

Detection of Obstructive Sleep Apnoea by ECG signals using Deep Learning Architectures

Haifa Almutairi, Ghulam Mubashar Hassan, Amitava Datta

Department of Computer Science & Software Engineering

The University of Western Australia

haifa.almutairi@research.uwa.edu.au, ghulam.hassan@uwa.edu.au, amitava.datta@uwa.edu.au

Abstract—Obstructive Sleep Apnoea (OSA) is a breathing disorder that happens during sleep and general anaesthesia. This disorder can affect human life considerably. Early detection of OSA can protect human health from different diseases including cardiovascular diseases which may lead to sudden death. OSA is examined by physicians using Electrocardiography (ECG) signals, Electromyogram (EMG), Electroencephalogram (EEG), Electrooculography (EOG) and oxygen saturation. Previous studies of detecting OSA are focused on using feature engineering where a specific number of features from ECG signals are selected as an input to the machine learning model. In this study, we focus on detecting OSA from ECG signals where our proposed machine learning methods automatically extract the input as features from ECG signals. We proposed three architectures of deep learning approaches in this study: CNN, CNN with LSTM and CNN with GRU. These architectures utilized consecutive R interval and QRS complex amplitudes as inputs. Thirty-five recordings from PhysioNet Apnea-ECG database have been used to evaluate our models. Experimental results show that our architecture of CNN with LSTM performed best for OSA detection. The average classification accuracy, sensitivity and specificity achieved in this study are 89.11%, 89.91% and 87.78% respectively.

Index Terms—Obstructive Sleep Apnoea, ECG, Deep Learning, Convolutional Neural Networks, Long Short Term Memory, Gated Recurrent Unit.

I. INTRODUCTION

Obstructive Sleep Apnoea (OSA) is a common chronic sleep syndrome which causes a person to breath abnormally. It occurs when the upper airflow is blocked for 5 to 10 seconds leading to failure to breath. However, when the upper airflow is just partially blocked, it is called Hypopnea [1]. The common symptoms of OSA are exhaustion during the day, snoring, morning headache, fatigue and impaired concentration. The prolonged existence of these symptoms can cause many health issues including high blood pressure, heart disease, irregular heartbeats called *arrhythmias*, stroke and diabetes. According to a study conducted by World Health Organization (WHO) in 2007, more than 100 million people were suffering from this syndrome around the world. In 2018, this number increased to 1 billion people [2]. In order to avoid these complications and diseases, early detection of OSA plays an important role.

The standard clinical procedure used for the diagnosis of OSA is called *polysomnography* (PSG) [3]. This procedure is inconvenient as it needs the patient to sleep in a laboratory or hospital for two to three nights with a large number of body electrodes and wires to record signals. Furthermore, the procedure needs expensive equipment, physicians, beds and nurses, which makes it an expensive procedure. The alternative solution is wearable sensors that can be used at home to record the physiological signals like Electrocardiography (ECG), Electromyogram (EMG) [4] and Electroencephalogram (EEG) [5].

Electrocardiography (ECG) is considered as a low-cost tool for detection of OSA. Different studies proved that ECG signals are useful tool to detect OSA an early stage [6], [7]. ECG records signals that represent the depolarization and re-polarization from the chambers of the heart. Each heartbeat has three primary waves: P, QRS complex and T, that represent specific intervals and measures in regular ECG signals [8]. When these distinct waves deviate from the habitual patterns associated with regular cardiac behavior, it is indicative of the problems which may lead to OSA and other heart diseases. These signals are then analyzed by an expert physician to detect OSA. However, it is a time-consuming process which requires a high-level skills. Therefore, automatic detection of OSA from ECG signals plays an important role in early OSA diagnosis and will help the physicians to make better decisions.

In 2000, PhysioNet website held a challenge to detect OSA from ECG signals. Different techniques were proposed to solve the problem [9], [10]. The dataset was made available publicly later for researchers to validate their proposed models to detect OSA and became a benchmark dataset for OSA detection.

