Analyzing event-related potentials in 8-channel EEG data using machine learning methods

A short explanation of my Bachelor's Thesis (at University of Osnabrück; supervisors: Olivera Stojanovic M. Sc., Prof. Dr. Gordon Pipa; handed in on 26.09.2018)

Ann-Kathrin Schalkamp

May 28, 2020

1 Motivation

Electroencephalography is one of the most used techniques in non-invasive brain research (Melnik et al., 2017). Electrical components are getting more powerful, cheaper, and smaller at the same time (Appel, 2018) allowing for the construction of mobile EEG systems, which can address new research questions.
In BCIs, performing single-trial analysis on EEG data, meaning an online classification based on the data of a single ERP, allows for the translation of the brain's activity into commands for the external world. Challenges here are having to learn efficiently from small data sets, a poor signal-to-noise ratio, and mislabels.
Because of the rise of mobile EEG systems and advances in the field of BCIs it may be time to question some standards in EEG research established over years. Mobile EEG devices often use a significantly smaller number of electrodes, some do not use the standard placement of electrodes, and others use dry electrodes. Rethinking standards of EEG research might lead to new ideas how to tackle with challenges in BCI research.

1.1 Aim of the Study

The main objective is to investigate whether EEG data acquired with the *Traumschreiber*, a portable high-tech sleep mask developed by Johannes Leugering and Kristoffer Appel at the University of Osnabrück under the supervision of Prof. Dr. Gordon Pipa, is valid for traditional and single-trial analysis. The established standard for the reference placement in EEG research will be reviewed in the context of BCI applications. From these insights new ideas to tackle the challenge of small data sizes in BCI applications will be developed and tested.

2 Analysis

2.1 Data

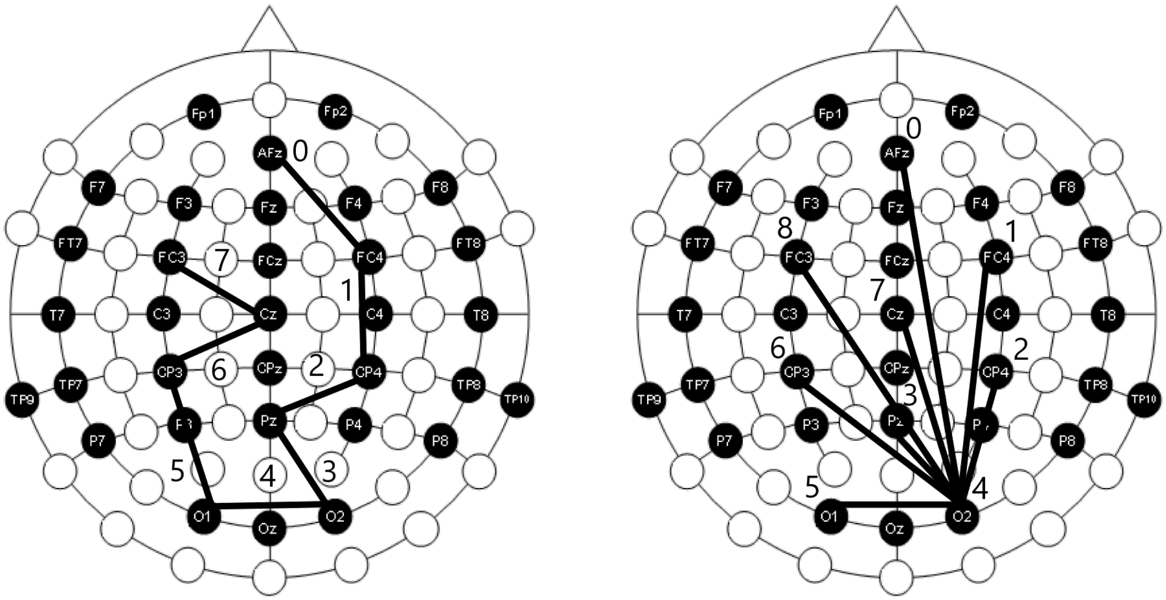
Performance on six paradigms (SSVEP, Motor Potential, N240, P300, vMMN, Auditory Potential), proposed as a benchmark for EEG systems (Melnik et al., 2017), is investigated. Data of eight participants (1 male, 7 females, mean age = 20.63 years, range: 19 - 21 years) was collected in cooperation with Merle Reimann and the data of three paradigms (SSVEP, Motor Potential, N240) is analyzed in this thesis. The *Traumschreiber* used for data acquisition differs from standard EEG systems as it was built to be maximally robust in case an electrode gets detached during sleep. It has a 220 Hz sampling rate and uses nine bipolar electrodes with one being the ground.

