Comparison of Machine Learning and Deep Learning Methods Based on Recurrence Analysis for Obstructive Sleep Apnea Detection

Daniele Padovano[1], Arturo Martinez-Rodrigo[1], Jose M Pastor[1], Jose J Rieta[2], Raul Alcaraz[1]

[1] Research Group in Electronic, Biomedical and Telecomm. Eng., Univ. of Castilla-La Mancha, Spain
[2] BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Spain

Abstract

Obstructive sleep apnea (OSA) is a substantial health risk often associated with severe cardiovascular diseases, yet it frequently goes undiagnosed by the prohibitive cost of gold standard's polysomnography. This study aims to improve OSA detection by comparing traditional machine learning (ML) methods with a modern deep learning (DL) approach based on AlexNet, specifically those utilizing recurrence quantification analysis (RQA) and general recurrence plots (GRPs) in heart rate variability (HRV). Publicly available PhysioNet databases were used, with ECG recordings divided into one-minute intervals. Following TRIPOD guidelines, the results show that the employed DL model outperforms traditional ML models in terms of overall accuracy. More importantly, the AlexNet-based model achieves a better balance between sensitivity and specificity compared to standard classifiers. This study emphasizes DL's potential in enhancing OSA detection through HRV analysis with RQA and GRPs, advancing research in this crucial healthcare area.

1. Introduction

Obstructive sleep apnea (OSA) is a sleep disorder characterized by repeated breathing interruptions during sleep [1]. Its prevalence varies widely, estimated between 9% and 38% in the general population [2]. OSA can lead to daytime drowsiness, impaired work performance, psychological issues, and severe accidents [3]. Furthermore, OSA is often underdiagnosed, especially in individuals with cardiovascular diseases (CVD), with reported rates exceeding 90%[4]. Given the strong correlation between OSA and CVD [5], the early detection of this condition is crucial [6].

Polysomnography (PSG) is the current gold standard for OSA diagnosis but is resource-intensive and highly uncomfortable [7]. PSG also requires clinical experts, limiting its availability and exacerbating the OSA detection challenge [8].

Consequently, researchers have explored alternatives to address PSG's limitations [9]. Particularly, the heart rate variability (HRV) analysis has emerged as a promising tool [6], reflecting autonomic nervous system activity, which is related to respiratory control [10]. In this regard, HRV extracted from single-lead electrocardiograms (ECG) has shown superiority over other kinds of measures[11, 12].

In the context of HRV-based methods, various works have dealt with OSA detection from several perspectives [13–19]. More precisely, recurrence plots (RP), a complexity-related tool [20], have been integrated with traditional machine learning (ML) classifiers [21, 22]. However, RP-based features have limitations, such as the distance threshold, focusing on a specific level of recurrence. The general recurrence plot (GRP) instead offers a holistic perspective across all recurrence levels, becoming the main interest of the present study [20].

Modern deep learning (DL) approaches have shown advancements in accurately distinguishing between normal and apneic episodes [6]. However, research exploring the synergy of GRP and DL in OSA detection remains scarce, especially when compared to single-lead ECG-based models. This scarcity underscores the opportunity for a more in-depth analysis into the capabilities of GRP and DL-based methodologies. Thus, in this study, we present an experimental proposal focused on these novel aspects, emphasizing their advantages over traditional ML methods.

2. Methodology

2.1. Databases

This study has employed three publicly available databases obtained from PhysioNet's repository: the CinC Challenge 2000 (Apnea-ECG) [23], the MIT Polysomnographic Database (MIT-BIH)[24], and the St. Vincent's University Hospital/University College of Dublin (UCD-DB) [25]. These databases contain multiple ECG recordings, each with different apneic annotations.

The Apnea-ECG database comprises 70 ECG recordings, ranging from 7 to 9 hours in length, collected from

 ISSN: 2325-887X DOI: 10.22489/CinC.2023.300

30 male and 5 female subjects aged between 27 and 63 years. This database includes minute-by-minute annotations by clinical experts, categorizing each minute of the ECG recording as either an apneic episode (A-labeled) or a normal episode (N-labeled).

The MIT-BIH database comprises 18 PSG recordings with durations ranging from 2 to 7 hours, obtained from 16 male subjects aged between 32 and 56 years. Similar to the Apnea-ECG database, clinical experts provided annotations, but these were assigned every 30 seconds.

