Application of Recurrent Neural Network for the Prediction of Target Non-Apneic Arousal Regions in Physiological Signals

Naimahmed Nesaragi[1], Shubha Majumder[2], Ashish Sharma[1],
Kouhyar Tavakolian[2], and Shivnarayan Patidar[1]

[1]Electronics and Communication Engineering Department, National Institute of Technology Goa,
Ponda, India
[2]Electrical Engineering Department, University of North Dakota, Grand Forks, USA

Abstract

This work presents a new method for detection of target non-apneic arousals by applying a recurrent neural network architecture on the various specified polysomnographic (PSG) signals. The proposed two stage architecture uses sequences of instantaneous frequencies and spectral entropies of the chosen PSG signals as feature vectors. At the first stage, these feature vectors are used to train several long-short term memory (LSTM) models. The LSTM networks can learn long-term relationships between time steps of time-frequency based sequences obtained out of physiological signals. As a second stage, some quadratic discriminant (QD) layers are modelled and appended to the trained LSTMs in groups. Subsequently, the outputs of all the QD layers are averaged for making final prediction. The models are trained using features obtained from one minute windows of the signals. However, the decision making on test signals involves inputs of one minute windows with half minute overlapping. When evaluated with 2018 PhysioNet/CinC Challenge dataset, the experimental outcomes demonstrate overall AUROC and AUPRC scores of 0.85±0.10 and 0.50±0.15 respectively for the training data. The generated test results indicate the AUROC and AUPRC scores of 0.624 and 0.10 respectively on a random subset of the test data.

Keywords: PSG, LSTM, Sleep arousal, sleep apnea

1. Introduction

Proper scoring of sleep stages can provide useful clinical information for diagnosis of patients with sleep disorders [1]. Traditionally, such diagnostic procedures are carried out in sleep laboratory settings, where in polysomnography (PSG) of the sleeping subject is carefully inspected by clinical experts to identify potential sleep disorders resulting from sleep arousal. Here, "arousal" signifies the change in state of sleep cycle which is mainly character-ized by abrupt changes in the pattern of brain wave activity measured by electroencephalography (EEG) which represents a shift from deep sleep also termed as Rapid Eye Movement (REM) sleep to light sleep (non-REM) sleep, or from light sleep to wakefulness. Thus, sleep arousals are the disturbances which affects the quality of sleep and lead to a medical condition termed as "sleep disorder". One of the most well-studied sleep disorders is Obstructive Sleep Apnea Hypopnea Syndrome (or simply, apnea).

Apneas are characterized by a complete collapse of the airways, leading to awakening, and consequent disturbances of sleep [2, 3]. Apart from apneic arousals, there are other sources of sleep arousals resulting from periodic limb movements, teeth grinding, chin movements, partial airway obstructions, or even snoring. So, one of the recent challenges in sleep stage scoring and classification is to accurately detect and identify these other sources of arousal (non-apneic) during sleep by monitoring a variety of physiological signals, collected during PSG sleep studies. PSG is the "golden standard" method for assessing sleep disorders in which manual recording of the subject is conducted overnight in a clinical environment [4]. The subject has to sleep while wearing multiple electrodes including that of EEG, electromyography (EMG), electrooculography (EOG), nasal probes, etc. PSG is a manual method which is prone to errors when compared from one expert to another. Moreover, visual inspection is a time-consuming process for a whole night recording. Hence, PSG is labor-intensive and also a costly technique that cannot be applied easily as a screening tool in its original form. [4, 5]. The challenge is to have uniform and standard automated process that can act as a benchmark for primary and expert level diagnosis agreeing completely with PSG ground truth with less deviation in results. For this several computer aided systems using signal processing techniques along with machine learning algorithms are developed to obtain useful information from physiological signals [6–8]. Such methods usually involves 4 steps that include preprocess-

ing, feature extraction, feature selection and classification.

This work aims to develop robust signal processing and machine learning based technique for primary and expert level diagnosis of sleep disorders by detecting arousals during different sleep stages from various physiological signals collected via PSG sleep studies as ground truth. The proposed method employs a two-stage architecture comprising several LSTMs and QD layers at the first and second stages respectively. The LSTM models of the first stage are fed with instantaneous frequency and spectral entropy based feature vectors. All the models are trained under optimal settings. The overall architecture has been trained and validated for classifying the regions of PSG signals into target arousals and non-arousal using the PhysioNet Challenge 2018 dataset.