2.2 Analysis Pipeline

The data is first segmented in epochs, artifacts are removed (200microV threshold), baseline correction is performed, and the data is re-referenced (see figure 2.2). After that, the standard averaging method and logistic regression for single-trial analysis are performed.

The choice of the reference is a highly discussed topic. Common choices are averaged mastoids or earlobes due to their proximity to the brain while recording nearly no brain activity (Cohen, 2014, p. 82). Here, we choose an electrode above the originating region of the respective ERP as the reference through re-referencing, a linear transformation which does not affect the data (see figure 2.1). This setup allows for an evaluation of the effect of various reference placements, as only one recording electrode is used, the one above the originating area, and the other electrodes serve as references. This is accomplished by calculating the cumulative sum over all channels and subtracting the channel which should be the reference. As a result time-series data of nine channels is obtained, where one's activity is equal to zero since it displays the difference from the reference electrode to itself.

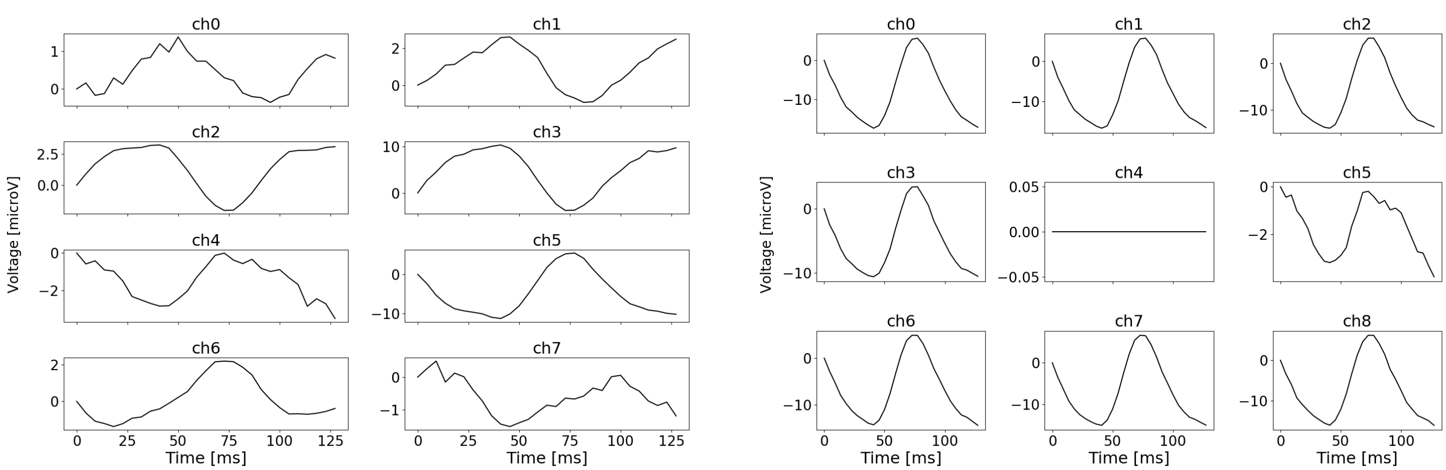
For the standard averaging method, the average over all epochs and subjects is computed for each channel. To check for the effect of the reference placement, the average over all channels is also computed. The single-trial analysis is performed with three different sets of input data for each paradigm. Firstly, a model is trained for each subject and channel individually. Secondly, data from the three best performing channels is concatenated in two ways: the data from the channels are assumed to add distinct information and the number of parameters is increased, or the data is assumed to show the same information and thus the sample size is increased. Thirdly, the data of all subjects is combined, and all previously described analyses are repeated with this increased data set. Therefore, the inter-individual heterogeneity is assumed to be low, allowing for ERPs being detected with the same model across subjects.



(a) Electrode Placement

(b) After re-referencing

Figure 2.1: EasyCap with used electrode placement and after re-referencing to electrode four. Adapted from (FieldTrip, 2018).