Bsoul et al. [7] suggested to use Support Vector Machine (SVM) to detect OSA. The technique used ECG and EDR (ECG-Derived Respiration) signals and achieved the accuracy of 89%. Xie et al. [11] suggested using AdaBoost method with Reduced Error Pruning Tree (REPTree) to detect OSA. They used Time and Spectral-domain features from ECG and SPO2 signals. The accuracy of this model was reported to be 77.74%. Nguyen et al. [12] proposed a statistical feature

called Recurrence Quantification Analysis (RQA). They extracted features from Heart Rate Variability (HRV) and deployed Support Vector Machine and Neural networks to detect OSA. The accuracy of this approach was reported to be 85.26%. Hassan et al. [13] proposed to decompose ECG signals into energy bands by using Tunable-Q Factor Wavelet Transform (TQFWT) which was further processed using asymmetric normal inverse Gaussian model. The features obtained were then fed into Adaboost classifier for OSA detection. The accuracy was reported to be dependent on the signal sampling rate and varied from 80 to 87%. Li et al. [14] proposed SVM and Hidden Markov model (HMM) to detect OSA. They used Sparse autoencoder feature and accuracy was reported to be 84.7%. Sharma et al. [15] used least squares support vector machine to detect OSA. They used a filter bank to select specific features from the signals as inputs and reported the accuracy to be 90.11%.

All the discussed approaches primarily rely upon prior knowledge of ECG-based signals and OSA to design and select features which is a limitation to completely automate the OSA detection process from ECG signals. Recently, deep learning techniques are found to be performing well in extracting features automatically from the input signals for clinical and nonclinical applications based on ECG and other physiological signals [14].

Deep learning is a machine learning process that uses neural networks with multi-layers to learn to detect and classify from input data [16]. A study by Dey et al. [17] used deep learning methods to detect OSA from ECG signals. Their model takes raw input signals which are fed into 1D-CNN model to detect OSA. The study used different ratios for training and testing partitions to evaluate the performance of their model. For each of the different ratios, the accuracy changed considerably which demonstrated the non-robustness of the model. This limitation can be contributed to the reason that the research did not use K-fold cross validation to evaluate their model.

In this paper, we propose deep learning based models for OSA detection with following details:

- Implementing three different architectures of Convolution Neural Networks (CNNs) to detect OSA: CNN, CNN with two consecutive Long Short Term Memory (LSTM) layers and CNN with two consecutive Gated Recurrent Unit (GRU) layers.
- Using 10 fold cross-validation to verify the accuracy of the model.
- Comparing the results of proposed deep learning techniques with other state of the art techniques.

The remaining part of this paper is organized as follows. Section II introduces the proposed approach of OSA detection including the dataset, data preprocessing, K-fold cross-validation method and deep learning model architectures. Section III presents the results and the discussion. The research is concluded in Section IV.

II. PROPOSED APPROACH

This section focuses on describing the dataset, preprocessing technique used in this study and the proposed deep learning models employed for detecting OSA from ECG signals.

A. Dataset

In this study, we have used PhysioNet Apnea-ECG database [18]. It contains 70 ECG recordings which were recorded from 32 subjects (25 males and 7 females). The duration of the recordings ranges from 6 to 8 hours (mean: 492 ± 32 min). The recordings were sampled at 100 Hz with 16-bit resolution, with one sample bit representing 5μ V. Each recording was segmented into 1-minute epochs by a human expert for labelling them as healthy or OSA. The 70 recordings had been split into two sets: *released* set for training and *withhold* set for testing. In this study, we used the released dataset which contains 35 recordings and the number of segments approximately equalled 16993 minutes. The total number of OSA and healthy segments are 6513 and 10480 respectively. Figure 1 shows ECG signals of a normal and an OSA subjects.