2. Methods

2.1. Time-frequency based feature extraction

Dimensionally reduced feature space with increased intelligibility can improve the time and space complexity of the algorithm for diagnostic decision making and can lead to improved classification performance.In literature, the time-frequency based image analysis of spectrograms using convolutional neural networks (CNNs) has been found useful for extracting diagnostic features in many applications [9, 10]. On the similar reasoning , LSTM can be used instead of CNNs by translating 2-D images into one dimensional signals. In practice, this can be achieved by extracting time-frequency moments from the spectrograms. This study explores two such moments in the time-domain namely instantaneous frequency and spectral entropy for detecting sleep arousals.

The instantaneous frequency is the time-dependent frequency of a signal which is estimated as the first moment of the power spectrogram by computing short-time Fourier transforms.For a 60 seconds segment of a given PSG signal sampled at 200Hz, a feature vector of 258 length is obtained by computing spectrograms over 258 time windows . The time outputs for the instantaneous frequency values correspond to the center of the time windows.

The spectral entropy measures the spikes or flatness present in the spectrum for a given signal. A signal with a spiky spectrum exhibits low spectral entropy. A signal with a flat spectrum, like white noise, exhibits high spectral entropy. The procedure to estimate spectral entropy is similar to that of instantaneous frequency and uses 258 time windows for a 60 seconds segment of the signal under study. Figures 1 and 2 illustrate the differences in the instantaneous frequency and spectral entropy for windowed arousal and non-arousal regions of the typical EMG and

EEG signals respectively.

2.2. Classification

The LSTM is preferred for handling raw or transformed time-series. Basically, neurons in LSTM keep the context of memory within their pipeline that allows capturing temporal information present in input sequences. LSTM networks were introduced by Sepp Hochreiter and Jrgen Schmidhuber [11]. The proposed method employs a two stage architecture as depicted in Figure 3. At the first stage, the time-frequency based feature vectors as described in Section 2 are used to train several LSTM models with optimal configurations. At the second stage, some quadratic discriminant (QD) layers are trained from the outputs of trained LSTMs in groups. The trained QD layers are used as the second classification stage. Eventually,the outputs of all the QD layers are averaged for making final prediction with improved machine learning performance.

3. Results and Discussion

The proposed work has been evaluated using 2018 PhysioNet/CinC Challenge dataset.The specified training and testing set contains 994 and 989 sets of signals respectively [12].The dataset has a variety of physiological signals recorded from subjects as they slept through the night including: EEG, EMG, EOG, ECG, abdominal(ABD),chest(thorax),airflow and oxygen saturation (SaO2). Excluding SaO2, all signals were sampled to 200Hz and were measured in microvolts.

