A Research Study On The Diagnostic Capabilities Of Deep Learning Regarding Sleep Apnea

Mrs. Minakshi Dhole[1*], Prof. Sachin Kharade[2], Mrs. Palaskar Kavita[3], Dr. Suvarna Nirmal[4]

[1*]Research Scholar, SunRise University Alwar, Rajashthan, Email: minakshi.kharade@gmail.com
[2]HOD, Computer Engineering, HSBPVT's Parikrama Polytechnic, Kashti, Maharashtra, India,
Email: sachindada.kharade@gmail.com
[3]HOD, Medical Electronics Enineering, HSBPVT's Parikrama Polytechnic, Kashti, Maharashtra, India
[4]Principal, HSBPVT's Parikrama Polytechnic, Kashti, Maharashtra, India

***Corresponding Author:** Mrs. Minakshi Dhole
*Research Scholar, SunRise University Alwar, Rajashthan, Email: minakshi.kharade@gmail.com

Abstract
Apnea of the Sleeping Disorders is a breathing disorder in which the patient stops breathing repeatedly for ten seconds or more while sleeping. In this study, we focus on a model or technique that employs deep learning using the CNN (computational neural network) approach. The principal disadvantages of type 1 full-night polysomnography compared to type 4 sleep investigations are the time commitment and the space requirements of a sleep lab. An alternative to costly and bulky polysomnography is a portable and affordable SPO2 sensor-based deep convolutional neural network model for sleep apnea detection. In all, 190,000 samples from 50 patients' SPO2 sensors were used. The accuracy of deep convolutional neural networks for snoring and apnea detection will be around approximately 92.3085% when using a loss rate of 2.3 and a cross-entropy cost function.

Keywords: Sleep Apnea, Computational Neural Network, Apnea Event Detectionand Deep Learning.

Introduction
Type 4 sleep studies, also known as continuous single bio-parameter or dual-bio parameter recording, are the basis for our method. Type 4 investigations need just one or two signal channels, such as oxygen saturation and airflow. Sleep cannot be scored in most type 4 investigations because of a lack of EEG and EMG data. It may be used to diagnose a breathing problem while you sleep. Type 4 sleep studies include other alternatives, such as the use of oxygen saturation and tracheal sound through acoustic sensor. The fundamental goal of training a deep convolutional neural network model for sleep apnea diagnosis is to acquire model parameters or features. The model's effectiveness may be determined whether or not a given signal sample comprises a sleep apnea episode depends on the parameters it haslearnt. For sleep apnea diagnosis, it is a difficult challenge to understand the connection between the interpretability of the various parameters and the performance of the deep convolutional model. While many convolutional neural networks have found utility in the image classification space, only a few have found use in the detection of audio or other signals. Because of the underlying mechanism of the CNN model, the creation andof a good convolutional neural network model is akin to the black box and requires a large number of trials for the Testing. In the deep learning model, the representation of the sensor signals is given a lot of weight. The amplitude of the sensors' wave data is interpreted as the digit vector figures in the proposed model, which may draws inspiration from picture classification. These figures may be transformed into any dimension using matrix vectorization or factorization. The continuous nature of the high frequency respiratory signal waves is the greatest difficulty in sleep apnea identification. The proposed approach aims to bridge the gap between the discrete character of deep learning and the continuous nature of biological signals from sensors. If you want to comprehend and learn the time series nature of the high frequency SPO2 signal, a convolutional neural network model is a viable and adequate approach. In Future results of the five empirical tests may be conducted to validate and evaluate sleep apnea detection demonstrate its ability to give relevant interpretation for the healthcare professional.

Following is a brief summary of the paper's most significant findings:
To determine whether or not sleep apnea occurs, a deep learning method using a convolutional neural network is suggested for use in the identification of the condition.
Different methods of detecting sleep apnea utilising the SPO2 signal will be used for examination with respect to their accuracy and loss rate.

