Semi-Supervised Neural Gas for Adaptive Brain-Computer Interfaces

Hannes Riechmann and Andrea Finke *

Bielefeld University - CITEC
Universitaetsstrasse 21, 33615 Bielefeld - Germany

Abstract. Non-stationarity is inherent in EEG data. We propose a concept for an adaptive brain computer interface (BCI) that adapts a classifier to the changes in EEG data. It combines labeled and unlabeled data acquired during normal operation of the system. The classifier is based on Fuzzy Neural Gas (FNG), a prototype-based classifier. Based on four data sets we show that retraining the classifier significantly increases classification accuracy. Our approach smoothly adapts to the session-to-session variations in the data.

1 Introduction

The non-stationarity of EEG data is one major issue in current BCI research and a severe impediment on the way to everyday usage of BCIs [1, 2]. There are two major sources of these instationarities: 1) Small changes between the experimental setup during training and evaluation sessions, e.g., in electrode placement and impedances. 2) Changes in neuronal activity, caused by alertness, fatigue, etc. These changes degrade the classification accuracy over time. Typically, this issue is addressed by retraining the system with newly acquired data - a time-consuming procedure that is not applicable for everyday usage of a BCI system.

Over the years, several adaptive BCI systems have been proposed which cope with non-stationarities in the data. There are two major types of adaptive BCI systems: One type focuses on efficiently retraining the classification system during the on-line run using labeled data. The drawback is that the true labels are not available in most scenarios [1]. Especially, in out-of-lab scenarios, where the user wants to control some device by means of a BCI, it is normally not possible to acquire true labels for the data. Other adaptive BCIs try to tune the classification system to the non-stationarities without using true labels. One approach is to use only the uncertain decisions of the current classifier as labels to train a new classifier. This is commonly used in P300-based BCIs, where the system output is computed as the average over several single-trial classifications. The labels for the single trials can then be reconstructed from the averaged result, leading to a higher overall accuracy. Still, the labels are prone to classification errors and can be misleading for the training of the new classifier. This averaging scheme, however, is not applicable in motor-imagery (MI) based BCIs, rendering on-line adaption more challenging in these systems. Purely unsupervised

*The study was partially funded by the German Research Council (DFG), EXC 277. This work was supported by the EU project MONARCA in the FP7 program.

classifiers have been proposed to cope with this issue, mostly based on Gaussian mixture models, e.g., [3, 4]. These classifiers assume a Gaussian class distribution and little class overlap, assumptions that are only true to some extent. Especially the class overlap is quite high for some data sets [4].

Here, we propose a novel, adaptive classification system for the detection of MI, based on semi-supervised learning and fuzzy neural gas (FNG). Using semi-supervised learning, labeled data acquired during one or more training sessions is combined with unlabeled data acquired during the on-line operation of the system. This combines the advantages of supervised and unsupervised learning. The labeled data from training allows for an initial modeling of the complex class distributions of the EEG data. Adding the unlabeled data acquired during normal operation, we can then adapt the distributions to the on-line variations during system operation. We use a prototype-based algorithm to model the data distribution. This allows us to successively increase the model complexity by adding more prototypes.

2 Semi-supervised Neural Gas

We extend the algorithm proposed by Villmann et al. to build an adaptive BCI system [5, 6]. The authors did a thorough theoretical analysis and applied it to some benchmark problems, but not to a real-world classification problem. In the following, we will describe the algorithm and our modifications for the detection of MI in EEG data.

2.1 Semi-supervised learning

Semi-supervised learning (SSL) combines labeled and unlabeled data to train a classifier [7]. While labeled data is more valuable to calculate the classification boundaries, it is rather time-consuming to acquire in a BCI system.