In this work, for preparing the first stage LSTM based models of the proposed two stage architecture, instantaneous frequency and spectral entropy based features of the all the signals except SaO2 are extracted. Each PSG signal contributes two feature vectors, therefore with a total of 12 PSG signals the LSTM network architecture has been specified with the input sequence size as 24. A bidirectional LSTM layer with an output size of 50 is specified. The two classes (non-apneic arousal and non-arousal) are specified by including a fully connected layer of size 2, followed by a softmax layer and a classification layer. It is noteworthy that the used feature vectors are found to have average values that differ significantly in magnitude. This can lead to slow learning and convergence of LSTM [13]. In view of these, the input feature vectors are standardized before feeding to bidirectional LSTM. Standardization, or z-scoring, is a popular way to improve network performance during training and has been used in this work. Furthermore, the LSTM models are trained with error back-propagation using Adaptive Momentum with an initial learning rate of 0.01. The optimizer and categorical cross-entropy is set as the loss function. 90 % of training data were used to train algorithm and the rest of 10 %
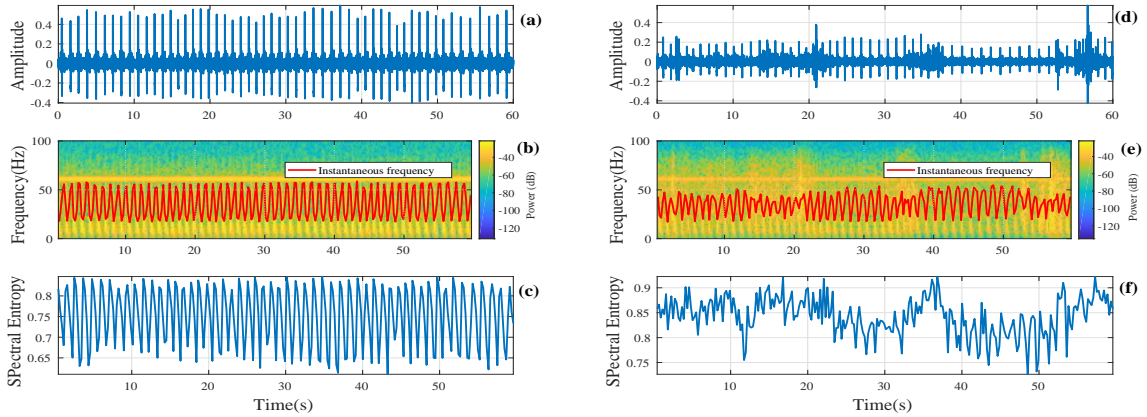
Figure 1. (a) and (d) show the arousal and non-arousal segments of an EMG signal.(b) and (e) are the plots of respective instantaneous frequencies, and (c) and (f) are the plots of spectral entropies.
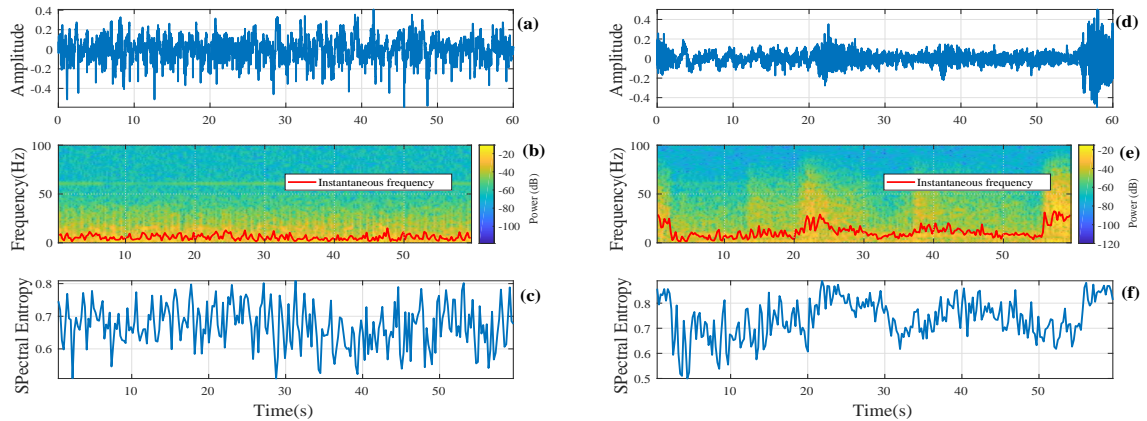

Figure 2. (a) and (d) show the arousal and non-arousal segments of an EEG signal.(b) and (e) are the plots of respective instantaneous frequencies, and (c) and (f) are the plots of spectral entropies.

were kept as a validation sub-set. While preparing the second stage of the proposed architecture, one QD layers is trained from the outputs of 10 trained LSTMs in groups. In all, 9 such QD layers are formed for second stage classification. Eventually,the outputs of all the QD layers are averaged for making final prediction with improved classification performance.

We have observed a great improvement in the training accuracy, which is greater than 90% with overall AUROC and AUPRC scores of 0.85±0.10 and 0.50±0.15 respectively. The generated test results on hidden test data indicate the AUROC and AUPRC scores of 0.624 and 0.10 respectively on a random subset of the data. Average running time (test set) was found to be 0.306% of quota when ran on PhysioNet challenge machine.

