proceedings

MDPI

Proceedings

Severe obstructive sleep apnea event detection from EEG recordings †

Lampros Kokkalas, Georgia Korompili, Nikolas-Alexander Tatlas, Stelios A. Mitilineos and Stelios M. Potirakis *

Department of Electrical and Electronic Engineering, University of West Attica, Attica 12243, Greece; lkokkalas@uniwa.gr (L.K.); gkorompili@uniwa.gr (G.K.); ntatlas@uniwa.gr (N.-A.T.); smitil@uniwa.gr (S.A.M.)

* Correspondence: spoti@uniwa.gr; Tel.:+30-210-538-1550

† Presented at 2nd International Electronic Conference on Applied Sciences, 15–31 October 2021; Available online: https://asec2021.sciforum.net/.

Abstract: Sleep apnea is a serious disorder where breathing stops frequently during sleep. Changes in brain activities that occur during apnea can be detected with an electroencephalogram (EEG). Although accurate detection of apnea events is very important, there is currently no algorithm that can efficiently measure the onset and end of apnea events based only on electroencephalogram signals. The number and duration of apnea events are used to calculate apnea-hypopnea index (AHI) and mean apnea-hypopnea duration (MAD), that are indicators of obstructive sleep apnea severity. Previous apnea detection algorithms usually focus on the classification of apnea patients and not specific apnea events, or perform a frame-by-frame analysis and classify each frame based on the global characteristics of the frame, instead of locating the onsets and ends of apnea events. Thus, the clinical significance of EEG signals for apnea detection is limited to sleep staging. The purpose of this study is to propose a method for sleep apnea event detection and event duration evaluation using Convolutional Recurrent Neural Networks, based only on EEG signals. Reference and estimated AHI are strongly correlated ($r=0.88$, $p<0.001$), whereas the sensitivity and positive predicted value for the individual events detection is 0.73 and 0.78, respectively. Reference and estimated MAD values are very highly correlated ($r=0.91$, $p<0.001$), and the absolute error between them is 2.05 ± 1.66 s. The proposed method has high accuracy in detecting individual apnea events from EEG signals, especially in severe apnea cases.

Keywords: EEG signals; sleep apnea; apnea–hypopnea index; mean apnea-hypopnea duration; deep neural network; convolutional recurrent neural network

Citation: Kokkalas, L.; Korompili ,G.; Tatlas, N. A.; Mitilineos, S. A.; Potirakis, S. M. Severe obstructive sleep apnea event detection from EEG recordings. *Proceedings* **2021**, *68*, x. https://doi.org/10.3390/xxxxx

Published: date

1. Introduction

Obstructive sleep apnea (OSA) is a sleep breathing disorder that is characterized by frequent cessations in breathing or reductions in breath amplitude [1]. Sleep apnea patients experience fragmented sleep and difficulty concentrating during the day that reduces quality-of-life and increases the risk of car accidents, while cardiovascular system strain due to low levels of blood oxygen is linked to cardiac diseases and reduced life expectancy [2,3]. The gold standard for diagnosing sleep-related breathing disorders is polysomnography that is a sleep study performed at a sleep center, which collects and records many physiologic parameters during sleep, including electroencephalogram, electrocardiogram, respiratory effort and blood oxygen levels [4]. The diagnosis of the disorder currently depends on the apnea-hypopnea index (AHI) that is the number of annotated sleep apnea and hypopnea events per hour of sleep. Based on the AHI, the severity of OSA is classified as follows: severe, where AHI is greater than 30, moderate, where AHI is between 15 and 30, mild, where AHI is between 5 and 15 and normal where there are less than 5 events per hour of sleep. In addition to AHI, mean ap-

nea-hypopnea duration (MAD) has been proposed lately as an indicator of the OSA severity [5,6].

Polysomnography is an inconvenient and costly process and many researchers have investigated methods for apnea detection based on single biological signals, such as the electrocardiogram [7,8], oxygen saturation [9], airflow [10], thoracic movement [11] and EEG. EEG detects electrical activity in the brain using electrodes attached to the scalp [12,13]. It has been successfully used for automatic sleep staging [14] and there are many efforts to use it also for sleep apnea severity estimation. Many of the reported methods consider classification between apnea and healthy subjects [15,16], while the rest try to discriminate apnea and non-apnea events using a frame-by-frame analysis, where frame duration is usually 30 s [17,18]. In frame-by-frame analysis the whole duration of a test frame is considered for feature extraction, which is not clinically meaningful for AHI estimation, since a frame may contain a partial apnea event or more than one apnea events.