(a) Not re-referenced

(b) Re-referenced to electrode 4

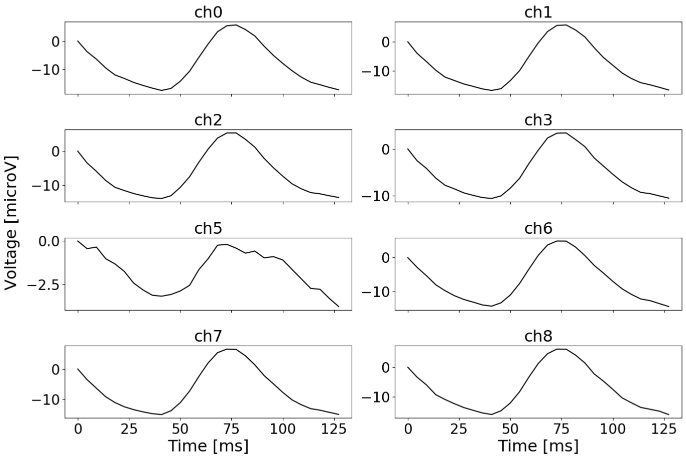
Figure 2.2: Effect of re-referencing

2.3 Results and Interpretation

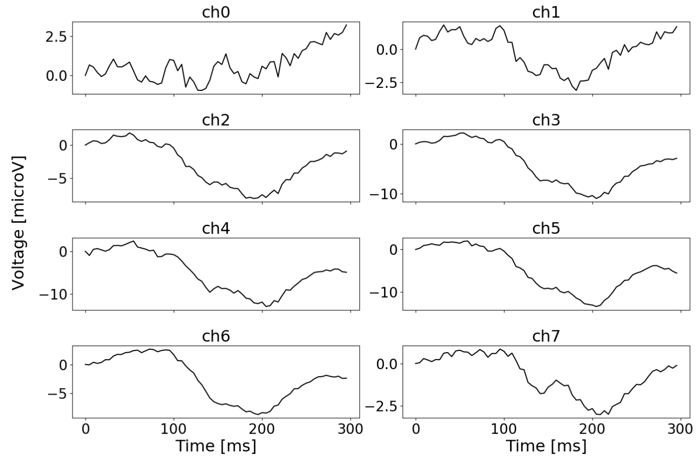
Overall, the results show that the *Traumschreiber* is capable of recording prominent ERPs. For the standard averaging method, the average of each channel and the average across all channels is computed. Because of the re-referencing, the ERP is visible in most of the channels suggesting that most reference placements work fine (see figure 2.3). In channels where the ERP is less clearly visible, the reference and the recording electrode are both placed near or above the originating area. Thus, the difference between these electrodes does not contain the signal itself. The average over all channels also depicts the ERP characteristics as described in the literature, suggesting that all channels recorded the same signal despite different reference placements (see figure 2.4).

The single-trial analysis shows average to good performance. See figure 2.5 for a comparison of the results between methods for each ERP. The placement of the reference does not influence the performance of the classifier significantly unless it is placed near or above the originating area as well. Placing the reference above the originating area provides the possibility to combine data from different channels, which all hold information about the ERP. Combining the data from the best three channels (in terms of AUC score) to increase the dimensionality reduced the AUC for all paradigms, except SSVEP significantly. For SSVEP more data points were available from each channel, enabling it to handle the increased dimensionality. The performance through this data concatenation was not increased, because all channels depict nearly the same activity and do not hold distinct information due to the applied re-referencing. However, combining the data of the best channels assuming they show the same information, increases the data size and led to AUC scores comparable to those of the single best channel. This proves, that the variations across channels was low and further shows that the reference placement is rather unimportant in BCI settings. Another way to increase the sample size is to combine data from subjects. This only led to high AUC scores if the task was low-level, like SSVEP. Cognitively more demanding tasks vary more across individuals, leading to a low AUC score if the data is assumed to hold the same information.