Lastly, the UCD-DB consists of 25 full overnight PSG recordings from 21 male and 4 female subjects aged between 28 and 68 years. Real-time single-lead ECG annotations were made following the Rechtschaffen and Kales rules, identifying various cardiorespiratory events in addition to apnea episodes.

To ensure uniformity in annotation criteria, all databases were adapted to the time resolution of the less detailed one, which is Apnea-ECG. Specifically, the MIT-BIH database was re-labeled by aligning annotations with the original labels every two blocks of 30 seconds, while the UCB-DB was re-labeled on a minute-by-minute basis. This approach harmonized the annotations across all databases.

2.2. Signal processing

The ECG recordings were resampled to 500 Hz for improved R-peak detection. To reduce noise, a second-order Chebyshev filter with a 100 Hz cutoff was applied, and a 0.5 Hz cutoff filtered low-frequency interference and base-line wandering. Both filters maintained the signal's phase and amplitude, preserving core features. The Pan-Tompkins algorithm was applied to detect R-peaks, and a sliding R-peak correction window accommodated missed peaks. ECG recordings were segmented into 1-minute intervals for HRV analysis. Eventually, a manual signal screening was performed to remove noisy segments, leaving more than 40,000 ECG segments, or approximately 680 hours of data.

2.3. Recurrence Analysis

The RP is a non-linear data analysis technique that visualizes the instances when a dynamical system's trajectory revisits the same phase space area [26]. It creates a binary matrix with black and white dots based on the pairwise distance between states in the phase space trajectory, compared to a specified threshold (ϵ). The time series ($s(n) = s(1), s(2), \ldots, s(N)$) is embedded into an m-dimensional space using Taken's time delay theorem [27], resulting in state vectors $\vec{x_i}$. Then, the recurrence matrix or RP ($R(i,j)$) can be defined as:

$$R(i,j) = \begin{cases} 1 & \text{if } \|\vec{x_i} - \vec{x_j}\| \leq \epsilon, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

The RP is visually represented with black dots for recurrent points $\|\vec{x_i} - \vec{x_j}\| \leq \epsilon$ and white dots otherwise. Features were extracted from RP ($\epsilon = 0.5$) constituting recurrence quantification analysis (RQA) [28], including recurrence rate, determinism, Shannon entropy, average diagonal line length, and divergence [29].

Without applying a threshold, the distance matrix, also referred to as the general recurrence plot (GRP), offers a continuous representation of recurrence. False nearest neighbors (FNN) and average mutual information (AMI) algorithms [30] were used to reconstruct the HRV phase space with $m = 3$ and $\tau = 2$.

2.4. Machine Learning Classifiers

On the one hand, multiple traditional ML algorithms were employed, such as support vector machine (SVM), k-nearest neighbors (KNN), and decision tree (TREE). SVM used a Gaussian radial basis function kernel, KNN utilized the Euclidean distance method, and TREE followed CART principles. All models used the same set of features extracted from the HRV's RP.

On the other hand, the AlexNet architecture was employed, accepting GRP representations of $227 \times 227 \times 3$ pixels. The pre-trained network parameters included learning rate, epoch, mini-batch size, and the adaptive moment estimation (ADAM) optimizer [6]. Training parameters included a learning rate of 0.001, 10 epochs, and a mini-batch size of 128.

2.5. Performance Assessment

Models were evaluated using accuracy (Ac), sensitivity (Se), and specificity (Sp). Testing followed a standardized framework, training models on balanced MIT-BIH and UCD-DB databases and testing on Apnea-ECG. This approach enabled obtaining results with reduced bias and adhered to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines [31].

3. Results

In the presented results (Table 1), the performance metrics for various classification models were assessed. The SVM exhibited an Ac of 67.13%, with a Se of 22.95% and Sp of 86.6%. The KNN model achieved an accuracy of 62.29%, with Se and Sp values of 37.21% and 73.37%, respectively. The TREE model yielded an Ac of 59.20%, with a Se of 43.23% and specificity of 66.27%. The DL model, AlexNet, displayed the highest Ac among the models, recording 71.30%, along with a Se of 58.95% and Sp of 76.76%.

Table 1: Table of results.

Model	Ac (%)	Se (%)	Sp (%)
SVM	67.13	22.95	86.6
KNN	62.29	37.21	73.37
TREE	59.20	43.23	66.27
AlexNet	71.30	58.95	76.76

4. Discussion

Among the conventional ML models, SVM demonstrated a considerable Sp but struggled with Se. Furthermore, KNN achieved a more balanced trade-off between Se and Sp, while TREE excelled in Sp but lagged behind in Se. In contrast, AlexNet have outperformed all other models in terms of Ac while maintaining a good balance between Se and Sp.