Objective of the study
In this research paper, we will examine and put into practice the most recent, cutting-edge research on using machine learning to diagnose sleep apnea. The implementation discusses the parameters and Multiple methods of Deep Learning.

Review of Literature

We have reviewed the current state of the art in sleep apnea detection using conventional machine learning techniques. Individual biological signs, such as SPO2, ECG, EOG, and EEG, have been the focus of several investigations for the diagnosis of sleep apnea. Most studies focus on using SPO2 and ECG data because they show an increase in heart rate and systolic blood pressure in response to apneic episodes. To diagnose OSA, for instance, feature engineering on SPO2 data yielded "ODI, total time below saturation levels (tsa), and six more features. Several different decision tree (DT) classifiers were used to attain the 93% accuracy. Pulse oximeter measurements are also used in to identify cases of sleep apnea. We used the SPO2 sensor's PPG readings to calculate the subject's respiratory rate and oxygen consumption. When Linear Discriminant Analysis will applied to SPO2 features and PPG features, the best classification result will be (87%). Statistics and time domain SPO2 and PPG characteristics were retrieved around SPO2 dips and averaged per patient in, another work that takes use of PPG measures taken from SPO2 readings. Here, we looked at how SPO2 and PPG characteristics affect OSA detection. A support vector machine (SVM) classifier was trained using three SPO2-based features and two PPG-based features. When SPO2 characteristics were added to the individuals' ages, a classifier was created with an accuracy of 77.7%; PPG features has no effect on the classifier's performance, in contrast to. This study demonstrates that age is a significant confounding variable due to its association with cardiovascular health, and that age alone may provide acceptable detection accuracy for OSA. To not only identify apnea but also determine its severity using just SPO2 information provided at the patient's home, four machine learning algorithms are compared in. Extraction of features, selection of features, and assessment of classifiers were the three components of the procedure. A Fast Correlation Based Filter feature selection method was fed a total of 16 features retrieved from SPO2 across statistical, spectral, and nonlinear domains, in addition to ODI. The highest performance in classifying apnea severity will be an AdaBoost model constructed using linear discriminants as basis classifiers. Mostafa et al.use a Deep Belief Network (DBN) to examine SPO2 signals in two publicly available datasets. According to the results, the trade-off between classifier performance and processing needs is not always worth it, even if accuracy does rise with the number of hidden neurons. Seven characteristics and SVM are used in another research to identify sleep apnea. This study uses a smart pillow to not only identify but also treat apneic episodes. A pulse oximeter-equipped wearable gadget, a mobile phone, and a flexible cushion make up the setup. A smart phone receives the SPO2 signal from a wearable device's pulse oximeter. When the smartphone detects a SPO2 desaturation event, it sends a command to the pillow to be readjusted. The form and height of the adjustable cushion change in response to the user's instructions. The pulse oximeter acts as a closed-loop feedback mechanism between monitoring and remedial measures by continuously monitoring and evaluating the adjustment's impact. Bluetooth allows for connection between the cushion and the wearable device. In, we see a summary of methods for sleep apnea detection based on pulse oximetry readings.Electrocardiograms are also often used as a diagnostic tool for sleep apnea. Hassan et al. evaluate several machine learning classifiers using data from an ECG sensor with a single lead. From the original data, both statistical moment-based and empirical mode decomposition characteristics were derived. Naive Bayes, k-NN, neural network, AdaBoost, Bagging, random forest, extreme learning machine (ELM), discriminant analysis (DA), and limited Boltzmann machine were evaluated after features were extracted. The highest precision (83.77%) was achieved by ELM. In addition, employed a single-lead ECG dataset to identify obstructive sleep apnea. In this investigation, frequency sub-bands were created from segments of ECG signals using dual-tree complex wavelet transform (DTCWT). The DTCWT output was analysed for its capacity to diagnose sleep apnea using three statistical features: variance, skewness, and kurtosis. A precision of 84.4% was reported by LogitBoost. Classifiers like DA, kNN, ANN, ELM, SVM, AdaBoost, and Bagging are also examined. Electrocardiogram (ECG) data have been used for both the identification and classification of sleep apnea.