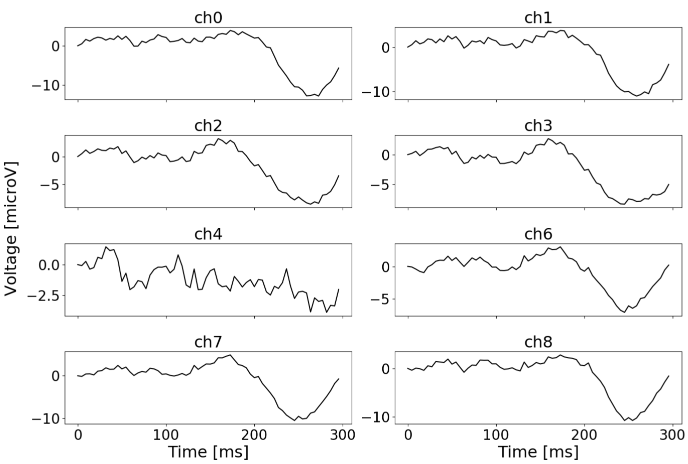
Overall, the results prove that the data acquired with the *Traumschreiber* allows for the standard averaging method and single-trial analysis. It does not deliver any information about the location of potentials as the spatial resolution with only eight channels is very poor. However, in BCI settings, the location of the ERP of interest is known as only well-studied ERPs are used as communication tools. The applied re-referencing method showed that the placement of the reference electrode is rather unimportant provided that the two electrodes do not record the same signal. Positioning the reference above the originating area allows for new approaches towards tackling the small sample sizes in BCIs. Combining data from many channels recording the same signal proves to be the most robust method to increase the data set size without increasing the number of trials.



