A Long Short Term Memory Network based method for Measurement of Sleep Arousals using Polysomnographic Time-Series

Naimahmed Nesaragi

Under the supervision of Dr.Shivnarayan Patidar

and co-supervision of Dr. T.Veerakumar

Department of Electronics and Communication Engineering
National Institute of Technology Goa

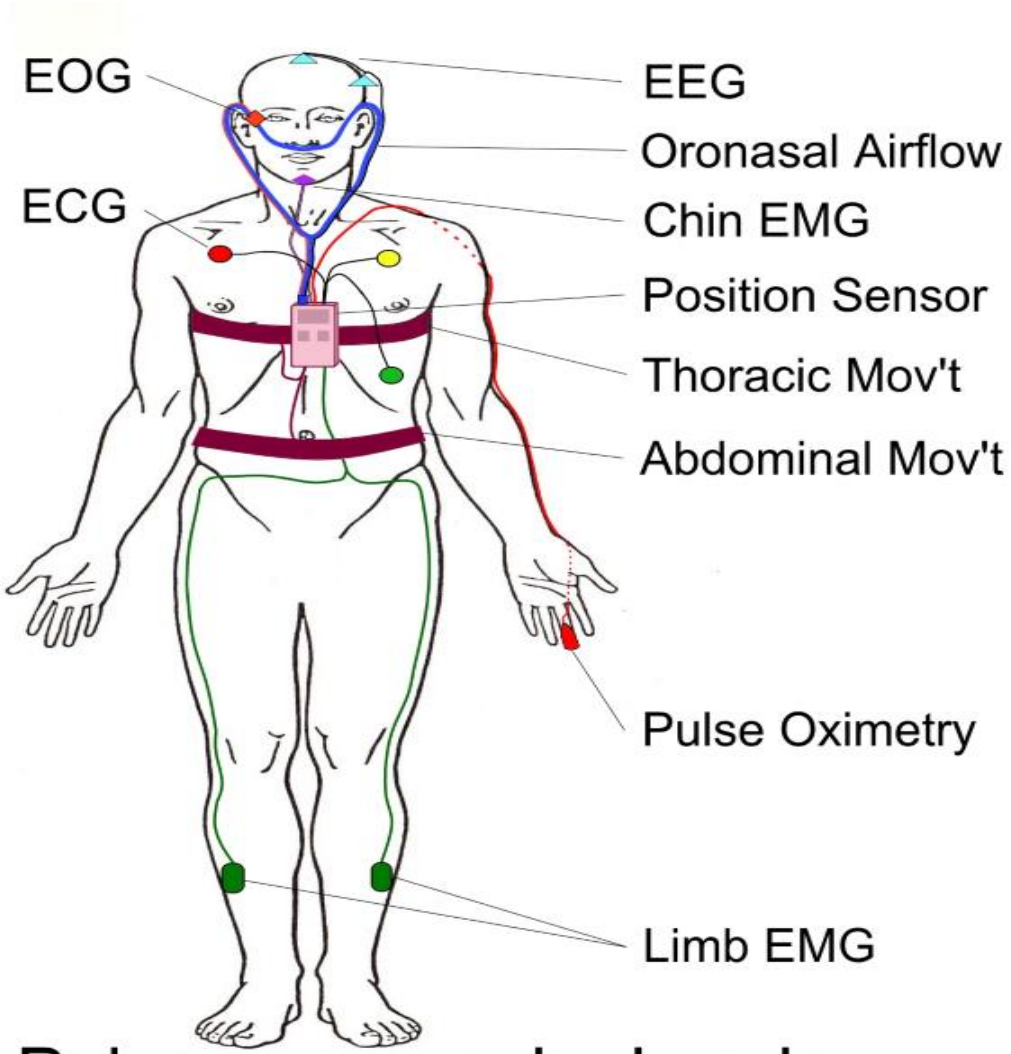
November 2018

Introduction

- In order to **improve the quality of sleep**, proper diagnosis of sleep disorders is of utmost importance [Lee-Chiong ,T et.al]. This in turn defines the good health and wellbeing of an individual.