This paper proposes the development of an apnea event detection model that depends only on EEG signals and detects individual events and their duration.

2. Materials and Methods

2.1. PSG data acquisition

The database used in this study consists of PSG data obtained at the Sismanoglio – Amalia Fleming General Hospital of Athens [19]. The study consists of 234 PSG examinations that were performed for the diagnosis of sleep-disordered breathing between April 2019 and November 2020. 174 cases were used to develop the neural network and the remaining 60 cases were used for testing. Of the 234 patients, 176 (75%) were male. The mean age of the patients was 57 years, the mean AHI was 57.93 events/h, and the mean MAD was 18.6 s. Overall 2 patients were normal, 6 patients had mild OSA, 29 patients had moderate OSA and 197 patients had severe OSA. Three EEG channels (C3-A2, C4-A1 and reference A1-A2) were used from the PSG data with a sampling ratio of 200 Hz. Sleep stages and respiratory events scoring was based on the general instructions for sleep stage labeling [20], using Sleepware G3 PSG software.

2.2. EEG signal preprocessing

EEG recordings from the three channels are divided into non-overlapping frames of 300 s width. For each frame, the spectrogram is produced and the generated time series data per frequency for each channel are considered as separate input features. Finally, Z-normalization (standardization) is applied that speeds up training and ensures convergence, especially when features have different dynamic range [21].

2.3. Convolutional recurrent neural network

The convolutional recurrent neural network (CRNN) is a combination of a convolutional neural network (CNN) and a recurrent neural network (RNN) [22-24]. In order to determine the exact time that each apnea starts and ends, a combination of 1D convolution layer (Conv1D) and Gate Recurrent Unit (GRU) is used. Conv1D extracts local frequency-time features, while the following GRU is used for the overall apnea time modeling (Figure 1).

2.4. Neural network training

EEG recordings are divided into non-overlapping frames and a label vector is generated for each frame. When the frame contains an annotated apnea event, as indicated in the PSG data, the corresponding time steps of the label vector are set to the value 1, otherwise they are set to the value 0. In inference mode, a batch of consecutive ones (1) is considered a predicted apnea if it corresponds to more than 10 s in duration and a pre-

dicted apnea is considered a True Positive if it overlaps with a scored apnea event, otherwise it is considered a False Positive.

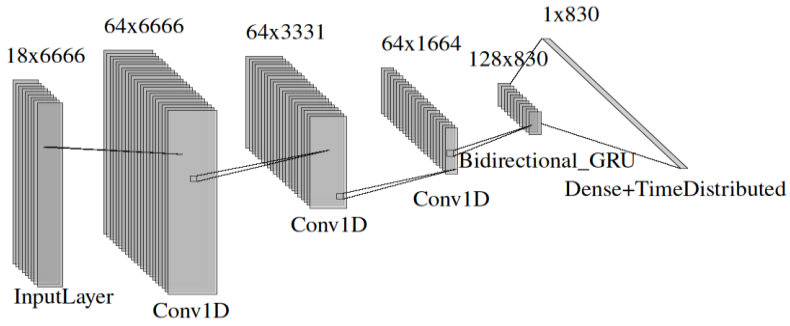

Figure 1. Proposed apnea event detection model architecture. Three CONV-1D layers followed by a Bidirectional GRU layer and finally a TimeDistributed Dense layer.

3. Results

3.1. Event duration detection

The agreement between the estimated MAD per patient and the reference MAD is presented with a scatter plot and a Bland-Altman plot (Figure 2). Estimated and reference MAD are very highly correlated (r=0.91, p<0.001), while the absolute error between the MAD values is 2.05 ± 1.66 s (mean ± standard deviation, SD). The Bland-Altman plot represents the difference between the reference and estimated MAD values against the averages of the two MAD values. The mean difference of the MAD values is -0.8 s, with a 95% confidential interval ranging from -5.7 to 4.1 s.

