SLEEP APNEA DETECTION USING DEEP LEARNING

Hnin Thiri CHAW, Sinchai KAMOLPHIWONG, Krongthong WONGSRISTRANG

Abstract: Sleep apnea is the cessation of airflow at least 10 seconds and it is the type of breathing disorder in which breathing stops at the time of sleeping. The proposed model uses type 4 sleep study which focuses more on portability and the reduction of the signals. The main limitations of type 1 full night polysomnography are time consuming and it requires much space for sleep recording such as sleep lab comparing to type 4 sleep studies. The detection of sleep apnea using deep convolutional neural network model based on SPO2 sensor is the valid alternative for efficient polysomnography and it is portable and cost effective. The total number of samples from SPO2 sensors of 50 patients that is used in this study is 190,000. The performance of the overall accuracy of sleep apnea detection is 91.3085% with the loss rate of 2.3 using cross entropy cost function using deep convolutional neural network.

Keywords: continuous single bio-parameter recording; deep convolutional neural network; deep learning; type 4 sleep study; portable sleep apnea detection

1 INTRODUCTION

The type of studies used in our system is type 4 sleep studies, which also refers to continuous single bio-parameter or dual-bio parameter recording. The minimum number of signals that can be used in this type 4 studies is one or two channels such as oxygen saturation and airflow [3-6, 12-30]. Normally type 4 studies do not have EEG and EMG signals, so scoring sleep is not possible. It can be used to identify respiratory sleep disorder during sleep time. The use of oxygen saturation and tracheal sound via acoustic sensor is another alternative to type 4 sleep studies. The primary purpose of using the deep convolutional neural network model for the detection of sleep apnea is to learn parameters or features for the model from the training dataset [7]. The performance of the model is related to the learned parameters in order to determine whether the sample of the signal contains sleep apnea event or not [7-9]. The relationship between the interpretability of the different parameters and the performance of the deep convolutional model is a challenging task for the detection of sleep apnea. A variety of different convolutional neural networks can be used for image classification task; however, only a smaller percentage of them is used for the detection relating to sound waves or the signal processing. The designing and testing of good convolutional neural network model is similar to the black box and needs to be done in a substantial amount of trials and errors because of the inner mechanism of the CNN model [9]. The representation of the signals from the sensor gains a lot of attention in the deep learning model. The proposed model is inspired from the image classification in which the amplitude from the sensors wave data can be seen as the digit vector figures, which can change to any dimensional shape for matrix vectorization or factorization. The main challenge of the sleep apnea detection is that the signal representation is not clear due to continuous nature of the high frequency breathing signal waves. The objective of proposing the biological signal to deep learning model is to fill the gap between the nature of deep learning and continuous nature of the biological signal from the sensors. The use of convolutional neural network model is a plausible and proper solution to interpret and learn the time series nature of the high frequency SPO2 signal [30-32]. After carrying out five empirical experiments for the validation and testing of sleep apnea detection, the result shows that it can provide meaningful interpretation for the clinical professional. The major contribution of the paper can be summarized as follows:

- A deep learning approach using convolutional neural network is proposed for the detection of sleep apnea in order to explore the segment of the time series data whether apnea occurs or not.
- The accuracy and loss rate SPO2 signal are compared with other approaches of sleep apnea detection using SPO2 signal.

The rest of this paper is organized as follows. The deep learning steps which include data collection, data preparation and segmentation, selection of deep learning model, training, testing, hyper-parameter tuning and prediction, are illustrated in section 2. The system design of the deep learning model can also be seen in Section 2. The performance evaluation results of our system are discussed in Section 3 with five empirical studies using subject validation of 50 patients. The studies contain 190,000 samples of the patients who have sleep apnea. However, the seriousness level of sleep apnea for each patient is different depending on the AHI index and the condition of the patients. Finally, Section 4 concludes and highlights the potential of future research direction for sleep apnea detection.