- 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/to sleep arousal**.

What is Polysomnography (PSG) ?



Polysomnography Leads

What is Polysomnography (PSG) ?

- PSG is the "golden standard" method for assessing sleep disorders in which **manual recording** of the subject is conducted overnight in a clinical environment [Malhotra A,et.al].
- The PSG monitors many body functions including brain(EEG),eye movements (EOG), muscle activity (EMG), and heart rhythm (ECG), during sleep.
- This manual method is usually prone to errors when carried out from one expert to another. Moreover, visual inspection is a time consuming process for a whole night recording [Malhotra A,et.al, Collop NA,et.al].

What is an Arousal/Sleep Arousal?

- "Arousals" are brief intrusions of wakefulness during sleep cycle, after which sleep resumes.
- Arousal reflects an abrupt change in the pattern of brain wave activity measured by EEG, which represents a shift from deep sleep(REM sleep) to light sleep(NREM 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.

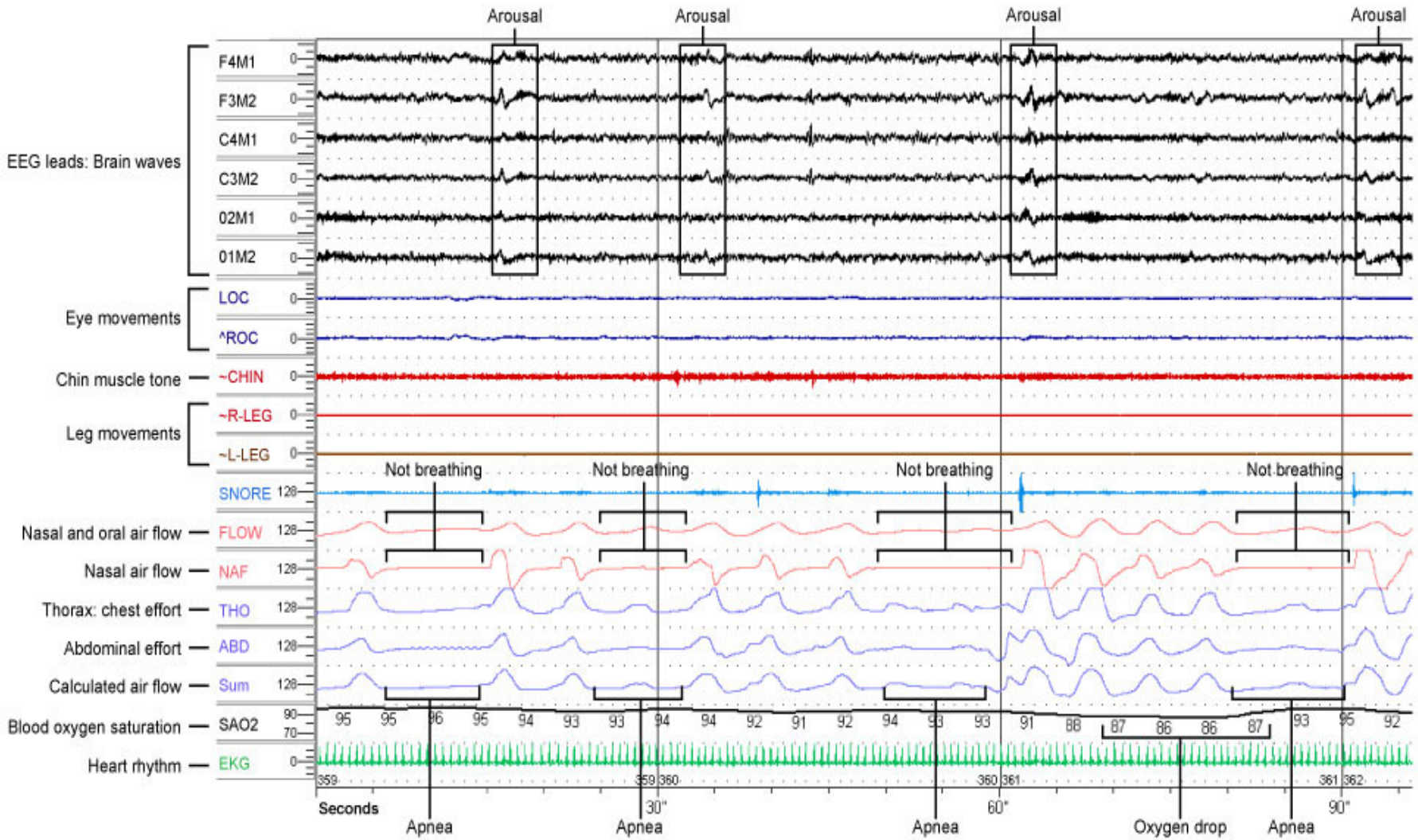
Sleep Disorders

- 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 airway, leading to awakening, and consequent disturbances of sleep. While apneas are arguably the best understood of sleep disturbances, they are not the only cause of disturbance/arousal [Marshall .NS et.al, Parati ,G et.al].
- Sleep arousals can also be spontaneous, result from teeth grinding, snoring, periodic limb movements etc.
- This causes Sleep fragmentation – frequent interruption of sleep by arousals results in daytime sleepiness, obesity, depression.