Classifiers for sleep apnea identification have been successfully constructed using a variety of ECG features, including IHR, HRV, BCG, and CPC. HRV measures have shown promise in improving OSA detection, according to the available literature." Khandoker et al. emphasised how HRV and EDR may be used with a support vector machine classifier to reliably identify apneic episodes. In addition to measuring OSA's relative severity, the SVM is used here. In order to diagnose sleep apnea, de Chazal et al. employ HRV, EDR, and CPC extracted from single lead ECG data. analyses HRV statistics using kNN, QDA, and SVM. Combining CPC features with HRV measures measured in the time domain resulted in the best classification performance (89.8% accuracy). The classifier's algorithm was multi-logistic discrimination. In, the authors detail a method for extracting 24 features of the ECG in the time and frequency domains. Frequency-domain features included normalised power at various frequencies and the vegetative balance index; time-domain features included the mean, median, standard deviation, and mode for each NN interval series. To train decision trees, discriminant analysis, logistic regression, support vector machines, version of KNN, and ensemble learning classifiers, we had to cut down the number of features to a manageable nine by getting rid of unnecessary data. Using sleep questionnaires and electrocardiograms, Seo et al. investigate how to best evaluate sleep quality and stability. From the ECG data, we were able to extract respiratory and CPC characteristics, and we discovered a strong association between AHI and CPC. Some research on interpreting EEG readings as a window into sleep is found in.

Deep Learning Based Solutions

It is becoming more common to employ deep learning methods for sleep apnea diagnosis, either on individual indicators or via sensor/feature fusion. These methods include the Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long-Short Term Memory (LSTM). The success of intelligent solutions relies heavily on the engineering and selection of features,which is particularly true in the biomedical field. Deep learning systems, which consist of layers of neurons, convolution, and pooling, may discover useful characteristics from raw data. To remove the reliance on subjective human knowledge on critical feature engineering issues, we employed unsupervised learning algorithms with sparse auto-encoders to learn features from ECG data. SVM and ANN are used for classification, and then decision fusion and HMM are used to improve classification performance. To automatically score sleep, a single electrooculogram (EOG) signal is employed. Features are extracted and labels are predicted using a DBN with three layers of 500, 200, and 100 neurons. An HMM model is trained using both the predicted labels and the actual labels. The degree of sleep apnea is diagnosed using LSTM-RNN. After feature extraction and selection, several LSTM-RNN combinations are employed for training. We use Deep Neural Networks (1D, 2D CNNs, RNNs, LSTMs, and GRUs) and Recurrent Neural Networks (RNNs) and Long Short-Term Memories (LSTMs) as classifiers in our project.

Proposed Methodology

The electrocardiogram (ECG) is the gold standard for diagnosing apnea. Electrocardiograms are recordings of the heart's electrical activity. It helps us know whether our hearts are functioning correctly or not. In this study, we normalised the whole ECG time series by breaking it up into regular intervals. In order to identify apnea in the allotted time, we turned to CNN-based deep learning models and ensemble learning.

Models using convolutional neural networks (CNN) are proposed:

Wang et al.'s suggested CNN framework
1. In their model for CNN, Sharan et al.
2. Developed by Almutairi et al., the CNN-LSTM network hybrid.
3. CNN based Model

In this work, we usedexisting CNN based models
(i) Wang et al.'s proposed CNN architecture,

Themodels have previously been proposed in this very domain. However, we made some minor changes to the original architecture.

