Journal of Neuroscience Methods xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines

Tarek Lajnef[a], Sahbi Chaibi[a], Perrine Ruby[b], Pierre-Emmanuel Aguera[b], Jean-Baptiste Eichenlaub[c], Mounir Samet[a], Abdennaceur Kachouri[a,d], Karim Jerbi[b,e,*]

[a] Sfax National Engineering School (ENIS), LETI Lab, University of Sfax, Sfax, Tunisia
[b] DYCOG Lab, Lyon Neuroscience Research Center, INSERM U1028, UMR 5292, University Lyon I, Lyon, France
[c] Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA
[d] Higher Institute of Industrial Systems of Gabes (ISSIG), University of Gabes, Gabes, Tunisia
[e] Psychology Department, University of Montreal, QC, Canada

ARTICLE INFO

Article history:
Received 20 June 2014
Received in revised form 15 January 2015
Accepted 16 January 2015
Available online xxx

Keywords:
Electroencephalography (EEG)
Sleep scoring
Oscillations
Polysomnography
Decision-tree
Support vector machine (SVM)
Linear Discriminant Analysis (LDA)
Hierarchical clustering
Machine learning
Dendrogram

ABSTRACT

Background: Sleep staging is a critical step in a range of electrophysiological signal processing pipelines used in clinical routine as well as in sleep research. Although the results currently achievable with automatic sleep staging methods are promising, there is need for improvement, especially given the time-consuming and tedious nature of visual sleep scoring.
New method: Here we propose a sleep staging framework that consists of a multi-class support vector machine (SVM) classification based on a decision tree approach. The performance of the method was evaluated using polysomnographic data from 15 subjects (electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) recordings). The decision tree, or dendrogram, was obtained using a hierarchical clustering technique and a wide range of time and frequency-domain features were extracted. Feature selection was carried out using forward sequential selection and classification was evaluated using *k*-fold cross-validation.
Results: The dendrogram-based SVM (DSVM) achieved mean specificity, sensitivity and overall accuracy of 0.92, 0.74 and 0.88 respectively, compared to expert visual scoring. Restricting DSVM classification to data where both experts' scoring was consistent (76.73% of the data) led to a mean specificity, sensitivity and overall accuracy of 0.94, 0.82 and 0.92 respectively.
Comparison with existing methods: The DSVM framework outperforms classification with more standard multi-class "one-against-all" SVM and linear-discriminant analysis.
Conclusion: The promising results of the proposed methodology suggest that it may be a valuable alternative to existing automatic methods and that it could accelerate visual scoring by providing a robust starting hypnogram that can be further fine-tuned by expert inspection.

1. Introduction

Sleep is characterized by continuous changes in brain, eye, muscle, respiratory and heart beat activity. These changes are monitored with polysomnographic recordings, which measure, during a full night of sleep, different types of physiological data typically including the electroencephalogram (EEG), electro-oculogram (EOG), electromyogram (EMG) and electrocardiogram (ECG). Physiologically speaking, sleep states are split into two broad types: rapid eye movement (REM sleep) and non-rapid eye movement (non-REM sleep). The latter consists of 4 stages (S1, S2, S3 and S4). These distinct sleep stages are associated with distinct physiological and neuronal features which are generally used to identify the sleep stage a person is in. This process called sleep scoring, or sleep staging, is a critical step in a range of electrophysiological signal processing pipelines used in clinical routine as well as in sleep research.

In clinical routine, sleep studies are usually performed for the diagnosis of pathologies, such as insomnia, hypersomnia, circadian rhythm disorders, epilepsy and sleep apnea. Sleep scoring often relies on visual analysis of the recordings to establish a hypnogram that depicts in time the different sleep stages. The analysis generally follows established guidelines for sleep stage classification, such as

* Corresponding author at: Psychology Department, University of Montreal, QC, Canada. Tel.: +1 438 969 2103.
E-mail address: karim.jerbi@umontreal.ca (K. Jerbi).

http://dx.doi.org/10.1016/j.jneumeth.2015.01.022
0165-0270/

the ones introduced by Rechtschaffen and Kales (1968), where each segment of 30 s is labelled as Awake, S1–S4 or REM. A more recent classification manual proposed by the American Academy of Sleep Medicine (AASM) in 2007 (Iber et al., 2007), combines the non-REM stages S3 and S4 into a single stage of deep sleep (called N3), also known as slow-wave sleep (SWS). Both manuals propose to use EEG derivations (2 in the R&K manual, and 3 in the AASM one), 2 EOG electrodes and one EMG electrode.

While visual scoring remains the gold-standard, recent years have witnessed a surge in method developments for automatic or semi-automatic sleep staging (e.g. Agarwal and Gotman, 2001, 2002; Becq et al., 2005; Berthomier et al., 2007; Ma et al., 2011; Itil et al., 1969; Koley and Dey, 2012; Krakovska and Kristina, 2011; Larsen and Walter, 1970; Schaltenbrand et al., 1996; Sheng-Fu et al., 2012; Shing-Tai et al., 2012; Stanus et al., 1987; Steinn et al., 2005; Huang et al., 2014). Although these results obtained so far are promising, there is still room and a need for improvement, especially given the time-consuming and tedious nature of visual sleep scoring. Across existing methods, a wide range of physiological signatures, or features, have been extracted from polysomnographic (PSG) signals, including time-domain, frequency-domain, and time–frequency-domain features, and both linear and non-linear features have been explored. While some studies rely only on one or two features to perform sleep stage classification (e.g. Fraiwan et al., 2010; Šušmáková and Krakovská, 2008), several studies provide evidence for the utility of searching for an optimum combination of features (e.g. Grozinger et al., 2001; Sheng-Fu et al., 2012; Khalighi et al., 2013).