Sleep Stages

- Sleep occurs in 3 basic states throughout the night. The 3 main stages of sleep are wakefulness (W), rapid eye movement sleep (REM) and non-rapid eye movement sleep (NREM).
- REM comprises the smaller portion of the sleep cycle.
- NREM has 3 sub stages N1,N2,N3 comprising the larger portion of the overall sleep cycle.
- Arousal occurs more often during NREM sleep.

PSG Waveforms



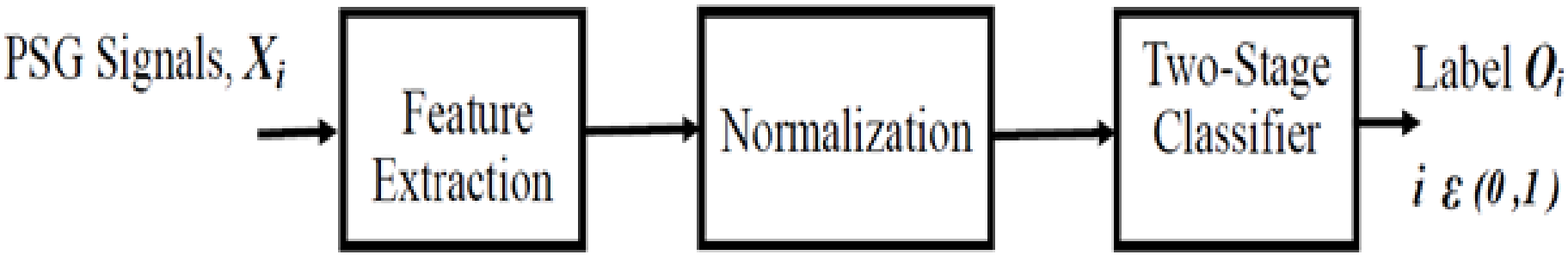
Existing Research Methods for Sleep Data Representation and Challenges

- Recent data mining techniques for sleep data representation deal with analysis and classification of different biomedical signals specially Polysomnography recordings for **automatic classification of sleep stages**.
- Limited amount of research is carried out for detection of sleep disorders dealing with **Apneic Arousals** classification only.
- The challenge faced with non-apneic arousal classification is **Data Imbalance**
- For this several computer aided systems using signal processing techniques along with machine learning algorithms are developed to obtain useful information from physiological signals [Boostani R,et.al, Gharbali AA,et.al, Khalighi S,et.al]

Problem Statement

- Therefore the challenge to the biomedical research community is to demonstrate the efficacy of automated methods for **non-apneic arousal detection** .
- Using a large, well-characterized, and representative set of PSG data as ground truth and advance the state of the art in this clinically significant problem.

