification

For validation, the 2 labelled sessions for each of the 15 participants were assigned to the train or validation set randomly (the third session was unlabelled and reserved for test set). The model was then trained and validated based on this split for all the participants at once, in a generalised manner. This approach was chosen because the dataset only contained 2 different labelled sessions per participant, making it hard for the model to generalise with a personalised approach. The training graphs can be seen in Figure 2. The overall accuracy on the validation set averaged over 3 experiments with random train/validation split was 46.46 %. The performance is consistently higher than the chance level of 33.33 % as seen in Figure 1 with the confidence interval, however this remains far from being robust, highlighting the difficulty of classification of unseen sessions. Finally, the model was re-trained on the 2 labelled sessions in order to produce the final results consisting of a prediction for each epoch of the unlabelled test set of each participant. The proposed method lets room for improvement, with future work possibly focused on fine tuning the model for each participant individually after the generalised training.

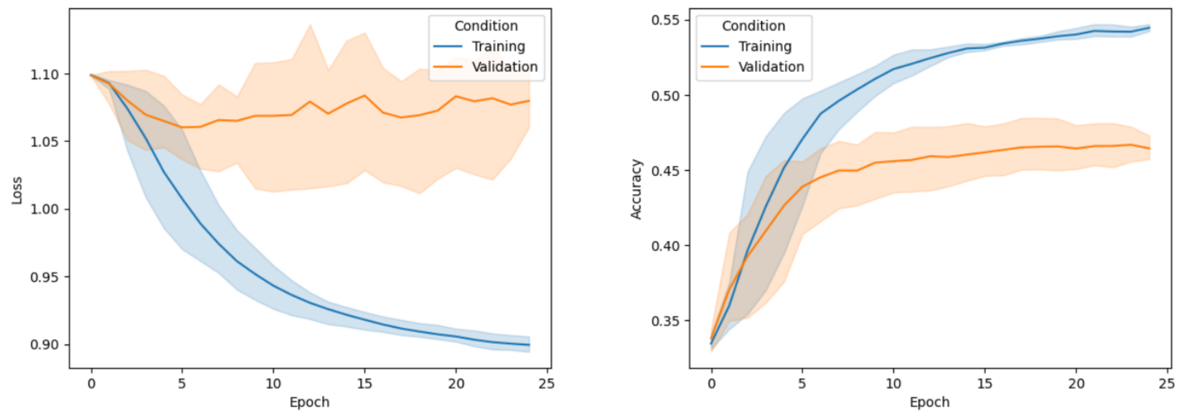

Figure 2. Training graphs of the CNN rank-1 model (loss on the left, accuracy on the right), using one labelled session for the train set and the other labelled session for the validation set (assigned randomly for each participant). The bands represent the 95 % confidence intervals on 3 repetitions of the procedure with random train/validation splits.

Acknowledgments:

This work was supported by the ERC under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665:CoBCoM : Computational Brain Connectivity Mapping). It has also been partly supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

References:

Clerc, M., Bougrain, L., & Lotte, F. (2016). Brain-Computer Interfaces 1. Wiley-ISTE.

Wang, S., Gwizdka, J., & Chaovalitwongse, W. A. (2015). Using wireless EEG signals to assess memory workload in the n-back task. IEEE Transactions on Human-Machine Systems, 46(3), 424-435.

Appriou, A., Cichocki, A., & Lotte, F. (2018, April). Towards robust neuroadaptive HCI: exploring modern machine learning methods to estimate mental workload from EEG signals. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-6).

Shalchy, M. A., Pergher, V., Pahor, A., Van Hulle, M. M., & Seitz, A. R. (2020). N-Back Related ERPs Depend on Stimulus Type, Task Structure, Pre-processing, and Lab Factors. Frontiers in human neuroscience, 14, 453.

Hinss, M. F., Darmet, L., Somon, B., Jahanpour, E., Lotte, F., Ladouce, S., & Roy, R. N. (2021). An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021 (Version 2) [Data set]. Neuroergonomics Conference, Munich, Germany. Zenodo.

Dupré La Tour, T., Moreau, T., Jas, M., & Gramfort, A. (2018). Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals. Advances in Neural Information Processing Systems (NIPS).

Kim, H., Yoon, J., Jeong, B., & Lee, S. (2018). Rank-1 convolutional neural network. arXiv preprint arXiv:1808.04303.