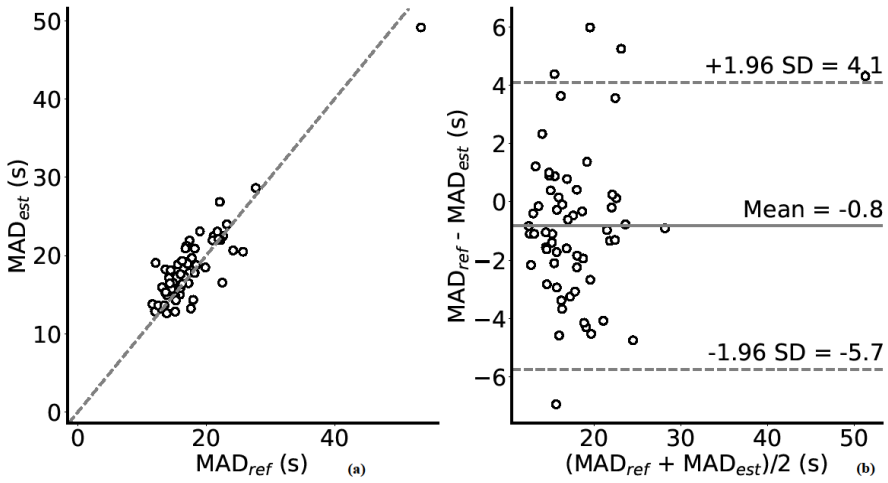

Figure 2. Agreement between the predicted and reference Mean Apnea-Hypopnea Duration. **(a)** Predicted vs. reference value; **(b)** Bland-Altman plot.

3.2. Apnea-hypopnea index

Reference and estimated AHI are strongly correlated (r= 0.88, p<0.001), while the mean difference of the AHI values is 7.3 events/h, with a 95% confidential interval ranging from –18.8 to 33.3 events/h (Figure 3).

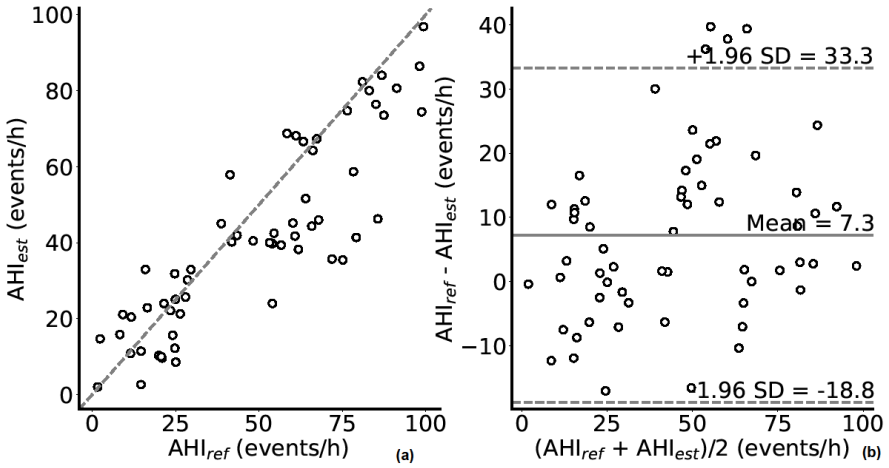

Figure 3. Agreement between the predicted and reference Apnea-Hypopnea Index. **(a)** Predicted vs. reference value; **(b)** Bland-Altman plot.

3.3. Individual apnea events detection

Per-apnea agreement for the test data recordings is shown in Table 2. Sensitivity and positive predictive value are 0.73 and 0.78, respectively. For the severe group, sensitivity is 0.87 and positive predictive value is 0.8, while both values decrease as we move from moderate to mild and finally to normal group. Hypopneas are the dominant type for the normal, mild and moderate group (83%, 58% and 51% of the total number of apneas respectively), while obstructive apneas are the dominant type for the severe group (61%).

Table 2. Performance results using individual apnea events for the 60 patients comprising the test set.

Group	Normal	Mild	Moderate	Severe	All
Group members	2	6	16	36	60
Annotated apneas	18	354	1593	8474	10439
True Positive	2	82	801	6757	7642
False Positive	69	353	772	1013	2207
Sensitivity	0.11	0.23	0.5	0.8	0.73
Positive Predictive Value	0.03	0.19	0.51	0.87	0.78

4. Discussion

The apnea event detection model proposed here, exhibits high performance in the MAD detection and estimates AHI with accuracy for all apnea severity groups. It also detects individual events with high sensitivity and positive predictive value in the severe apnea group.

The model exhibits high performance in the MAD detection and correlation is strong for all patient groups, normal, mild, moderate and severe. The absolute error between the MAD for the test set is 2.05 ± 1.66 s. This error is quite small, if we take into account that oxygen desaturation affects brain activity several seconds after the apnea onset. Since MAD can vary significantly in patients with the same AHI, its value can be used with AHI, to estimate the possible sleep apnea implications.