The above observations imply that the tested DL-based model exhibits strong performance for accurate OSA detection through the analysis of HRV using GRP. The combination of HRV features and recurrence analysis offers a distinctive, with AlexNet emerging as the leader in both Ac and Se, positioning it as a valuable tool for early OSA detection. Generally, ML-based classifiers may struggle to achieve a balance between Se and Sp values compared to DL-based classifiers due to their reliance on the chosen feature extraction strategy. While ML models require human intervention in feature extraction, DL models have the inherent capability to extract previously unidentified features from input data, thus enhancing classification Ac by prioritizing the features that contribute most to the input data.

It is important to highlight that these results have undergone rigorous validation procedures in accordance with the TRIPOD guidelines [31]. Diverging from overly optimistic outcomes obtained with cross-validation frameworks, the models in the present study underwent training on two distinct databases and were subsequently tested on a third one, ensuring robust and clinically relevant performance assessment. Nevertheless, further research and validation efforts are imperative to fully harness the potential of these models in real-world clinical applications.

5. Conclusions

The present study emphasizes the potential of pretrained deep learning architectures like AlexNet, in detecting obstructive sleep apnea (OSA) through heart rate variability analysis with general recurrence plots. Rigorous validation and superior sensitivity highlights their value in improving OSA detection in episodes of one minute-length, addressing a significant public health concern.

Acknowledgments

This research has received financial support from public grants PID2021-00X128525-IV0 and PID2021-123804OB-I00 of the Spanish Government 10.13039/501100011033 jointly with the European Regional Development Fund (EU), SBPLY/17/180501/000411 and SBPLY/21/180501/000186 from Junta de Comunidades de Castilla-La Mancha, and AICO/2021/286 from Generalitat Valenciana. Moreover, Daniele Padovano holds a predoctoral scholarship 2022-PRED-20642, which is co-financed by the operating program of European Social Fund (ESF) 2014-2020 of Castilla-La Mancha.

References

[1] Rundo JV. Obstructive sleep apnea basics. Cleveland Clinic Journal of Medicine September 2019;86(9 Suppl 1).

[2] Senaratna CV, Perret JL, Lodge CJ, Lowe AJ, Campbell BE, Matheson MC, Hamilton GS, Dharmage SC. Prevalence of obstructive sleep apnea in the general population: A systematic review. Sleep Medicine Reviews Aug 2017; 34:70–81.

[3] Veasey SC, Rosen IM. Obstructive Sleep Apnea in Adults. New England Journal of Medicine April 2019; 380(15):1442–1449.

[4] Yeghiazarians Y, Jneid H, Tietjens JR, Redline S, Brown DL, El-Sherif N, Mehra R, Bozkurt B, Ndumele CE, Somers VK, null n. Obstructive Sleep Apnea and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation July 2021;144(3):e56–e67.

[5] Organization WH, et al. Noncommunicable diseases country profiles 2018. Publications on NCDs 2018;.

[6] Bahrami M, Forouzanfar M. Sleep Apnea Detection from Single-Lead ECG: A Comprehensive Analysis of Machine Learning and Deep Learning Algorithms. IEEE Transactions on Instrumentation and Measurement 2022;1–1.

[7] Bsoul M, Minn H, Tamil L. Apnea MedAssist: real-time sleep apnea monitor using single-lead ECG. IEEE Transactions on Information Technology in Biomedicine 2011; 15(3):416–427.

[8] Osman AM, Carter SG, Carberry JC, Eckert DJ. Obstructive sleep apnea: current perspectives. Nature and Science of Sleep January 2018;10:21–34.

[9] Ramachandran A, Karuppiah A. A Survey on Recent Advances in Machine Learning Based Sleep Apnea Detection Systems. Healthcare July 2021;9(7):914.

[10] Chandra S, Sica AL, Wang J, Lakticova V, Greenberg HE. Respiratory effort-related arousals contribute to sympathetic modulation of heart rate variability. Sleep Breathing Dec 2013;17(4):1193–200.

[11] Chang HY, Yeh CY, Lee CT, Lin CC. A Sleep Apnea Detection System Based on a One-Dimensional Deep Convolution Neural Network Model Using Single-Lead Electrocardiogram. Sensors January 2020;20(15):4157.