Methodology



- 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 **intelligent time series** which are the transformed version of the PSG raw time series. The transformed series are the sequences of Instantaneous frequencies and spectral entropies

Time-frequency based feature extraction

- 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 [Pons J,et.al, Wang D,et.al].
- 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.

Spectral entropy (SE)

- 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. Steps to calculate SE as a moment from spectrograms:

  1.Calculate the spectrum $\mathbf{X}(w_i)$

  2.Calculate the PSD $\mathbf{P}(w_i) = \frac{1}{\mathbf{N}} |X(w_i)|^2$

  3.Normalize the calculated PSD to PDF

$$p_i = \frac{P(w_i)}{\sum_i P(w_i)}$$

  4.Spectral entropy

$$SE = -\sum_{i=1}^{n} p_i log_2 p_i$$

Case Specific calculations

- 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 spectral entropy values correspond to the center of the time windows.

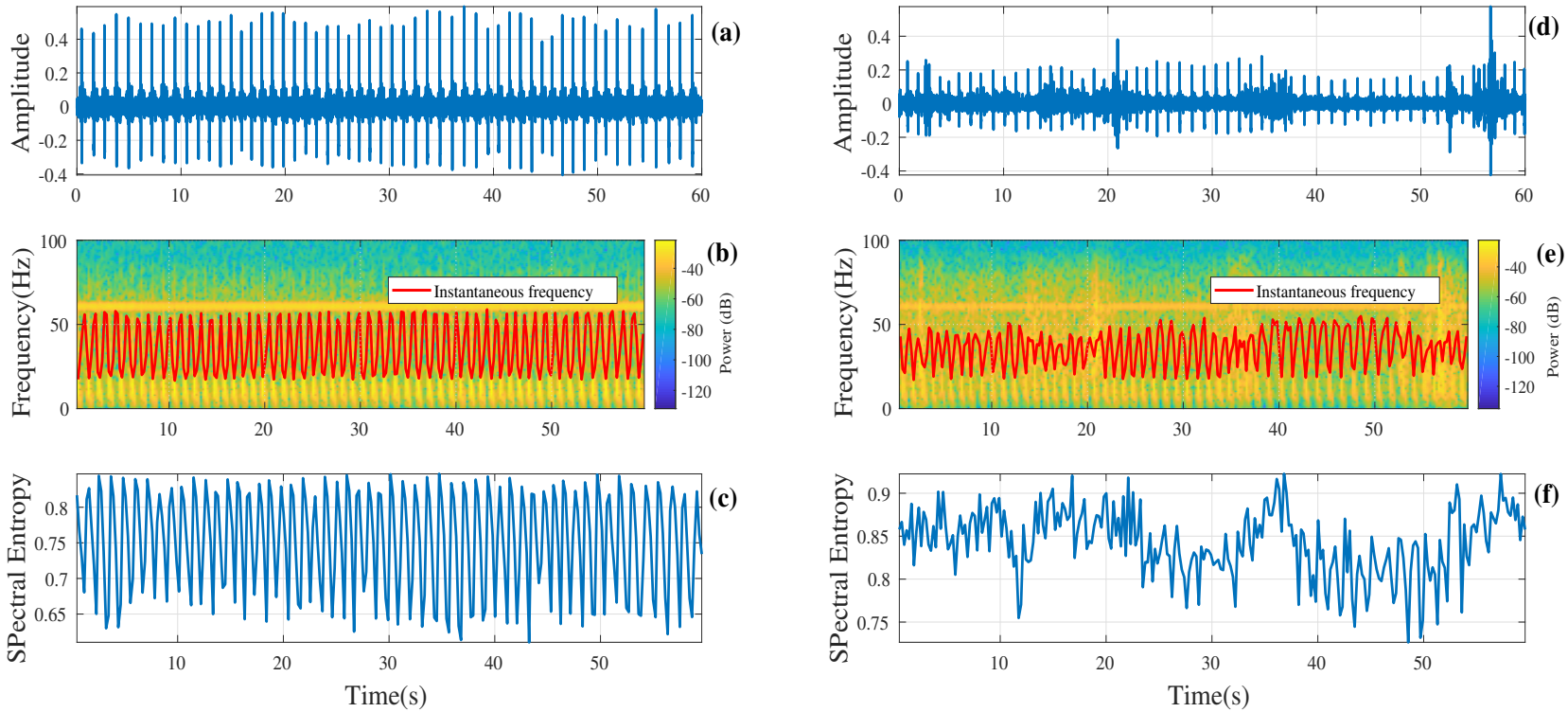
Instantaneous Frequency

- It can be obtained from a time-frequency distribution (TFD)as the first conditional moment in frequency, suggesting that the IF is the average frequency at each time **t** [Cohen,et.al].
  This is achieved by computing short-time Fourier transforms.

$$f_i(t) = \frac{\int_{-\infty}^{\infty} fP(t,f)df}{\int_{-\infty}^{\infty} P(t,f)df}$$

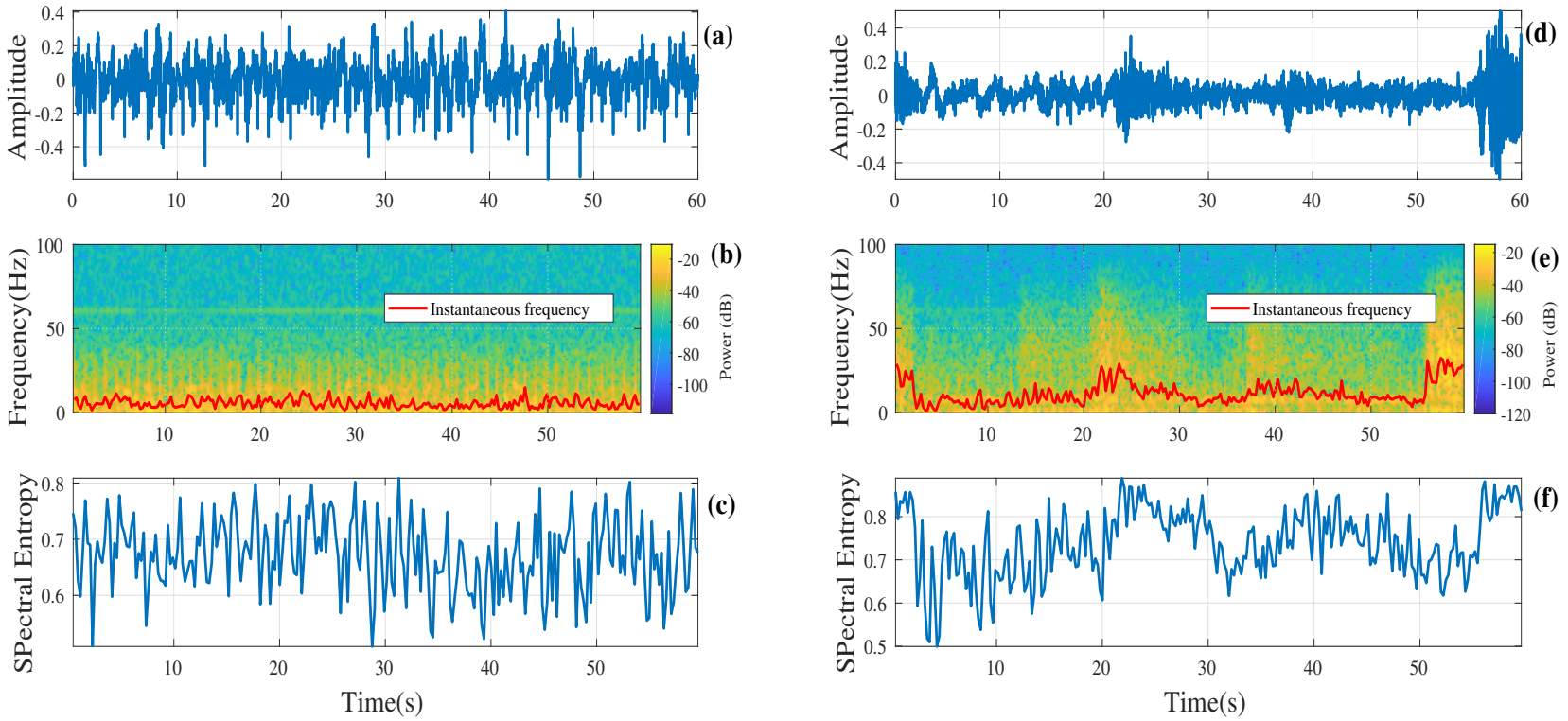
where P(t,f) is Time frequency representation