Model generated AHI has strong correlation with the reference AHI, but is underestimated, since it is calculated based on recording time, whereas reference AHI is calculated based on sleep time. A future improvement that distinguishes sleep/wake status using the EEG signals, would improve further the AHI estimation performance.

Individual apnea events detection for patients characterized by severe apnea-hypopnea syndrome has high sensitivity and positive predicted value, while both metrics are decreased in patients that belong to the other three groups, moderate, mild and normal. There are two possible causes for these results. The first one is that the dataset imbalance between the severity classes may strongly affect the way the model is trained and therefore the provided accuracy. Our neural network model is trained mostly on the severe group and it is for this group that it achieves the best performance. The second reason is that normal and mild groups are dominated by hypopneas. According to the American Academy of Sleep Medicine manual, a hypopnea is detected when there is a reduction of more than 30 % in nasal pressure, in relation to pre-event baseline, that lasts more than 10 s and is associated with more than 3 % oxygen desaturation or an arousal. This definition induces a level of uncertainty in hypopnea scoring, due to variability in flow measurements and visual inspection [25].

A limitation of this study, as mentioned before, is the imbalanced data that is used. The dataset has been retrieved from complete sleep studies at the hospital, which is an examination usually prescribed to patients complaining for severe apnea symptoms, such as loud snoring or excessive daytime sleepiness. Thus, patients characterized with severe apnea-hypopnea syndrome are the majority among the examined subjects.

5. Conclusions

The reported study proves the capacity to accurately determine the Mean Apnea-Hypopnea Duration and the Apnea-Hypopnea Index depending only on EEG recordings. Our method can detect most of the events in the severe apnea group, the group where accurate event detection is vital, while event-by-event detection increases the confidence of the method.

Author Contributions: L.K.: conceptualization, methodology, software, validation, writing—original draft preparation, G.K., N.-A.T., S.A.M. and S.M.P.: writing—review and editing, S.M.P.: conceptualization and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T1EDK-03957_Automatic Pre-Hospital, In-Home, Sleep Apnea Examination).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Sismanoglio – Amalia Fleming General Hospital of Athens on March 16, 2017, with protocol number 05/16.03.2017.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available in PSG-Audio at https://doi.org/10.11922/sciencedb.00345

Acknowledgments: We acknowledge the contribution of the medical team of Sismanogleio General Hospital of Athens—Amfilochiou Anastasia, Kastanakis Emmanouil, Maniou Chrysoula and Tsiafaki, Xanthi—for events annotation and characterization of all polysomnogram recordings.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. McNicholas, W.; et al. Public health and medicolegal implications of sleep apnoea. *Eur. Respir. J.* **2002**, *20*, 1594-1609; DOI:10.1183/09031936.02.00404502.