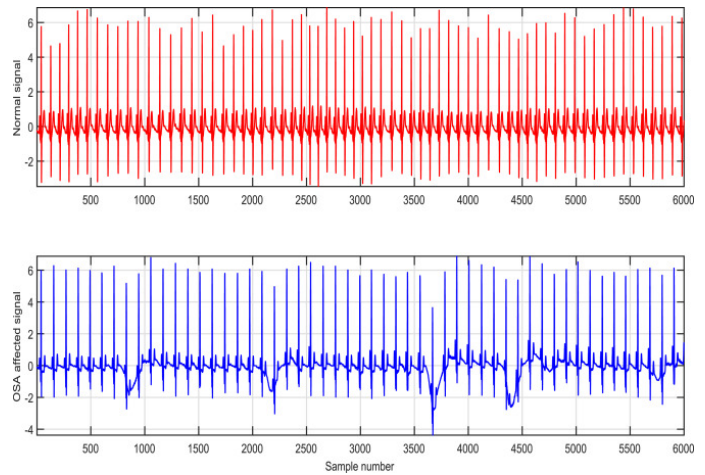


Fig. 1. a normal (in red) and an OSA (in blue) subjects. [15].

B. Preprocessing

In the pre-processing process, QRS complex amplitudes and the intervals between successive R were calculated. We computed QRS time and found the R-peaks by using Real-time QRS detector algorithm [19]. We eliminated the false intervals by using cubic spline interpolation. To attenuate the

high frequencies, z-score was used to normalize all the signals. Eq. (1) presents the z-score as [20]:

$$z - score = \frac{(signal - E\{signal\})}{\alpha signal} \quad (1)$$

Where E is the mean signal and α is the standard deviation of the $signal$.

The total number of segments after preprocessing were 16461 consisting of 6211 OSA and 10250 healthy segments.

C. Deep Learning model

In deep learning model, we used K-fold cross-validation method and implemented Convolution Neural Networks (CNNs) in three different architectures. An overview of these architectures is shown in Figure 2.

- 1) **K-fold cross-validation method** : In general, our model has two statistical inputs: R-R intervals and QRS amplitudes. We applied k-fold cross-validation on all the data. It is a re-sampling technique used to estimate predictive models by randomly partitioning the dataset into k equal sub-sampling [21]. This simple method generates results having less bias. In this study, we selected k = 10 folds; one fold for testing the model and (k-1=9) folds for training the model.
- 2) **CNN Architecture**: The first deep learning model contained three 1-dimensional convolution layers. These layers used filters (Kernels) that calculated the input features in the form of a feature map. For each convolutional layer, we used kernels of size 3 and different parameter values as listed in Table I. We added Max-pooling layer between two convolutional layer to minimize the number of parameters of the feature map. The max-pooling window calculated the maximum or the average value over the feature map. We added dropout regularization after each max pooling to reduce over-fitting. ReLU was the activation function used with both 1-D convolution layers. The mathematical operation of ReLU used is presented as [22]:

$$Max(0, x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{otherwise } x > 0 \end{cases} \quad (2)$$

where x is the input.

- 3) **CNN with two consecutive Long Short Term Memory (CNN+LSTM) Architecture**: In the second architecture, two consecutive LSTM layers were added to above mentioned CNN architecture. LSTM was introduced to increase the length of term dependencies from the input and handle the gradient vanishing problem. Each LSTM layer contained three gates: *input* gate, *forget* gate and *output* gate. The input gate updates the memory cell according to the new input. The forget gate sets the cell memory to zero and the output gate is

TABLE I
PARAMETERS OF DEEP LEARNING MODEL

Layers	Kernel	Activ
Input	480	
Architecture 1 : CNN		
1D-Convolution	64	ReLU
1D-Max-Pooling	2	
Dropout	0.2	
1D-Convolution	128	ReLU
1D-Max-Pooling	2	
Dropout	0.2	
1D-Convolution	16	ReLU
1D-Max-Pooling	2	
Dropout	0.2	
Architecture 2 : Architecture 1 + LSTM		
LSTM	64	Tanh
Dropout	0.2	
LSTM	128	Tanh
Dropout	0.2	
Architecture 3 : Architecture 1 + GRU		
GRU	64	Tanh
Dropout	0.2	
GRU	128	Tanh
Dropout	0.2	
Dense Layer	64	ReLU
Dense Layer	1	Sigmoid

used to provide activation function for prediction. We added dropout regularization after each LSTM layer to reduce over-fitting. Tanh was the activation function used in both LSTM layers.