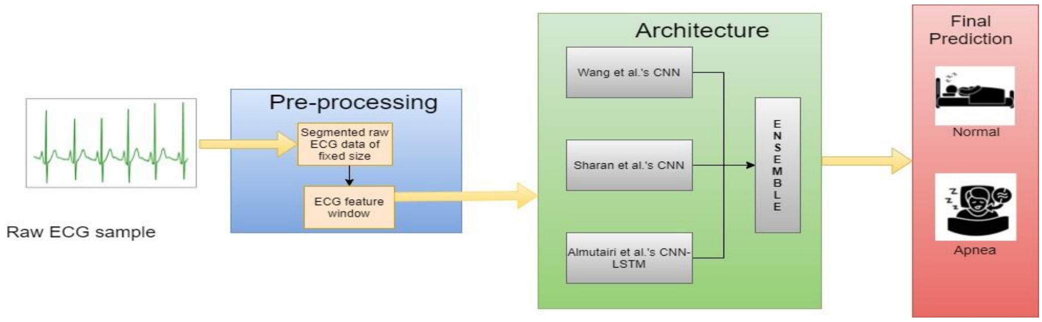

Figure 1 - Proposed model

2.Apnea/Hypopnea Event Detection

Because the ups and downs of the signal will be observed according to minute, second, and hour views, "flat" signals do not indicate that the signal is very flat. Before and after the apnea event, the time segment before and after the occurrence is represented as t1 and t3 respectively, and the amplitude before and after the apnea event is represented as A_1 and A_2 respectively, in which t1 <t2< t3. The baseline amplitude for the apnea is the previous 3 to 6 breaths. If the maximum relative amplitude for the apnea threshold is set to 10 percent, this indicates that the ratio of A_2 to A_1 should be lower than 10 percent. Additionally, an apnea event is considered to have occurred if the amplitude of the event is twenty percent or lower than the data that came before it, as shown in figure 2. According to the most recent version of the AASM manual scoring, the highest threshold for the hypopnea event is less than 70 percent.

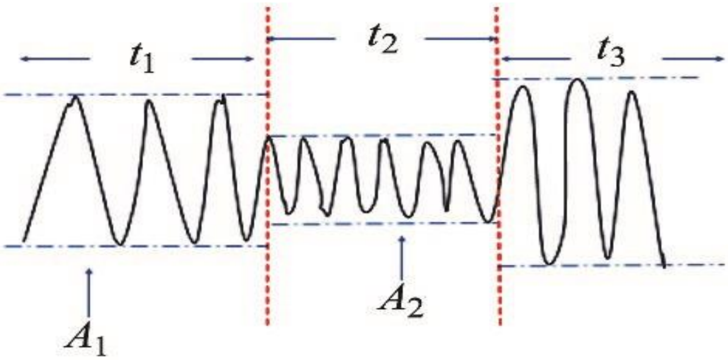

Figure 2 Amplitude Threshold Apnea Event.

As shown in Fig. 3, an event must last for at least 10 seconds before it can be classified as an apnea or hypopnea. *T*equals 10 seconds is indicated in red inside the surrounding area.

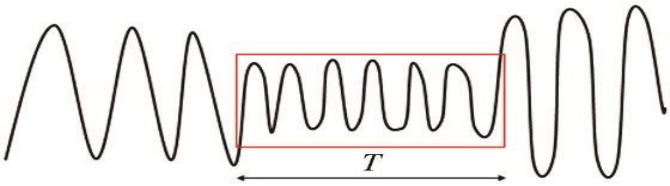

Figure 3 Amplitude Threshold for the Apnea Event

2.3 Deep Convolutional Neural Network

In Fig. 4, we see a simplified representation of a convolutional neural network, in which the first layer serves as input, the last layer serves as output (identifying whether a given sensor segment is apneic or not), and the layers in between serve as hidden layers.