Beyond the specific electrophysiological features used, existing methods also differ in the type of classification framework used. Some machine learning techniques such as artificial neural networks have been widely used for sleep staging (Kerkeni et al., 2012; Marina et al., 2012; Ronzhina et al., 2012; Ma et al., 2011). A disadvantage of this method is the fact that the exact decision procedure remains hidden or implicit. Classification methods based on Bayesian probability (linear and quadratic discrimination, *k*-nearest neighbour), have also been used in sleep scoring (Fraiwan et al., 2010; Krakovska and Kristina, 2011) The requirement of a Gaussian distribution of data in these methods can sometimes be a limitation. Other approaches for automatic sleep scoring based on mathematical modeling and hidden Markov Models have also been proposed (Doroshenkov et al., 2007; Shing-Tai et al., 2012). Support vector machines (SVM) classification has also been used for sleep scoring (Steinn et al., 2005; Koley and Dey, 2012). Support vector machines, introduced in the early 90s (Boser et al., 1992; Cortes and Vapnik, 1995) are used in a wide range of learning problems such as pattern recognition, text categorization and medical diagnosis and they continue to draw a lot of attention in many fields including basic and clinical neuroscience.

In this paper, we propose a sleep staging procedure that achieves multi-class classification by embedding multiple SVMs in a decision tree (dendrogram) framework. The dendrogram-SVM method is applied to polysomnography data from 15 individuals and its performance is compared to expert visual scoring. In addition, to assess the added value of the proposed methodology, its results were benchmarked against two standard classification methods, which are "one-against-all" SVM and linear-discriminant analysis (LDA). In brief, the proposed classification pipeline, which is described in more detail below, consists of three main steps: (i) feature extraction from all EEG, EOG and EMG data of 15 individuals (covering both time and frequency domain features, combining linear and non-linear measures), (ii) dimension reduction and feature subset selection using forward selection and cross-validation within the training step and (iii) classification using a multi-class SVM based on a decision tree obtained via ascendant hierarchical clustering (AHC).

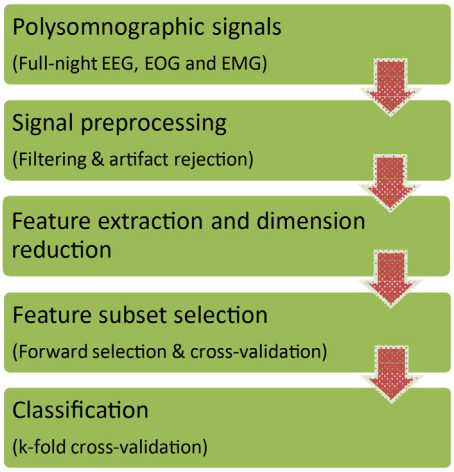

Fig. 1. Overview of the sleep stage classification pipeline.

In the following, we describe the full pipeline and the individual steps in detail. We then present the performance of the proposed DSVM method as measured by sensitivity, specificity and accuracy (using expert scoring as ground truth). Next, we report the comparison between the DSVM framework and standard "one-against-all" SVM and LDA applied to the same polysomnographic data set. Finally, we discuss the components of our method that could explain its higher performance compared to standard methods and also address potential limitations and ideas for further improvements.

2. Materials and methods

2.1. Polysomnographic data base

The data used in this study consists of polysomnographic (PSG) records in 15 healthy subjects aged 29.2 ± 8 years, which were collected at the DyCog Lab of the Lyon Neuroscience Research Center (Lyon, France) as part of a larger study exploring cognition during sleep (Eichenlaub et al., 2012, 2014; Ruby et al., 2013a,b). Each record contains EOG, EMG and 21 scalp-EEG channels. The EEG electrodes were positioned according to the International 10–20 system, the EOG electrodes were placed diagonally on the outer edges of the eyes, the EMG electrodes were positioned on the chin. All signals were recorded with a sampling frequency of 1000 Hz. The 15 PSG sleep recordings were visually scored by an expert in successive windows of 30-s using the R&K guidelines.

2.2. Methods

The sleep stage classification process can be divided in 5 distinct steps (Fig. 1). Once the polysomnographic signals are acquired, the EEG, EMG and EOG signals were filtered and all segments contaminated by artefacts were excluded (through a combination of automatic thresholds and visual inspection). This pre-processing step is comparable to standard EEG pre-processing and was fast (only taking a few hours to complete for all data sets). Different time and frequency domain features were calculated (see detailed below) and the most relevant ones were identified by straightforward statistical analysis (*t*-tests with *p* values revealing significant modulation by sleep stage). This is referred to here as the feature-space dimension reduction step. Next, the most discriminating features subsets were selected using a standard sequential forward selection procedure: Starting from the feature that provides the highest accuracy, we continue sequentially searching for the next feature that will then collectively provide the highest increase in

Please cite this article in press as: Lajnef T, et al. Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. J Neurosci Methods (2015), http://dx.doi.org/10.1016/j.jneumeth.2015.01.022