- 4) **CNN with two consecutive Gated Recurrent Unit (CNN+GRU) Architecture**: The third architecture is similar to CNN+LSTM architecture, where LSTM is replaced by GRU. Each GRU layer contains two gates: *reset* gate and *update* gate. The reset gate determines how much old information to forget. The update gate works similar to the forget and input gates in LSTM layer and determines what information should forget and what information should pass to the output. We added dropout regularization after each GRU layer to reduce over-fitting. Tanh was the activation function used in both GRU layers.

For all architectures, we added batch normalization layers to improve the speed and performance of our models. Also, we added fully connected layers to give the final probabilities for each class. Sigmoid was the activation function used in fully connected layers. The mathematical operation of sigmoid used is presented as [22].

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3)$$

where z is the output of the previous layer.

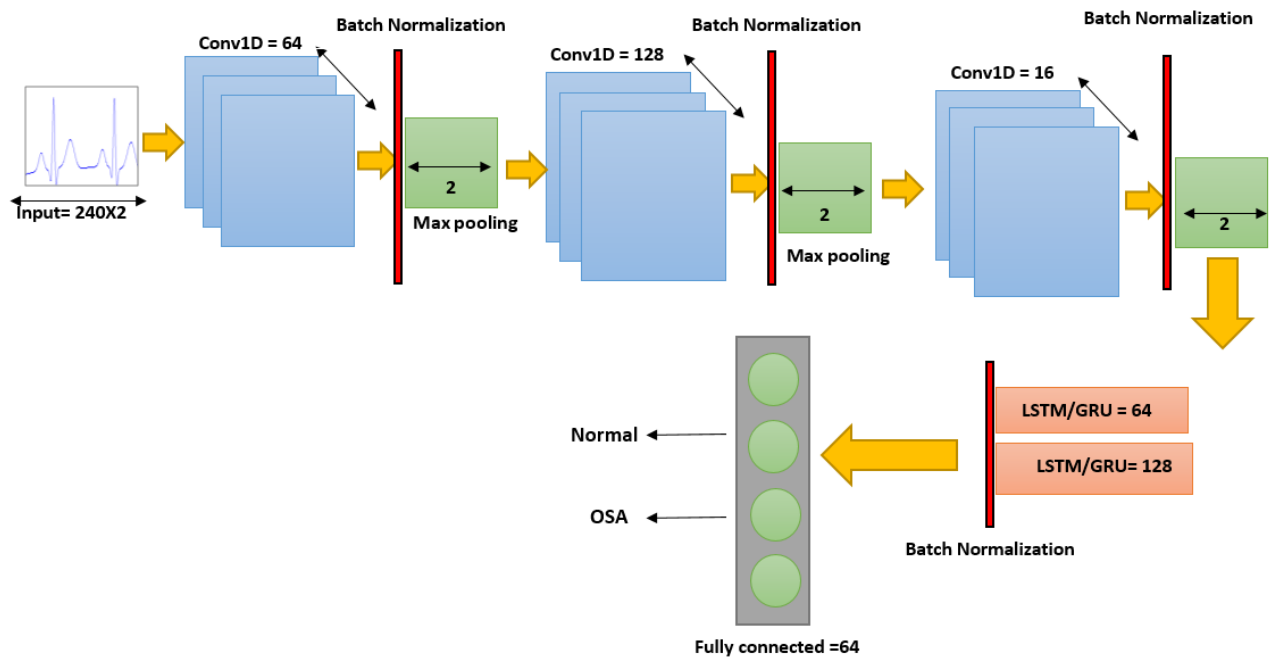


Fig. 2. The architectures of deep learning models.

III. RESULTS AND DISCUSSION

This experiment was performed on a system having an Intel (R) Core (TM) 3.6GHz (i7-7700) processor with 8 GB RAM using Python 3.7 version. Our model used Adam optimization and cross-entropy loss functions to minimize the model parameters.