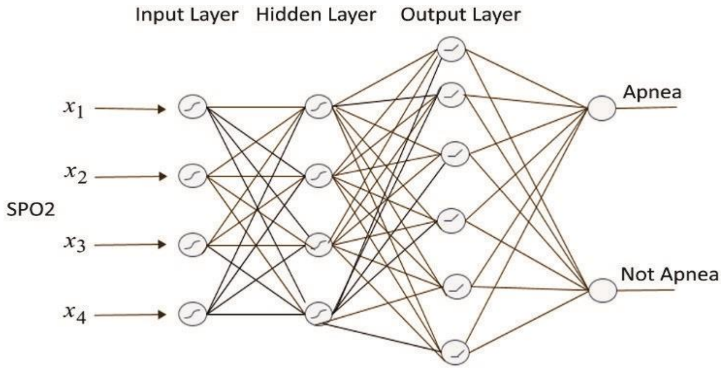

Figure 4 Simple representation of CNN for SPO2

Tab. 1 below provides a synopsis of the ten-layer deep convolutional neural network architecture. In addition, there are three max pooling layers and three convolutional layers, all of which operate in two dimensions. There are 32,614 trainable parameters for the convolutional neural network after running the deep convolutional neural network model for each empirical research, with 794 samples used for training and 3,178 samples used for testing. The SPO2 signal in each sample consists of 484 data points collected at 16 Hz. The suggested model makes use of convolutional layers (both pooling and two-dimensional), as well as a flatten layer, a thick layer, and a dropout layer. First convolutional layer parameters number 320, second convolutional layer parameters number 9,248, and third convolutional layer parameters total 18,496. For the first dense layer, there are 4,160 parameters, and for the third dense layer, there are 390. The purpose of the convolutional layer is to learn the unique qualities of the input signal via feature extraction. Maximum-value pooling layers lower the model's dimensionality while preserving its features and structure. The completely linked layer, which may be employed as a last layer for prediction, is sometimes referred to as the dense layer.

Table 1 Deep Convolutional Neural Network Design

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 20, 20, 32)	320
max_pooling2d_1	(MaxPooling2(None,10,10,32)	0
conv2d_2 (Conv2D)	(None, 8, 8, 32)	9248
max_pooling2d_2	(MaxPooling2(None,4,4,32)	0
conv2d_3 (Conv2D)	(None, 2, 2, 64)	18496
max_pooling2d_3	MaxPooling2 (None, 1, 1, 64)	0
flatten_1 (Flatten)	(None, 64)	0
dense_1 (Dense)	(None, 64)	4160
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 2)	390
Total params: 32,614		
Trainable params: 32,614		
Non-trainable params: 0		
Train on 794 samples, validate on 3178 samples		

3 Performance Evaluation

The proposed deep convolutional neural network model is evaluated in five separate empirical tests, with each research including 10 patients and collecting data from 3,700 to 4,000 samples. There are two sets of visual representations for each empirical study: one for precision and one for error. To determine the rate of decay, the cross entropy is used. The suggested model has an overall performance of 91.3085 percent. In addition to evaluating the models' overall performance, this section compares the suggested convolutional neural network model against the alternatives. Included are analyses of several classifiers like LDA and SVM and representation trees and artificial neural networks trained on SPO2 signals and the suggested convolutional neural network of deep learning.

Five empirical investigations comparing training accuracy are summarised in Fig.5; the orange, red, green, purple, and blue trend lines represent scenarios 1, 2, 3, 4, and 5, respectively. Ten patients of varying ages, BMIs, weights, and AHI indices are included in each scenario. All the lines of best fit showed some degree of wiggle room and modest growth between 1% and 20 iterations. After twenty iterations, the percentage of accuracy begins to stabilise; after forty iterations, the accuracy stabilises around 90%. From this, we learn that between 40 and 100 training cycles provide a stable state for the data.