Typical Binary classification in case of Chin-EMG signal



(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.

Typical Binary classification in case of EEG signal



(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.

LSTM-Long Short term Memory

- The LSTM networks can learn long-term relationships between time steps of time-frequency based sequences obtained out of physiological signals [Schmidhuber J,et.al]. Basically, neurons in LSTM keep the context of memory within their pipeline that allows capturing temporal information present in input sequences.

LSTM input Layer

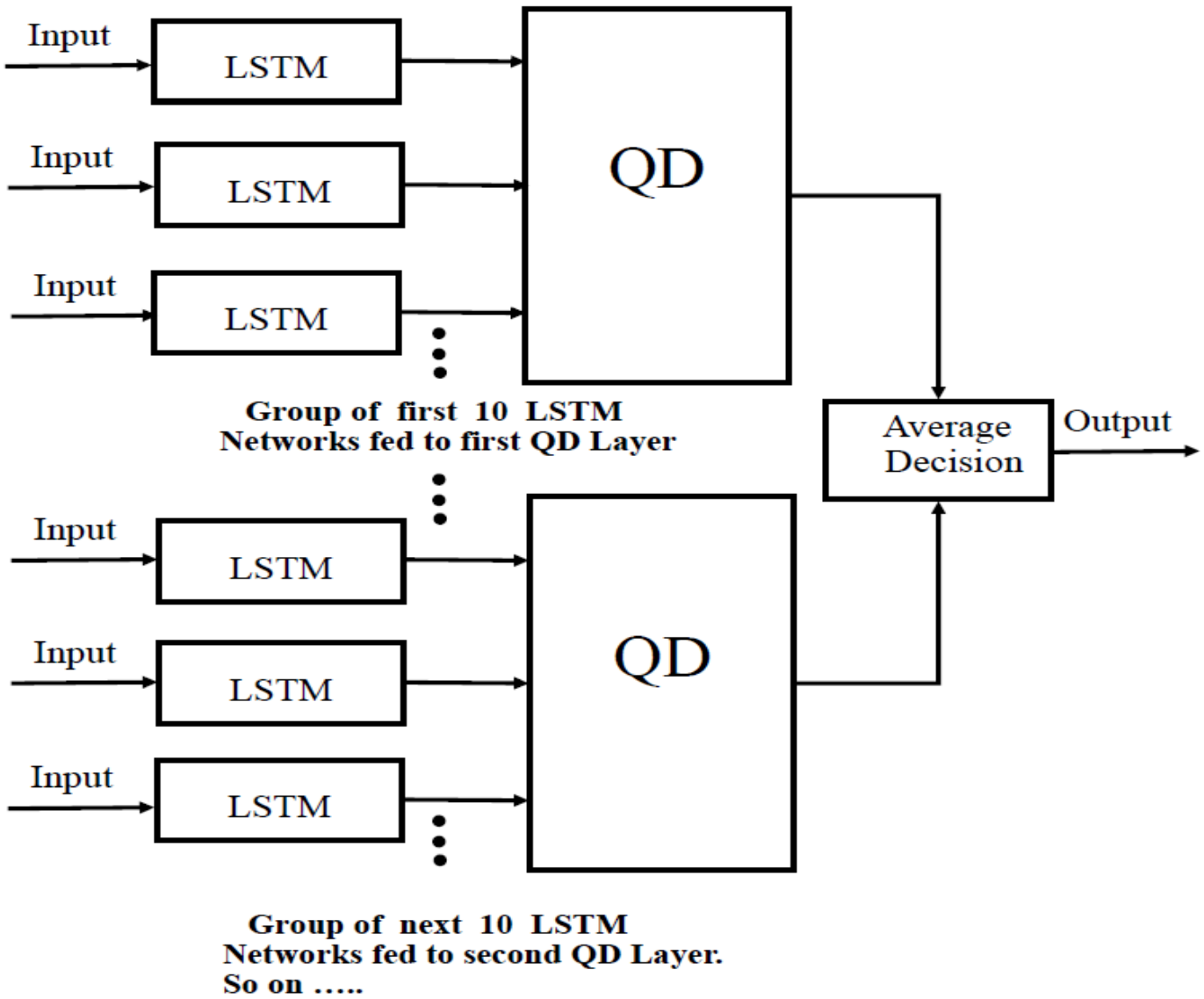
- The LSTM input layer is specified by the "*input* – *shape*" argument on the first hidden layer of the network.
- The input to every LSTM layer must be three-dimensional.The three dimensions of this input are:

  Samples: One segment/epoch corresponds to one sample. A batch is comprised of one or more samples.

  Time Steps: One time step is one point of observation in the sample.

  Features: One feature is one observation at a time step.

Proposed Architecture for 2-Stage Classifier



Work-Flow of Algorithm

- The proposed method involves design of 90 such LSTM models from first 900 training records (subjects), each model is being trained with data of 10 records.
- Sequences of Instantaneous frequencies and Spectral entropies are used as compatible inputs for LSTM.
- The LSTM output has predictions as sequences of 0s (non-arousal) and 1s (Target apneic arousals).So out of 90 LSTM models we have 90 sequences of 0s and 1s.
- Next ,these 90 predictions are converted to 9 groups of 10 sequences each and are used to model 9 QD layers as a second stage.
- So overall , at first stage, ninety LSTM models are trained then, nine QD layers are modeled one for set of ten LSTMs.
  Note: However, the method rely on LSTM to learn and come up with suitable function as features from training data that can discriminate the 2 classes.

Challenge Data

- The proposed work has been evaluated using 2018 PhysioNet/CinC Challenge dataset, named "You Snooze - You Win".
  The specified training and testing set contains 994 and 989 sets of signals respectively [Moody BE,et.al].
  The arousal labels for 994 of the recordings were made available in a public training set while 989 labels were retained in a hidden test set.
- The dataset has a variety of 13 physiological signals recorded from subjects as they slept through the night which includes EEG(6 channels),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.

Expert Labeling

- The EEG signals were scored in non overlapping 30-second epochs according to the AASM standards as one of five stages: wake (W), rapid eye movement (REM), non-REM stage 1 (N1), non-REM stage 2 (N2), and non-REM stage 3 (N3). Subject waveforms were also annotated for the presence of arousals that interrupted their sleep.
- The annotated arousals were classified as either:spontaneous arousals, respiratory effort related arousals(RERA), bruxisms (teeth grinding),apneas (central, obstructive and mixed), vocalizations, snores, periodic leg movements, partial airway obstructions.
- Sleep arousal annotations are contained in a sample-wise vector made up of 1s 0s and -1s.The target non-apneic arousal regions are marked by "1". Apneic regions are marked by a "-1"(not scored) where subject is sleeping are marked by "0".