Three different deep learning architectures were developed to detect OSA for this study. We achieved the best performance using CNN+LSTM architecture. The average classification accuracy was found to be 89.11% with 10 fold cross-validation. The Average classification sensitivity, average classification specificity and F1 measures were found to be 89.91%, 87.87% and 91.41% respectively. Table II shows the experimental results of the three architectures and confusion matrix for each architecture. From the confusion matrix of CNN+LSTM, we can observe that this model can identify correctly 93% of normal and 83% of OSA ECG segments.

Table III presents the comparison of our results of CNN+LSTM architecture with state of the art techniques [12], [13], [15] for the same dataset. The state of the art classification techniques for OSA detection are machine learning methods which rely on feature engineering to extract selected features from ECG signals. This makes the feature selection as a critical step and restricts the machine learning methods to learn from selected features. However, recent literature suggests that deep learning models outperform human selection in extracting features from the data in many applications [23]. Whereas, feature engineering is

laborious task and requires human ingenuity and domain knowledge [24]. Therefore, in this study, we proposed deep learning models to extract the important features from the signals automatically and learns from their patterns. This make the process simple and avoids time consuming feature engineering.

It can be observed from the comparative results presented in Table III that CNN+LSTM architecture performs better than rest of the presented techniques except Sharma et al.'s. [15] technique. The model proposed by Sharma et al. [15] used feature engineering which involved feature extraction, feature selection, feature reduction and feature classification. Whereas, our proposed model learns the discriminant features automatically and the performance is very close to Sharma et al. [15].

IV. CONCLUSION

In this paper, we proposed the OSA detection model based on deep learning. We implemented three architectures of deep learning: CNN, CNN+LSTM and CNN+ GRU. In the study, we focused on extracting features automatically using deep learning techniques rather than feature engineering. The results of this study showed that our proposed method of CNN+LSTM is efficient for OSA detection and the accuracy of our approach was found to be 89.11%. In this study, the used dataset had imbalanced class distribution with 62% of the observations classified as healthy and 38% as OSA. In future work, we will focus on using the large and balanced dataset.

TABLE II
THE EXPERIMENTAL RESULTS OF THE THREE ARCHITECTURES.

10-fold Cross-validation						Confusion matrix			
Classifier	AC(%)	SE(%)	SP(%)	F1(%)	AE	TP	FP	FN	TN
CNN	86.31	89.38	81.58	88.95	45.7	9083	1167	1085	5126
CNN+GRU	88.55	91.34	84.35	90.73	25.5	9248	1002	882	5329
CNN+LSTM	89.11	89.91	87.87	91.41	26.5	9535	715	1076	5135

*AC = Accuracy, SE = Sensitivity, SP = Specificity, AE = Average Epoch

TABLE III
COMPARISON BETWEEN OSA DETECTION PERFORMANCE OF OUR MODEL AND EXISTING WORKS.

Reference	Input(num)	Classifier	AC (%)	SE (%)	SP (%)	K-fold
Nguyen et al. [12].	Feature Engineering (32)	SVM, NN	85.26	86.37	83.47	3-fold
Hassan et al. [13].	Feature Engineering (36)	Adaptive Boosting	87.33	81.99	90.72	10-fold
Sharma et al. [15].	Feature Engineering (12)	LS-SVM	90.11	90.87	88.88	35-fold
Proposed model	Automated feature (480)	CNN+LSTM	89.11	89.91	87.87	10-fold