4. Conclusion

In this work, we have explored the strength and applicability of instantaneous frequency and spectral entropy based features when used with LSTM classifier for detection of target non-apneic arousals. Our method has shown sparking performance on a large and diverse dataset with high variant records. With some advancements, this work has clinical potential to be realized into an automatic real-time system for detection of sleep disorders. The future scope of the work includes (a) application of some transforms (say, Tunable-Q wavelet transform) to obtain relevant sub-bands that may carry more clear underlying physiological information on arousals for more meaningful feature extraction. (b) tuning the involved LSTM models for optimal operations and (c) inclusion of more feature vectors for passing more information into the system.
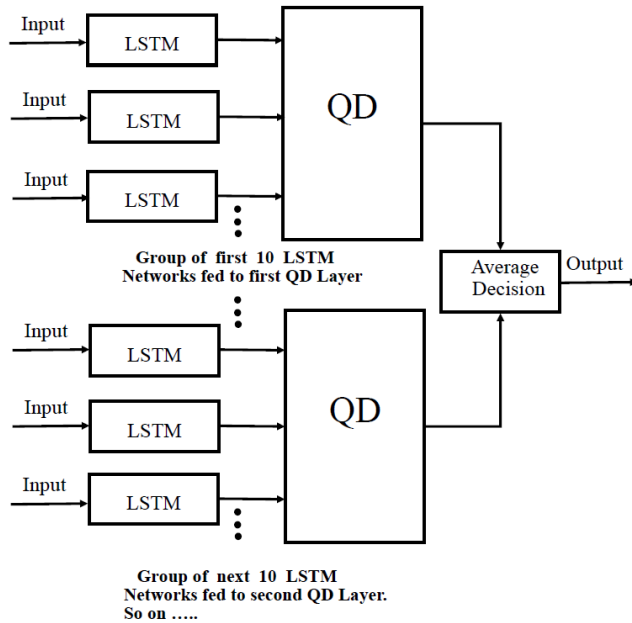
Figure 3. The proposed architecture for classification.

Acknowledgements

References

[1] Lee-Chiong T. Sleep Medicine: Essentials and review. Oxford University Press, 2008.

[2] Marshall NS, Wong KK, Cullen SR, Knuiman MW, Grunstein RR. Sleep apnea and 20-year follow-up for all-cause mortality, stroke, and cancer incidence and mortality in the busselton health study cohort. Journal of Clinical Sleep Medicine 2014;10(04):355–362.

[3] Parati G, Lombardi C, Hedner J, Bonsignore MR, Grote L, Tkacova R, Lévy P, Riha R, Bassetti C, Narkiewicz K, et al. Recommendations for the management of patients with obstructive sleep apnoea and hypertension. European Respiratory Journal 2012;erj02267–2011.

[4] Malhotra A, Younes M, Kuna ST, Benca R, Kushida CA, Walsh J, Hanlon A, Staley B, Pack AI, Pien GW. Performance of an automated polysomnography scoring system versus computer-assisted manual scoring. Sleep 2013; 36(4):573–582.

[5] Collop NA. Scoring variability between polysomnography technologists in different sleep laboratories. Sleep Medicine 2002;3(1):43–47.

[6] Boostani R, Karimzadeh F, Nami M. A comparative review on sleep stage classification methods in patients and healthy individuals. Computer Methods and Programs in Biomedicine 2017;140:77–91.

[7] Gharbali AA, Najdi S, Fonseca JM. Investigating the contribution of distance-based features to automatic sleep stage classification. Computers in Biology and Medicine 2018; 96:8–23.

[8] Khalighi S, Sousa T, Pires G, Nunes U. Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels. Expert Systems with Applications 2013;40(17):7046–7059.

[9] Pons J, Lidy T, Serra X. Experimenting with musically motivated convolutional neural networks. In Content-Based Multimedia Indexing (CBMI), 2016 14th International Workshop on. IEEE, 2016; 1–6.

[10] Wang D. Deep learning reinvents the hearing aid. IEEE Spectrum 2017;54(3):32–37.

[11] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation 1997;9(8):1735–1780.

[12] Ghassemi MM, Moody BE, Lehman LH, Song C, Li Q, Haoqi S, Roger GM, Westover BM, Clifford GD. You snooze, you win: the physionet/computing in cardiology challenge 2018. volume 45. IEEE, 2018; 1–4.

[13] Brownlee J. How to scale data for long short-term memory networks in python. 7 july 2017.

Address for correspondence:

Dr Shivnarayan Patidar
Department of Electronics and Communication Engineering
National Institute of Technology Goa, Ponda
India, 403401
shivnarayan.patidar@nitgoa.ac.in