 DOI:10.3217/978-3-85125-260-6-144

On Classifying Artifactual Independent Components: Generalization Ability to Different Electrode Setups

I. Winkler[1], E. Waldburger[1], S. Haufe[1], M. Tangermann[1]
[1]*Department Machine Learning, Berlin Institute of Technolgy, Germany*

Correspondence: I. Winkler, Dept. Machine Learning, Marchstr. 23, 10587 Berlin, Germany. E-mail: i.winkler@tu-berlin.de

Abstract. BCI system development relies on data from healthy subjects, who unconsciously might utilize artifacts for BCI control. As these systems are typically developed for people with severe motor disabilities, a high sensitivity level of a BCI system for artifacts must be considered problematic. A robustness analysis of a state-of-the-art classification approach for the automatic rejection of artifactual independent components reveals, that this method robustly performs for a wide range of electrode setups, and that simple re-training ensures high rejection accuracy even for drastically reduced electrode numbers.

Keywords: Artifact Removal, EEG, Independent Component Analysis (ICA), Blind Source Separation (BSS), Machine Learning

1. Introduction

The analysis of EEG signals is often impeded by muscular or external artifacts, especially for EEG with small data set sizes. Consequently the reliability of single-trial analysis methods (as in BCI) and data visualizations for introspection may suffer. A common counter-measure is to decompose the original EEG into independent source components (ICs) and reconstruct it after dismissal of hand-selected artifactual ICs [Jung et al., 2000].

Avoiding this time-consuming process, recently proposed algorithms classify ICs into artifactual and non-artifactual components. Demonstrating good performance on similar validation data, the question arises how well these methods generalize to data acquired under novel experimental conditions. First studies suggest that generalization is possible (e.g. [Viola et al., 2009; Winkler et al., 2011]), but a detailed assessment of robustness is lacking. Here, we take a step forward by analyzing the generalization ability of an IC classification algorithm we recently proposed.

2. Material and Methods

2.1. Experimental setup, ICA unmixing and data split

The artifact classifier was set up using expert-labeled independent components gained from several conditions of a reaction time study [Winkler et al., 2011]. EEG data from 121 approx. equidistant sensors was available for eight healthy, right-handed male subjects. In total, 43 runs of 10 minutes duration were available, of which 28 from five subjects were used as training sets, and 15 runs from three subjects as test sets. After high-noise channels were rejected based on a variance criterion, they still had 104 electrodes in common. Prior to the IC computation via TDSEP [Ziehe et al., 2004], a 2 Hz high pass filter was applied, and a dimensionality reduction to 30 PCA components was performed in order to reduce artificial splits of sources. Two experts hand-labeled the 30 ICs per data set into artifactual and non-artifactual components, resulting in 840 training- and 450 test ICs.

2.2. The artifact classifier

The artifact classifier was a linear classifier based on six features that were selected in a feature selection procedure described in [Winkler et al., 2011]. The mean local skewness aims to detect outliers in the time series of an IC. Three features describe a $1/f$ fit of the IC to the spectrum and its log band power in the α band (8–13 Hz). Contrary to these first four features, the two remaining ones directly depend on the electrode setup, as they extract information about the scalp pattern of an IC: (1) Range Within Pattern characterizes the difference between the minimal and maximal activation in a pattern. (2) Current Density Norm is derived from the source localization of an IC, which is based on its pattern. We considered 2142 locations arranged in a 1 cm grid and computed the source distribution with minimal l_2-norm. This norm was used as a feature. The underlying idea is that noisy patterns and patterns originating outside the brain represent more complicated sources, which are characterized by larger l_2-norms.

 DOI:10.3217/978-3-85125-260-6-144

2.3. Analyzing the classification performance for different electrode setups

Two different classification strategies were compared: One *fixed* IC-classifier was pre-trained on features of the full 104-channel data. Its performance was estimated on test data of setups varying from 16 to 104 channels (all approx. equidistant and covering the whole scalp). Alternatively, a *re-training* of the IC-classifier was performed for each montage, on features computed on training patterns cut to the specific montage. The performance of the re-trained classifiers again was tested on the full and the reduced setups.