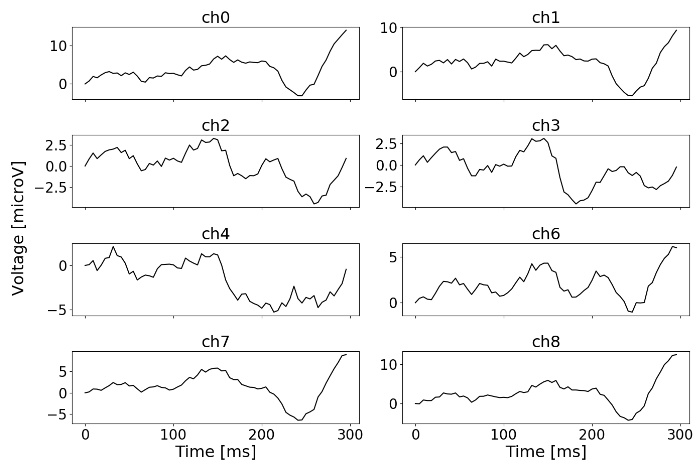
(a) SSVEP for each channel



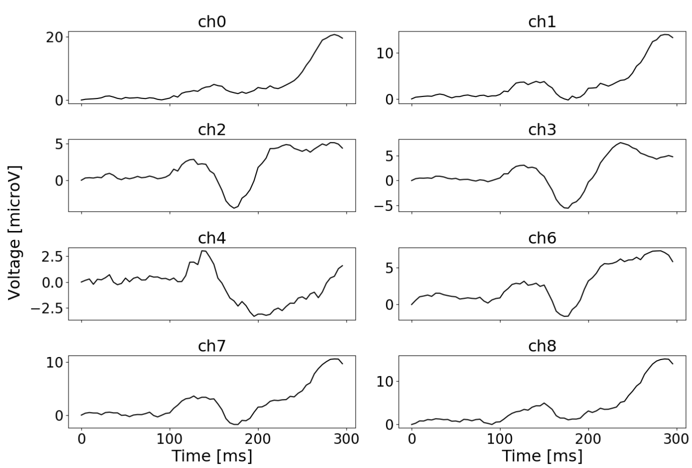
(b) MP for each channel



(c) N240 target minus distractor for each channel

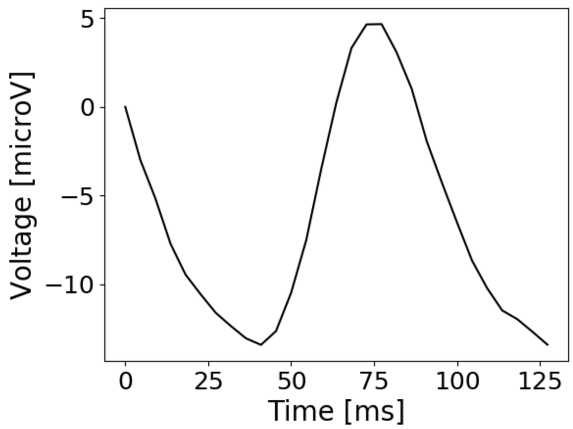


(d) N240 target for each channel

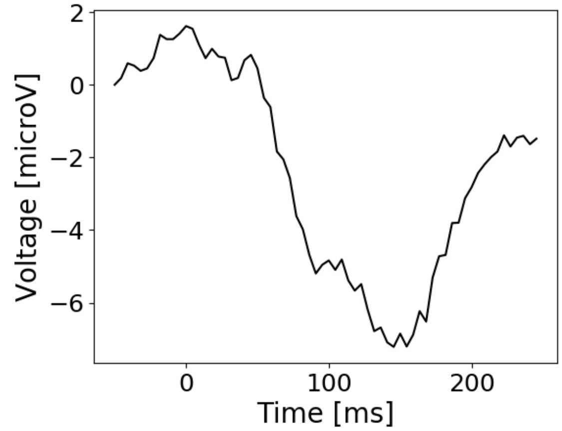


(e) N240 distractor for each channel

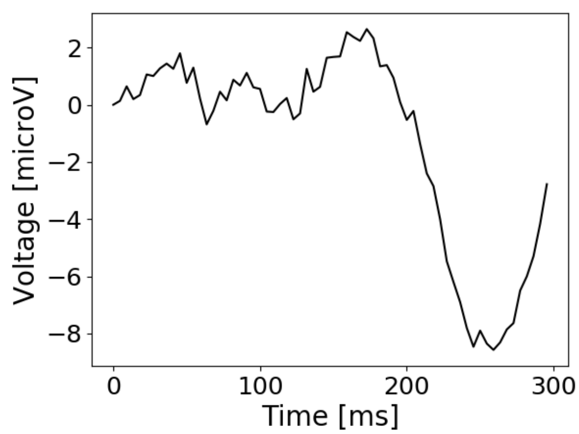
Figure 2.3: Standard Averaging Method for each channel