2.2 Fuzzy Neural Gas

Since the introduction of self-organizing maps (SOM) and learning vector quantization (LVQ), prototype based models have been thoroughly investigated [8]. The general idea is to model the data distribution by a set of prototypes. This can be used for compression, but also for classification, where each prototype is assigned a class membership. In contrast to many other classification approaches, it is possible to directly influence the complexity of the classification boundary by setting the number of prototypes. In Neural Gas (NG), in contrast to LVQ, all prototypes are updated according to a neighborhood function at the presentation of a new sample. This way, a gradient descent on the error function, which measures the distances between each sample and the nearest prototype, robustly converges [9]. When using Neural Gas for classification, each prototype is given a label and, after training, new samples are assigned the label of their prototype. Fuzzy-labeled NG [5] and semi-supervised NG [6] were introduced to cope with fuzzy or unknown labels in the training data. In both approaches, the class membership of a prototype is an n-dimensional label $\mathbf{c}$ with $\mathbf{c}_j \in [0,1], j = 1, ..., N$, where N is the number of classes. For each class, the corresponding entry $\mathbf{c}_j$ indicates the confidence that the prototype belongs to this class.

The error function from the original NG is adapted to take the label distances into account:

$$\begin{aligned} E = & \sum_j \int P(\mathbf{v}) h(k_j(\mathbf{v}, \mathbf{w}_j)) D(\mathbf{v}, \mathbf{w}_j, \gamma) d\mathbf{v} \\ D(\mathbf{v}, \mathbf{w}_j, \gamma) = & (\gamma \cdot \delta(\mathbf{c}_\mathbf{v}, \mathbf{y}_j) + \epsilon_\delta) \cdot ((1-\gamma) \cdot (d(\mathbf{v}, \mathbf{w}_j) + \epsilon_d)) - \epsilon_\delta \cdot \epsilon_d \end{aligned}$$

$\mathbf{v}$ are the data samples with their distribution $P(\mathrm{v})$, $\mathbf{w}_j$ the prototypes, $\mathbf{y}_j$, $\mathbf{c}_\mathbf{v}$ the labels for the prototypes and samples. $h(k_j(\mathbf{v}, \mathbf{w}_j)) = exp(-\frac{k_j(\mathbf{v}, \mathbf{w}_j)}{2\sigma^2})$ is the rank function, which uses the neighborhood function $k_j(\mathbf{v}, \mathbf{w}_j)$ based on a distance measure, usually the euclidean metric. Differentiating the error function leads to the following update equations for the prototype positions and labels:

$$\triangle \mathbf{w}_j = -(1-\gamma)(\gamma \cdot \delta(\mathbf{c}_\mathbf{v}, \mathbf{y}_j) + \epsilon_\delta) \cdot h(k_j(\mathbf{v}, \mathbf{w}_j)) \cdot \frac{\partial d(\mathbf{v}, \mathbf{w}_j)}{\partial \mathbf{w}_j}$$

$$\triangle \mathbf{y}_j = -\gamma((1-\gamma) \cdot d(\mathbf{v}, \mathbf{w}_j) + \epsilon_d) \cdot h(k_j(\mathbf{v}, \mathbf{w}_j)) \cdot \frac{\partial \delta(\mathbf{c}_\mathbf{v}, \mathbf{y}_j)}{\partial \mathbf{y}_j}$$

Training is stopped when the error function converges or a maximal number of steps is reached.

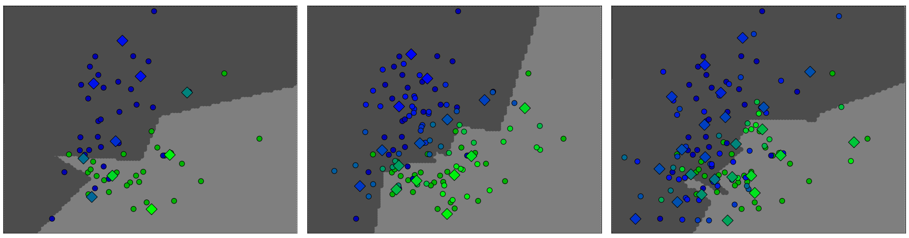

Fig. 1: Fuzzy Neural Gas. The prototypes are depicted by diamonds, the samples by circles, color indicates class. The background shows the classification boundarys.

Figure 1 shows an example using real EEG data. The left figure depicts the situation after the first training, the input data samples have a known, discrete training label. Two clusters are visible, but show a strong overlap. When more and more training samples (with fuzzy labels assigned by the classifier) arrive, the class distributions slightly move and become more complex. For example, the region at the lower left is mostly occupied by the green data class after the first training and is later occupied by the blue class. Meanwhile, the green class shifts to the right. The right picture also shows that the cluster assumption, which states that class boundaries run through regions with low data density, is violated for this data set. Additionally, the class boundary shows a cut on the lower left. Over-fitting might be an explanation, but the cut exists consistently throughout the re-trainings.