2. Drager, L.; McEvoy, R.; Barbe, F.; Lorenzi-Filho, G.; Redline, S. Sleep Apnea and Cardiovascular Disease. *Circ.* **2017**, *136*, 1840-1850; DOI:10.1161/CIRCULATIONAHA.117.029400.
3. Terán-Santos, J.; Jimenez-Gomez, A.; Cordero-Guevara,J. The Association between Sleep Apnea and the Risk of Traffic Accidents. *N. Engl. J. Med.* **1999**, *340*, 847-851; DOI:10.1056/NEJM199903183401104.
4. Medical Advisory Secretariat. Polysomnography in patients with obstructive sleep apnea: an evidence-based analysis. *Ont. Health Technol. Assess. Ser.* **2006**, *6*, 1-38.
5. Butler M.; et al. Apnea–Hypopnea Event Duration Predicts Mortality in Men and Women in the Sleep Heart Health Study. *Am. J. Respir. Crit. Care Med.* **2019**, *199*, 903-912; DOI:10.1164/rccm.201804-0758OC.
6. Wu, H.; Zhan, X.; Zhao M.; Wei, Y. Mean apnea–hypopnea duration (but not apnea–hypopnea index) is associated with worse hypertension in patients with obstructive sleep apnea. *Med.* **2016**, *95*, 5493; DOI:10.1097/MD.0000000000005493.
7. Urtnasan, E.; Park, J.; Joo, E.; Lee, K. Automated Detection of Obstructive Sleep Apnea Events from a Single-Lead Electrocardiogram Using a Convolutional Neural Network. *J. Med. Syst.* **2018**, 42; DOI:10.1007/s10916-018-0963-0.
8. Cheng, M.; Sori, W. J.; Jiang, F.; Khan, A.; Liu, S. Recurrent Neural Network Based Classification of ECG Signal Features for Obstruction of Sleep Apnea Detection. *Conf. Proc. IEEE Eng. Emb. Ubi. Comp.* **2017**, 199-202; DOI:10.1109/CSE-EUC.2017.220.
9. Park, J.; Lee, H.; Lee, J.; Urtnasan, E.; Kim, H.; Lee, K. Automatic classification of apnea/hypopnea events through sleep/wake states and severity of SDB from a pulse oximeter. *Physiol. Meas.* **2015**, *36*, 2009-2025; DOI:10.1088/0967-3334/36/9/2009.
10. Choi, S.; et al. Real-time apnea-hypopnea event detection during sleep by convolutional neural networks. *Comput*. Biol. Med. **2018**, *100*, 123-131; DOI:10.1016/j.compbiomed.2018.06.028.
11. Ferrer-Lluis, I.; Castillo-Escario, Y.; Montserrat J. M.; Jané R. Analysis of Smartphone Triaxial Accelerometry for Monitoring Sleep-Disordered Breathing and Sleep Position at Home. *IEEE* Access. **2020**, *8*, 71231-71244; DOI:10.1109/ACCESS.2020.2987488.
12. Beniczky, S.; Schomer, D. Electroencephalography: basic biophysical and technological aspects important for clinical applications. *Epileptic Disorders.* **2020**, *22*, 697-715; DOI:10.1684/epd.2020.1217.
13. Davidson, P. R.; Jones, R.D.; Peiris, M. T. EEG-based lapse detection with high temporal resolution. *IEEE. Trans. Biomed. Eng.* **2007**, *54*, 832–839; DOI:10.1109/TBME.2007.893452.
14. Korkalainen, H.; *et* al. Accurate Deep Learning-Based Sleep Staging in a Clinical Population with Suspected Obstructive Sleep Apnea. *IEEE J. Biomed. Health Inform.* **2019**, 1-1; DOI:10.1109/JBHI.2019.2951346.
15. Zhou, J.; Wu, X.; Zeng, W. Automatic detection of sleep apnea based on EEG detrended fluctuation analysis and support vector machine. *J. Clin. Monit. Comput.* **2015**, *29*, 767-772; DOI:10.1007/s10877-015-9664-0.
16. Vimala, V.; Ramar, K.; Ettappan, M. An Intelligent Sleep Apnea Classification System Based on EEG Signals. *J. Med. Syst.* **2019**, 43; DOI:10.1007/s10916-018-1146-8.
17. Saha, S.; Bhattacharjee, A.; Fattah, S. Automatic detection of sleep apnea events based on inter-band energy ratio obtained from multiband EEG signal. *Health. Technol. Lett.* **2019**, *6*, 82-86; DOI:10.1049/htl.2018.5101.
18. Wang, Y.; Ji, S.; Yang, T.; Wang, X.; Wang, H.; Zhao, X. An Efficient Method to Detect Sleep Hypopnea- Apnea Events Based on EEG Signals. *IEEE Access.* **2021**, *9*, 641-650; DOI:10.1109/ACCESS.2020.3038486.
19. Korompili, G.; Amfilochiou, A.; Kokkalas, L.; et al. PSG-Audio, a scored polysomnography dataset with simultaneous audio recordings for sleep apnea studies. *Sci. Data.* **2021**, *8*, 197; DOI:10.1038/s41597-021-00977-w.
20. Berry, R.; et al. AASM Scoring Manual Updates for 2017 (Version 2.4). *J. Clin. Sleep Med.* **2017**, *13*, 665-666; DOI:10.5664/jcsm.6576.
21. Panigrahi, S.; Behera, H. S. Effect of normalization techniques on univariate time series forecasting using evolutionary higher order neural network. *Int. J. Eng. Adv. Technol*. **2018**, *3*, 280-285.
22. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *ArXiv.* **2014**, 1409-1556.
23. Choi, K.; Fazekas, G.; Sandler, M.; Cho, K. Convolutional recurrent neural networks for music classification. *Conf. Proc. IEEE Ac. Sp. and Sign. Proc.* **2017**, 2392-2396; DOI:10.1109/ICASSP.2017.7952585.
24. Sherstinsky, A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D.* **2020**, *404*, 132306; DOI:10.1016/J.PHYSD.2019.132306.
25. Shamim-Uzzaman, Q.; Singh, S.; Chowdhuri, S. Hypopnea definitions, determinants and dilemmas: a focused review. *Sleep Sci. and Prac.* **2018**, *2*, 1-12; DOI:10.1186/s41606-018-0023-1.