*AC = Accuracy, SE = Sensitivity, SP = Specificity

REFERENCES

- [1] P. Peppard, T. Young, J. Barnet, M. Palta, E. Hagen, and KM.Hla, "Increased prevalence of sleep-disordered breathing in adults," *American Journal of Epidemiology*, vol. 177, no. 9, pp. 1006–1014, 2013.
- [2] A. Benjafield, K. Valentine, N. Ayas, P. Eastwood, R. Heinzer, M. Ip, S. Patel, P. Peppard, S. Sinha, S. Tufik, C. Nunez, and A. Malhotra, "Global prevalence of obstructive sleep apnea in adults: Estimation using currently available data," in *International Conference, 2018 - San Diego, CA*. American Thoracic Society, 2018, p. B67.
- [3] K. E. Bloch, "Polysomnography: a systematic review," *Technology and Health Care*, vol. 5, no. 4, pp. 285–305, 1997.
- [4] G. Biagetti, P. Crippa, L. Falaschetti, S. Orcioni, and C. Turchetti, "Wireless surface electromyograph and electrocardiograph system on 802.15. 4," *IEEE Transactions on Consumer Electronics*, vol. 62, no. 3, pp. 258–266, 2016.
- [5] W. Rosenberg, T. Chanwimalueang, V. Goverdovsky, D. Looney, D. Sharp, and D. Mandic, "Smart helmet: Wearable multichannel eeg and eeg," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 4, 2016.
- [6] C. Guilleminault, R. Winkle, S. Connolly, K. Melvin, and A. Tilkian, "Cyclical variation of the heart rate in sleep apnoea syndrome: mechanisms, and usefulness of 24 h electrocardiography as a screening technique," *The Lancet*, vol. 323, no. 8369, pp. 126–131, 1984.
- [7] M. Bsoul, H. Minn, and L. Tamil, "Apnea medassist: real-time sleep apnea monitor using single-lead eeg," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 416–427, 2010.
- [8] S. C. S. V. and M. Y. Narasimha, "Ecg signal analysis using wavelet transforms," *Bulg. J. Phys.*, vol. 35, no. 1, pp. 68–77, 2008.
- [9] J. McNames, A. Fraser, and A. Rechtsteiner, "Sleep apnea classification based on frequency of heart rate variability," 2000.
- [10] Z. Shinar, A. Baharav, and S. Akselrod, "Obstructive sleep apnea detection based on electrocardiogram analysis," in *Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163)*. IEEE, 2000, pp. 757–760.
- [11] B. Xie and H. Minn, "Real-time sleep apnea detection by classifier combination," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 469–477, 2012.
- [12] H. D. Nguyen, B. A. Wilkins, Q. Cheng, and B. A. Benjamin, "An online sleep apnea detection method based on recurrence quantification analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1285–1293, 2013.
- [13] A. Hassan, "computer-aided obstructive sleep apnea detection using normal inverse gaussian parameters and adaptive boosting," *Biomed. Signal Process. Control.*, vol. 29, no. 1, pp. 22–30, 2016.
- [14] K. Li, W. Pan, Y. Li, Q. Jiang, and G. Liu, "A method to detect sleep apnea based on deep neural network and hidden markov model using single-lead eeg signal," *Neurocomputing*, vol. 294, pp. 94–101, 2018.
- [15] M. Sharma, S. Agarwal, and U. R. Acharya, "Application of an optimal class of antisymmetric wavelet filter banks for obstructive sleep apnea diagnosis using eeg signals," *Computers in Biology and Medicine*, vol. 100, pp. 100–113, 2018.
- [16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [17] D. Dey, S. Chaudhuri, and S. Munshi, "Obstructive sleep apnoea detection using convolutional neural network based deep learning framework," *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 95–100, 2018.
- [18] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger, and J. H. Peter, "The apnea-ecg database," in *Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163)*. IEEE, 2000, pp. 255–258.
- [19] J. Pan and W. Tompkins, "A real-time qrs detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, 1985.
- [20] S. I. Kim and T. S. Suh, *World Congress of Medical Physics and Biomedical Engineering 2006: August 27-September 1, 2006 COEX Seoul, Korea*. Springer Science & Business Media, 2007, vol. 14.
- [21] D. Rodriguez, A. Perez, and A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569–575, 2009.
- [22] J. Xue, J. Li, and Y. Gong, "Restructuring of deep neural network acoustic models with singular value decomposition," in *Interspeech*, 2013, pp. 2365–2369.
- [23] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Information Fusion*, vol. 42, pp. 146–157, 2018.
- [24] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.