Scoring

- Final algorithm was graded for its binary classification performance on target arousal and non-arousal regions, (designated by $+1$ and 0 respectively) as measured by the area under the precision-recall curve (AUPRC). Precision (p_j) and recall (r_j) were defined as follows:

$$p_j = \frac{TP}{TP + FP} = \frac{|A \cap P_j \cap \overline{\mathbb{N}}|}{|P_j \cap \overline{\mathbb{N}}|}$$

$$r_j = \frac{TP}{TP + FN} = \frac{|A \cap P_j \cap \overline{\mathbb{N}}|}{|A \cap \overline{\mathbb{N}}|}$$

where $\mathbb{N}$ indicates the set of non-scored samples, A indicate the set of target arousal samples, and P_j indicates the set of samples for which the predicted arousal probability was at least $\frac{j}{1000}$.

Results

Proposed Method

Metrics	Scores (Test Data)	Train Data
AUROC	0.624	0.85
AUPRC	0.10	0.35
Average run time (test set):	0.306%	of quota
Maximum run time (test set):	0.32%	of quota

Top 3 Teams	Test Data AUROC	Test Data AUPRC
1	0.90	0.45
2	0.85	0.42
3	0.78	0.21

Future Scope and Conclusion

- In this work, we have explored the strength and applicability of instantaneous frequency and spectral entropy based time series when used with LSTM classifier for detection of target non-apneic arousals.
- Our method has shown decent 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.

Research Publications

Nesaragi, N., Majumder, S., Sharma, A., Tavakolian, K., Patidar, S.(2018)

Application of recurrent neural network for the prediction of target non-apneic arousal regions in physiological signals.

Computing In Cardiology, 45^{th} IEEE Conference,Sep 24-26, 2018,Maastricht, Netherlands.

References

Lee-Chiong T.(2008)

Sleep Medicine: Essentials and review.

Oxford University Press, 2008.

Marshall NS, Wong KK, Cullen SR, Knuiman MW, Grunstein RR. (2014)

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),pp:355-362.

Parati G, Lombardi C, Hedner J, Bonsignore MR, Grote L, Tkacova R, Levy P, Riha R, Bassetti C, Narkiewicz K (2011)

Recommendations for the management of patients with obstructive sleep apnoea and hypertension.

European Respiratory Journal 2012;erj02267, 2011.

Malhotra A, Younes M, Kuna ST, Benca R, Kushida CA, Walsh J, Hanlon A, Staley B, Pack AI, Pien GW (2013)

Performance of an automated polysomnography scoring system versus computer-assisted manual scoring.

Sleep 2013; 36(4), pp:573-582.

Collop NA.(2002)

Scoring variability between polysomnography technologists in different sleep laboratories.

Sleep Medicine, 2002;3(1),pp::43-47.

Boostani R, Karimzadeh F, Nami M. (2017)

A comparative review on sleep stage classification methods in patients and healthy individuals.

Computer Methods and Programs in Biomedicine 2017;140,pp:77-91.

Gharbali AA, Najdi S, Fonseca JM.(2018)

Investigating the contribution of distance-based features to automatic sleep stage classification.

Computers in Biology and Medicine 2018;96:8-23.

Khalighi S, Sousa T, Pires G, Nunes U.(2013)

Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels.

Expert Systems with Applications 2013;40(17):7046-7059.

Pons J, Lidy T, Serra X.(2016)

Experimenting with musically motivated convolutional neural networks.

Content-Based Multimedia Indexing (CBMI), 2016 14^th International Workshop on IEEE, 2016;pp: 1-6.

Wang D. (2017)

Deep learning reinvents the hearing aid.

IEEE Spectrum 2017;54(3),pp:32-37.

Hochreiter S, Schmidhuber J.(1997)

Long short-term memory.

Neural Computation 1997;9(8),pp:1735-1780.

Ghassemi MM, Moody BE, Lehman LH, Song C, Li Q, Haoqi S, Roger GM, Westover BM, Clifford GD.(2018)

You snooze, you win: the physionet/computing in cardiology challenge 2018. volume 45. IEEE, 2018; 14.

Computing In Cardiology 2018. volume 45. IEEE, 2018;pp: 1-4.

Cohen, Leon.(1995)

Time-frequency analysis.

Prentice Hall PTR Englewood Cliffs, N.J, 1995.

Thank You