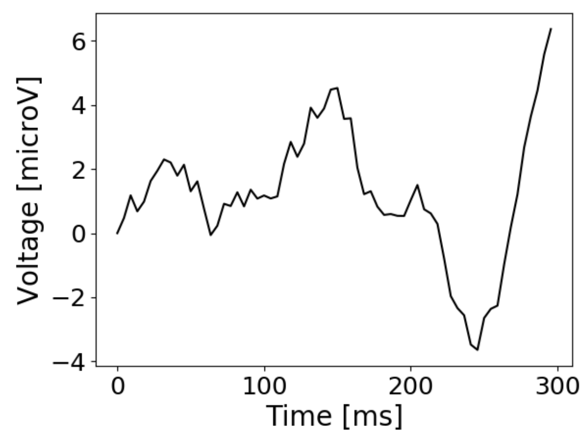
(a) SSVEP over channels



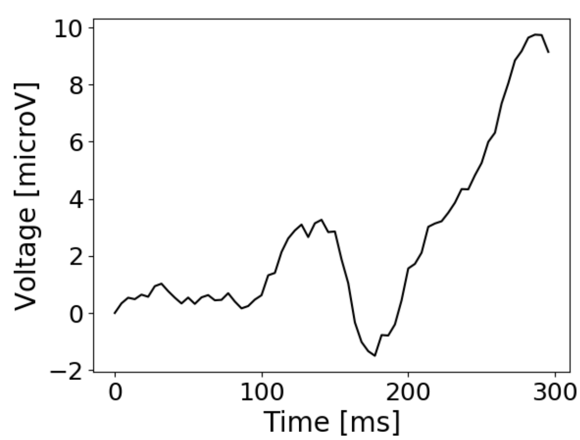
(b) MP over channels



(c) N240 target minus distractor over channels

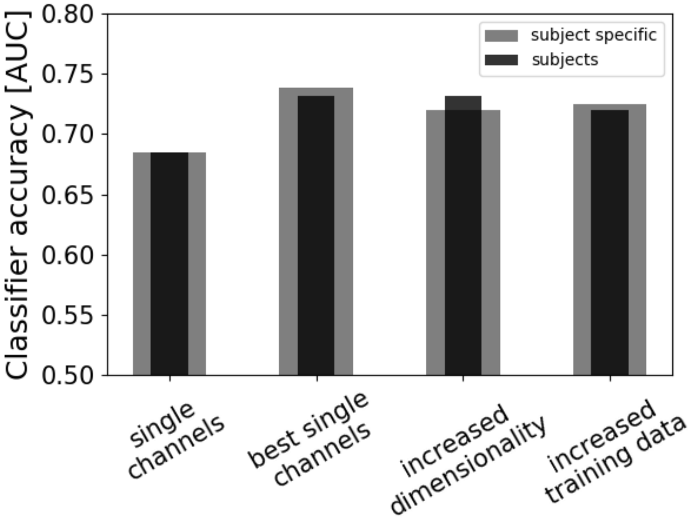


(d) N240 target over channels

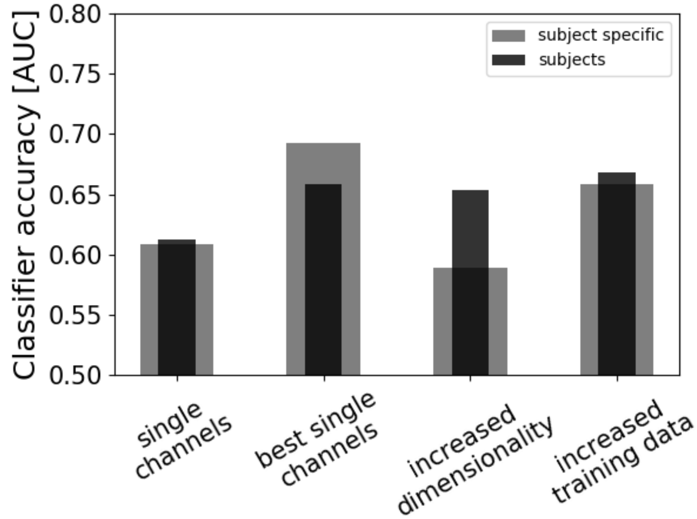


(e) N240 distractor over channels

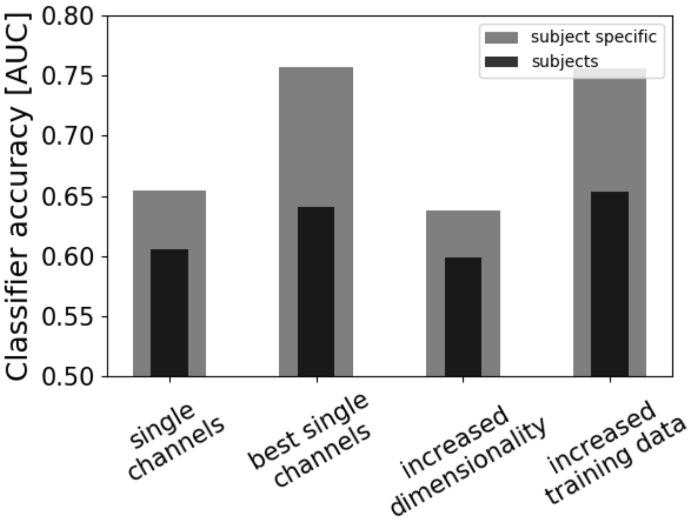
Figure 2.4: Standard Averaging Method across channels



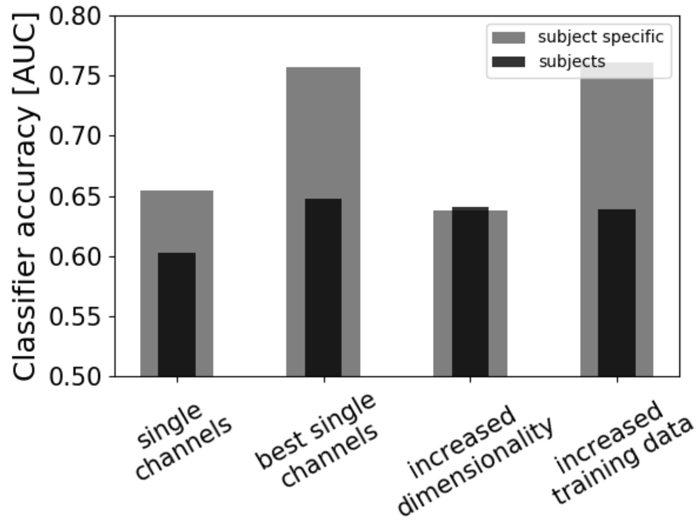
(a) SSVEP



(b) motor



(c) N240 target



(d) N240 distractor

Figure 2.5: Comparison of AUC scores

References

Appel, K. (2018). *The traumschreiber system: Enabling crowd-based, machine learning-driven, complex, polysomnographic sleep and dream experiments* (Unpublished doctoral dissertation). University of Osnabrück.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice.* MIT press.

FieldTrip. (2018). *easycapm3 - extended 10/20-system with 30 channels.* Retrieved from `http://www.fieldtriptoolbox.org/_media/template/easycapm3.png` ([accessed: 17-09-2018])

Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., & König, P. (2017). Systems, subjects, sessions: to what extent do these factors influence eeg data? *Frontiers in human neuroscience*, *11*, 150.