[12] Neha, Sardana HK, Kanwade R, Tewary S. Arrhythmia detection and classification using ECG and PPG techniques: a review. Physical and Engineering Sciences in Medicine December 2021;44(4):1027–1048.

[13] Yang Q, Zou L, Wei K, Liu G. Obstructive sleep apnea detection from single-lead electrocardiogram signals using one-dimensional squeeze-and-excitation residual group network. Computers in Biology and Medicine January 2022;140:105124.

[14] Taghizadegan Y, Jafarnia Dabanloo N, Maghooli K, Sheikhani A. Obstructive sleep apnea event prediction using recurrence plots and convolutional neural networks (RP-CNNs) from polysomnographic signals. Biomedical Signal Processing and Control August 2021;69:102928.

[15] Zhao Dong, Xiang Li, Wei Chen. Frequency network analysis of heart rate variability for obstructive apnea patient detection. Ieee Journal of Biomedical and Health Informatics November 2018;22(6):1895–1905.

[16] Lin CY, Wang YW, Setiawan F, Trang NTH, Lin CW. Sleep apnea classification algorithm development using a machine-learning framework and bag-of-features derived from electrocardiogram spectrograms. Journal of Clinical Medicine 2021;11(1):192.

[17] Sheta A, Turabieh H, Thaher T, Too J, Mafarja M, Hossain MS, Surani SR. Diagnosis of Obstructive Sleep Apnea from ECG Signals Using Machine Learning and Deep Learning Classifiers. Applied Sciences January 2021;11(14):6622.

[18] Tang L, Liu G. The novel approach of temporal dependency complexity analysis of heart rate variability in obstructive sleep apnea. Computers in Biology and Medicine August 2021;135:104632.

[19] Zarei A, Asl BM. Automatic detection of obstructive sleep apnea using wavelet transform and entropy-based features from single-lead ECG signal. IEEE Journal of Biomedical and Health Informatics May 2019;23(3):1011–1021.

[20] Marwan N, Wessel N, Meyerfeldt U, Schirdewan A, Kurths J. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. American Physical Society August 2002;66(2):026702.

[21] JeyaJothi ES, Anitha J, Rani S, Tiwari B. A Comprehensive Review: Computational Models for Obstructive Sleep Apnea Detection in Biomedical Applications. BioMed Research International February 2022;2022:e7242667.

[22] Sofia Martin-Gonzalez, Juan L. Navarro-Mesa, Gabriel Julia-Serda, Jan F. Kraemer, Niels Wessel, Antonio G. Ravelo-Garcia. Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder. Computers in Biology and Medicine December 2017;91:47–58.

[23] Penzel T, Moody GB, Mark RG, Goldberger AL, Peter JH. Apnea-ECG database, 2000. URL `https://physionet.org/content/apnea-ecg/`.

[24] Ichimaru Y, Moody GB. MIT-BIH polysomnographic database, 1992. URL `https://physionet.org/content/slpdb/`.

[25] McNicholas W, Doherty L, Ryan S, Garvey J, Boyle P, Chua E. St. Vincent's University Hospital / University College Dublin Sleep Apnea Database, 2004. URL `https://physionet.org/content/ucddb/`.

[26] Marwan N, Carmen Romano M, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. Physics Reports January 2007;438(5):237–329.

[27] Noakes L. The takens embedding theorem. International Journal of Bifurcation and Chaos December 1991; 01(04):867–872.

[28] Martín-González S, Navarro-Mesa JL, Juliá-Serdá G, Ramírez-Ávila GM, Ravelo-García AG. Improving the understanding of sleep apnea characterization using recurrence quantification analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold. Public Library of Science April 2018;13(4):e0194462.

[29] Nguyen HD, Wilkins BA, Cheng Q, Benjamin BA. An Online Sleep Apnea Detection Method Based on Recurrence Quantification Analysis. IEEE Journal of Biomedical and Health Informatics July 2014;18(4):1285–1293.

[30] Wallot S, Mønster D. Calculation of Average Mutual Information (AMI) and False-Nearest Neighbors (FNN) for the Estimation of Embedding Parameters of Multidimensional Time Series in Matlab. Frontiers in Psychology 2018;9.

[31] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. British Journal of Surgery 2015;102(3):148–158.

Address for correspondence:

Daniele Padovano
Campus Universitario S/N. 16071 – Cuenca (Spain)
daniele.padovano@uclm.es