3. Results

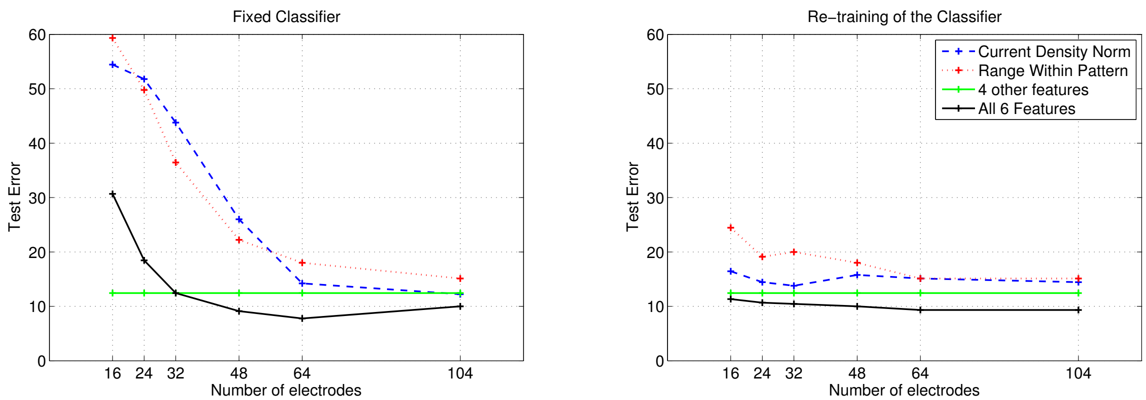

Figure 1: *Classification error estimated on the test sets for different channel setups for a fixed classifier (left plot) and a classifier re-trained for each channel setup (right plot). The chance level performance is at 50 %.*

For the 104 channel setup, a classifier using the full six features achieves a low error rate of 10.0 % only, which outperforms the use of only four pattern-independent features (12.4 %). The *fixed* classifier generalizes robustly over a large range of 104 to 48 electrodes in the test sets. The increased error of up to 30.6 % for the smallest set of 16 electrodes is associated with the bad performance of both single features which are based on the pattern (over 50 %).

For the re-training strategy, the error increase of the single *Range Within Pattern* feature was less pronounced (from 15.1 % to 24.4 %), and the *Current Density Norm* feature even remained relatively stable. Using the re-trained classifier, the overall error for 16 electrodes remained at 11.33 %, which is comparable to inter-expert disagreements. For this reduced setup, the classifier weight of the *Range in Pattern* dropped, while the weight for *Current Density Norm* remained stable.

4. Discussion

We have analyzed the generalization ability of an IC classification algorithm we recently proposed to different electrode setups. For this analysis, two human experts judged the components based on patterns showing 104 electrodes, while we restricted the electrodes that the classifier saw. We showed that classification was relatively robust to a decrease from 104 to 48 electrodes - roughly half the number of electrodes - from training to testing. However, performance dropped after more electrodes were removed. We demonstrated that recomputing the features and retraining the classifier based on the specific electrode montage of the test set alleviates the problem.

Acknowledgments

This work is supported by the European ICT Programme (Project FP7-224631 *TOBI*), by the German Federal Ministry for Education and Research (BMBF) (Grant 01GQ0850) and by the Federal State of Berlin.

References

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiol*, 37:163–178.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., and Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *Clin Neurophysiol*, 120:868–877.

Winkler, I., Haufe, S., and Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behav Brain Funct*, 7:30.

Ziehe, A., Laskov, P., Nolte, G., and Müller, K.-R. (2004). A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *J Mach Learn Res*, 5:801–818.