2.3 Adaptation for BCIs

To use Fuzzy Neural Gas for classification of MI in EEG data, some adaptations were necessary. First, in Neural Gas and its variants classification is normally done by a winner-takes-all strategy. A new sample is simply assigned the label of its closest prototype. Here, we take into account several prototypes and calculate a weighted average over their labels to obtain the label for the new data point. First, the score is computed:

$$Score_v = \sum_{j \in N(v)} (|N(v)| - k_j(\mathbf{v}, \mathbf{w}_j)) \cdot \mathbf{y}_j$$

$\mathbf{v}$ is the new sample, $N(v)$ gives the set of indices of the prototypes in the neighborhood of $\mathbf{v}$ and $\mathbf{y}_j$ is the label vector of prototype j. The score is a vector with values in $[0, 1]$. Each value codes the degree to which the new sample belongs to that class. The final label is given by $argmax(Score_v)$.

Second, we dynamically adapt the number of prototypes. As more and more unlabeled data arrive, the class distributions become more complex due to the non-stationarity of the data. We gradually add more prototypes to compensate for this effect. If the system runs over a long time, there will also be the need to discard some training data and prototypes, but we do not consider this here.

Third, like most prototype-based classification approaches, FNG is quite affected by the prototype initialization. We initialize the prototype positions and labels with some of the data positions and labels. Every nth sample is drawn from the sequentially sorted training set. This procedure ensures a good representation of the whole data space.

3 Methods and Experimental Setup

The data used to evaluate our method was acquired using a 16 channel gUSBamp (Guger Technologies) amplifier sampling at 256 Hz. Data was bandpass filtered with subject-specific cut-off frequencies, typically between 9 and 27 Hz. We applied Common Spatial Patterns (CSP) [10, 11] as spatial filter. We computed the log-variance for every output channel of the CSP. The resulting feature vectors were fed into the classifier.

The data was acquired in repeated sessions on different days. During the training sessions, the subjects were instructed to imagine movements of the left and right hand. On-line feedback was given.

For the off-line analysis of the adaptive BCI system the data was split into five parts in sequential acquisition order. This scheme simulates the situation of an on-line run for off-line analysis. The first 20 percent of the data is used for an initial classifier training using the correct known labels. The other data are gradually added to the training set, their fuzzy labels are given by the classifier from the last (re)training, as shown in Figure 2.

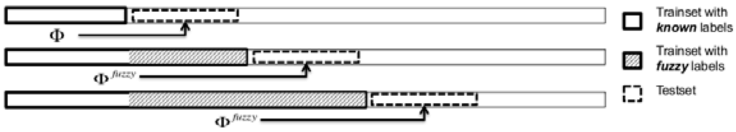

Fig. 2: Off-line simulation of an on-line run.

4 Results

Figure 3 shows the classification accuracies for the different conditions. To show the advantages of the semi-supervised FNG approach, we computed the classification accuracies for four conditions: 1) static FNG (sFNG) where the first 20 percent are used to train an FNG classifier and no retraining is applied. 2) static FDA (sFDA), same as condition 1 but using standard FDA. 3) adaptive FNG (aFNG), where the retraining procedure is applied as described above, 4) like 3, but using FDA (aFDA). The latter condition is a self-training FDA similar to the pseudo-supervised approach in [3]. The error-prone labels from the previous classifier are used to train a new classifier which can be beneficial, but might, for some data sets, lower the accuracy.

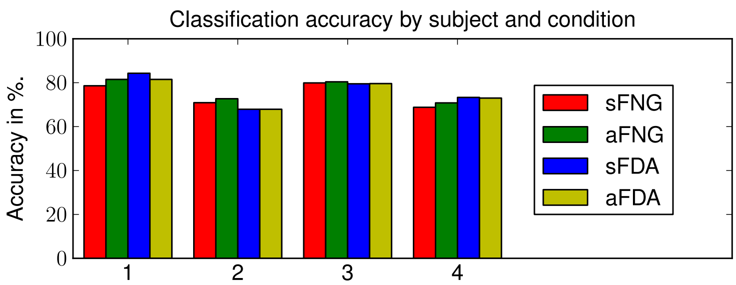

Fig. 3: Classification accuracy per subject for the 4 experimental conditions.