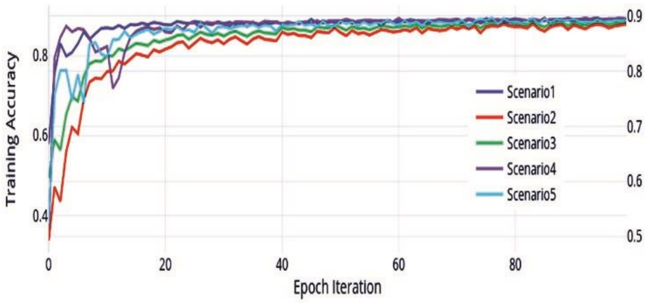

Figure 5 Accuracy of training: a comparison of five empirical studies

Figure 6 summarises the findings of a comparison of five empirical studies (from top to bottom: Scenario 1, Scenario 2, Scenario 3, Scenario 4, and Scenario 5; blue, orange, green, purple, and cyan trend lines) in terms of training prediction loss. The research employed 50 patients and subject-specific scenario validation with a split rate of 0.2 to detect sleep apnea. In each case, ten patients are selected, each representing a range of ages, BMIs, weights, and AHI indices. All lines of training loss decline until iteration 20, but beyond that point, it fluctuates unpredictably. They are much more durable after 20 cycles. The training prediction loss is stable between 20 and 100 iterations.

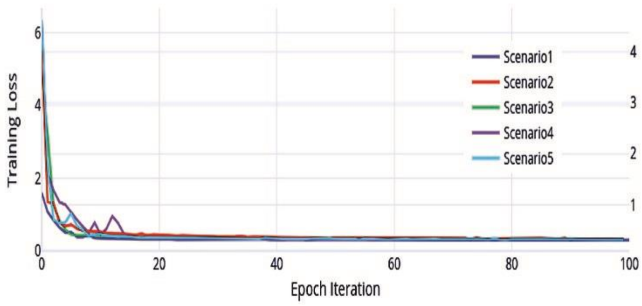

Figure 6 Comparison of training prediction loss for five empirical studies

Table 2 Classifiers used to identify sleep apnea were compared

S.no.	Type of Classifier	Split Rate	Training Data	Testing Data	Accuracy
1	LDA	0.5	50%	50%	86.5
2	SVM	0.3	70%	30%	90
3	Bagging Rep Tree	0.1	90%	10%	84.80
4	Artificial Neural Network	0.17	83%	17%	90.3
Proposed Model	CNN	0.2	20%	80%	91.3085

Table. 2 shows how the SPO2 signals are utilised to evaluate the proposed model in contrast to other classifiers. With a split rate of 0.2, whereby 20% of the data is used for training (794 samples), and 80% is used for testing (3178 data samples), the suggested model outperformed the previous models.

Classifiers such Linear Discriminant Analysis, Support Vector Machines, Bagging representation trees, and artificial neural networks are all compared and contrasted in Tab. 2. The results show that a deep convolutional neural network achieves a 91.3085 percent success rate. In other words, out of a total of 100,000 data points, the model performs very well for 913085 of them. The alternative model uses a small fraction of data for testing but a much larger fraction for training. To prevent overfitting, the suggested model for sleep apnea diagnosis only uses 20% of available data during training and 80% during testing. Considering that each patient's sleep data is recorded for an average of 8 hours, the 20% utilised for training is really too large for the suggested model. There are 50 participants in these 5 empirical investigations. While the competing model relies on a little amount of data from a generic database, the suggested model employs its own database populated with the medical records of fifty people diagnosed with sleep apnea with the aid of sleep lab technicians.

4 Conclusion

Before feeding data into deep learning, it is essential to clean and organise the model data. In order to prevent the overfitting issue, we use SPO2 with subject validation and a split rate of 0.2 to evaluate our deep CNN (convolutional neural network) model for deep learning experiments. When we will apply subject validation to the complete model, we will get an results approx. 92, with testing showing a clear improvement over training.

