

Using Autoencoders to Denoise Cross-Session Non-Stationarity in EEG-Based Motor-Imagery Brain-Computer Interfaces

O. Shriki^{1*}, O. Almagor¹, O. Avin¹, R. Rosipal²

¹Ben-Gurion University of the Negev, Beer-Sheva, Israel; ²Slovak Academy of Sciences, Bratislava, Slovakia

mail: shrikio@bgu.ac.il

Introduction:

A major problem in motor-imagery (MI) brain-computer interfaces (BCIs) relates to the non-stationarity of brain signals. Consequently, the performance of a classifier trained for an individual subject on a certain day deteriorates during the following days. The traditional approach is to recalibrate the algorithm every session, limiting the wide use of BCIs.

Materials, Methods, and Results:

Here, we use an autoencoder (AE) convolutional neural network to identify a low dimensional representation of the EEG signals from the first day (or days) and show that this allows for stable decoding performance on the following days without resorting to recalibration. Furthermore, we demonstrate that the residual signals, namely the difference between the original and reconstructed EEG, can be used to accurately discriminate among different recording sessions. In line with that, the reconstructed EEG cannot be used to discriminate among recorded sessions. This implies that the reconstructed EEG reflects an invariant representation of the subject's intent, whereas the residual signals reflect a non-stationary component, which differs from one session to another. Results from an application to MI data from a stroke patient are depicted in Figure 1. As can be seen, the performance of our algorithm with no recalibration (green) is equivalent to that of a benchmark classifier which is recalibrated (blue). In contrast, training a Naïve classifier with no recalibration (red) achieves poorer performance.

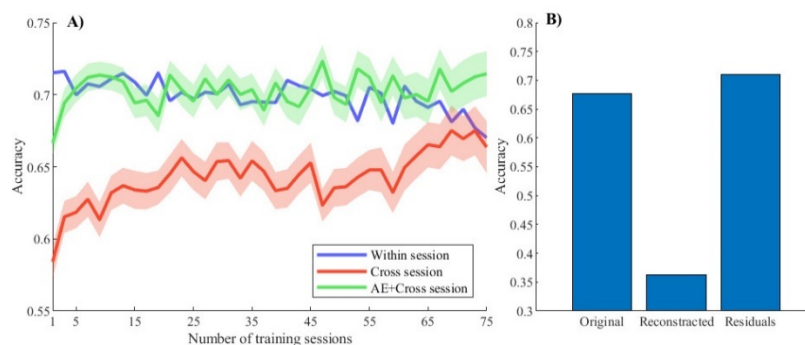


Figure 1: Analysis of longitudinal data (134 daily sessions) from a stroke patient. **A)** Mean accuracy of different models. The cross-session (red) and AE + cross-sessions (green) models were trained on an increasing size of the training set and were tested on the remaining sessions. The within-session benchmark (blue) is the score of the cross validation on the test set (no inference on new sessions). The shaded area represents the standard error of the accuracy across sessions. **B)** Mean accuracy of origin session classification using the original signals, reconstructed de-noised, and residuals signals.

Discussion:

Autoencoders can be used to identify a low-dimensional invariant representation of subject intent, which can minimize or eliminate recalibration. Analyzing the residual signals can provide further insights into the sources of signal non-stationarity. The approach can be generalized to incorporate weakly-supervised methods.

Significance:

The approach relies on an unsupervised learning framework which can be easily incorporated (without additional data collection) to increase the stability of BCIs and bring them closer to wide adoption.