Adaptively retraining the classifier (condition aFNG) does obviously never hamper the classification accuracy. In total, a significant increase is reached, comparing sFNG (mean accuracy: 74.4%) with aFNG (mean accuracy: 76.4%), (independent t-test: $t \approx 3.6$, $p < 0.02$). A comparison of the two FDA conditions shows that retraining lowers classification accuracy for data sets 1 and 4. On average, retraining slightly decreases accuracy (mean 76.3% vs. 75.5%), although this difference is not significant ($p > 0.17$).

5 Discussion & Conclusion

We have presented a novel approach to deal with the non-stationarity of EEG data for motor-imagery based BCI. We use a Fuzzy Neural Gas classifier which is iteratively updated using only unlabeled data which can be gathered during normal operation of a BCI system.

To show the effectiveness, we tested our approach on four data sets. For two of the four data sets the adaptive FNG performs best. For the other two data sets classification accuracy increases as well compared to the static FNG, but does not reach the accuracy of the static FDA. A possible explanation might be that those two data sets are quite short and thus the non-stationarity of the EEG data is not as strong as for the other two data sets. While the expressiveness of this study is limited due to the small number of data sets, our results show that the approach is promising and worth further investigation.

In comparison with Gaussian approaches for un- or semi-supervised learning, the FNG seems to be more robust. While the adaptive FNG always outperforms the static FNG, the adaptive FDA actually lowers accuracy for some data sets, a known problem of many un- or semi-supervised approaches [4].

In contrast to purely discriminative approaches like FDA, FNG not only provides class labels, but also a confidence that a new data sample belongs to a certain class. We do not use this information here. In future work, we will explore whether this information allows to detect outliers.

References

[1] P. Shenoy, M. Krauledat, B. Blankertz, R. P.N Rao, and K. R Müller. Towards adaptive classification for BCI. *Journal of Neural Engineering*, 3:R13, 2006.

[2] D.J. Krusienski, M. Grosse-Wentrup, F. Galán, D. Coyle, K.J. Miller, E. Forney, and C.W. Anderson. Critical issues in state-of-the-art brain-computer interface signal processing. *Journal of Neural Engineering*, 8:025002, 2011.

[3] J.Q. Gan. Self-adapting bci based on unsupervised learning. In *3rd International Workshop on Brain-Computer Interfaces*, pages 50–51, 2006.

[4] SE Eren, M. Grosse-Wentrup, and M. Buss. Unsupervised classification for non-invasive Brain-Computer-Interfaces. In *Proc. of Automed Workshop (Dusseldorf, Germany: VDI*, pages 65–6, 2007.

[5] T. Villmann, B. Hammer, FM Schleif, and T. Geweniger. Fuzzy labeled neural gas for fuzzy classification. In *Workshop on Self-Organizing Maps*, pages 283–290, 2005.

[6] M. Kästner and T. Villmann. Fuzzy supervised neural gas for semi-supervised vector quantization - theoretical aspects. Technical report, Computational Intelligence Group, University of Applied Sciences Mittweida, 2011.

[7] K. Nigam, A. K McCallum, S. Thrun, and T. Mitchell. Text classification from labeled and unlabeled documents using EM. *Machine learning*, 39(2):103–134, 2000.

[8] T. Kohonen. *Self-Organizing Maps.* Springer, 3 edition, 2001.

[9] T. M. Martinetz, S. G. Berkovich, and K. J. Schulten. 'Neural-gas' network for vector quantization and its application to time-series prediction. *Neural Networks, IEEE Transactions on*, 4(4):558–569, July 1993.

[10] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clinical neurophysiology*, 110(5):787–798, 1999.

[11] Yijun Wang, Shangkai Gao, and Xiaorong Gao. Common spatial pattern method for channel selection in motor imagery based brain-computer interface. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 5392 –5395, jan. 2005.