References

1. Burgos, A., Goñi, A., Illarramendi, A. & Bermudez, J. (2010). Real-Time Detection of Apneas on a PDA. *IEEE Trans. Inf. Technol. Biomed*., *14*, 995–1002.
2. De Chazal, P. & Sadr, N. (2016). Sleep apnoea classification using heart rate variability, ECG derived respiration and cardiopulmonary coupling parameters. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc*., 3203–3206.
3. Erdenebayar, U., Kim, Y. J., Park, J. U., Joo, E. Y. & Lee, K. J. (2019). Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram. *Comput. Methods Programs Biomed*. *180*, 105001.
4. Gottlieb, D. J. & Punjabi, N. M. (2020). Diagnosis and Management of Obstructive Sleep Apnea. *323*, 1389–1400.
5. Gula, L. J., Krahn, A. D., Skanes, A., Ferguson, K. A., George, C., Yee, R. & Klein, G. J. (2003). Heart Rate Variability in Obstructive Sleep Apnea: A Prospective Study and Frequency Domain Analysis. *Ann. Noninvasive Electrocardiol*., *8*, 144–149.
6. Gupta, V. & Pachori, R. B. (2021). FBDM based time-frequency representation for sleep stages classification using EEG signals. *Biomed. Signal Process. Control*. *64*, 102-265.
7. Gutierrez-Tobal, G.C., Alvarez, D., Crespo, A., Del Campo, F.&Hornero, R. (2108). Evaluation of Machine-Learning Approaches to Estimate Sleep Apnea Severity From At-Home Oximetry Recordings. *IEEE J. Biomed. Health Inform.*,*23*, 882–892.
8. Hassan, A. R. (2015). A comparative study of various classifiers for automated sleep apnea screening based on single-lead electrocardio- gram. *In Proceedings of the 2015 International Conference on Electrical & Electronic Engineering (ICEEE), Rajshahi, Bangladesh*, 45–48.
9. Hsu, C. C. & Shih, P. T. (2010). An intelligent sleep apnea detection system. In Proceedings of the 2010 International Conference on Machine Learning and Cybernetics, *Qingdao, China*, *6*, 3230–3233.
10. Ivanko, K., Ivanushkina, N. & Rykhalska, A. (2020). Identifying episodes of sleep apnea in ECG by machine learning methods. *In Proceedings of the 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine*, 588–593.
11. Javaheri, S., Barbé, F., Campos-Rodriguez, F., Dempsey, J. A., Khayat, R., Javaheri, S., Malhotra, A., Martinez-Garcia, M. A., Mehra, R. & Pack, A. (2017). Somers, Sleep Apnea: Types, Mechanisms, and Clinical Cardiovascular Consequences. *J. Am. Coll. Cardiol.*, *69*, 841–858.
12. Jayawardhana, M. & De Chazal, P. (2017). Enhanced detection of sleep apnoea using heart-rate, respiration effort and oxygen saturation derived from a photople-thysmography sensor. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 121–124.
13. Khandoker, A. H., Palaniswami, M. & Karmakar, C. (2008). Support Vector Machines for Automated Recognition of Obstructive Sleep Apnea Syndrome from ECG Recordings. *IEEE Trans. Inf. Technol. Biomed*.
14. Mendonca, F., Mostafa, S. S., Ravelo-Garcia, A. G., Morgado-Dias, F. & Penzel, T. A. (2018). Review of Obstructive Sleep Apnea Detection Approaches. *IEEE J. Biomed. Health Inform*. *23*, 825–837.
15. Novak, D., Mucha, K. & Al-Ani, T. (2008). Long short-term memory for apnea detection based on Heart Rate Variability. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 5234–5237.
16. Reed, M. J., Robertson, C. E. & Addison, P. S. (2005). Heart rate variability measurements and the prediction of ventricular arrhythmias QJM. *Int. J. Med.*, *98*, 87–95.
17. Sadek, I., Heng, T. T. S., Seet, E. & Abdulrazak, B. (2020). A New Approach for Detecting Sleep Apnea Using a Contactless Bed Sensor: Comparison Study. *J. Med. Internet Res*., *22*, 182-197.
18. Sharma, M., Tiwari, J. & Acharya, U. (2021). Automatic Sleep-Stage Scoring in Healthy and Sleep Disorder Patients Using Optimal Wavelet Filter Bank Technique with EEG Signals. *Int. J. Environ. Res. Public Health*.