2 DEEP LEARNING STEPS

The biological neural network of human brain decision making inspires the proposed deep learning model using convolutional neural network [7, 9, 31, 32]. It is based on the single stimulus in which a number of neurons carries messages through electrochemical process, which is required for decision making. In the real world, the identification and prediction of sleep apnea by deep learning method would
facilitate patients with sleep apnea disease [1]. The deep learning model allows to learn from the feature to represent the nature of the data and its patterns. There are two types of trends in deep learning known as supervised learning, in which training data includes both input and desired output, and unsupervised learning, in which training data contains input but not the desired output [31]. Fig. 1 represents the deep learning process in general such as data collection, data preparation and segmentation, deep learning model design such as deep convolutional neural network, training, testing, hyper-parameter tuning and prediction.

For one-dimensional feature vector as shown in Fig. 3, the segmented signal will be labelled before being put in the system. For example, if \( X \) is assumed as one segmented SPO2 signal, \( X = \{x_1, x_2, x_3, \ldots, x_n\} \) in which \( n \) = sampling rate * one epoch (30 sec) and \( x_n \in X \), then \( x_n \in X \) where \( x \) is a random variable representing the signal data per second. Before putting it into the deep learning model, each segmented signal is stored in the feature vector as follow.

\[
\begin{array}{cccccc}
X_1 & X_2 & X_3 & \ldots & X_{n-1} & X_n & \text{Apnea/ Not Apnea}
\end{array}
\]

According to Fig. 4, the figure illustrates how the sleep apnea event occur. There are three consecutive time events known as \( t_1, t_2, \) and \( t_3 \). When breathing stops, the brain sends some signals to the muscle in order to do breathing. This kind of arousal event can be seen in the EEG electrodes’ signal that are attached to the patients’ head. This event can be seen and is marked in \( t_2 \) time segment. The arousal event from the EEG electrode, respiratory signals flat event can be seen in the following Fig. 5 and Fig. 6.

As for the event relating to the time \( t_1 \) from Fig. 4, it is significant that the signals coming from the nasal pressure, thermistor, thoracic and abdomen decreased and the signals become flat compared to the normal trend line. This kind of event can be seen in Fig. 5. In other words, these four signals

2.1 System Design

The design of the detection of sleep apnea can be seen in three layers in which the first layer, second layer and third layer are data collection layer, analysis layer and diagnosis layer respectively and can be seen in Fig. 2.

The first layer is for collecting the data on the oxygen saturation from the patient in order to detect sleep apnea. The second layer is for analysis of the detection of sleep apnea. In the third diagnosis layer, the number of events of sleep apnea occurrence can be identified together with how serious the sleep breathing disorder is for the patients. The third layer is called the diagnosis layer in which health care professional identifies whether the identification of sleep apnea is correct or not.
drop significantly compared to the normal breathing condition. The first four signals highlighted in blue are thermistor, nasal pressure, thoracic and abdominal signal.

After occurrence of the signal drop in these four signals, the oxygen desaturation occurs in SPO2 signal highlighted in red. The abrupt shift of the brain signal from the electrodes with the frequency of less than 16% which does not include spindle and it can last for at least 3 sec. When the sleep apnea occurs, this event is highlighted using green and a red box in Fig. 6.

As a result of the lung stop breathing, the occurrences of oxygen desaturation can be seen in the \( t_3 \) time event of SPO2 signals relating to Fig. 6. The C3-M2 means the central electrode on the left side of the brain is referenced to the right pre-auricular electrode near the right ear. The C4-M1 means the central electrode on the right side of the brain is referenced to the left pre-auricular electrode near the left ear. The F3-M2 means the frontal electrode on the left front side of the brain is referenced to the right pre-auricular electrode near the right ear. The F4-M1 means the frontal electrode on the right side of the brain is referenced to the left pre-auricular electrode near the left ear. The O1-M2 means the occipital electrode on the left front side of the brain is referenced to the right pre-auricular electrode near the right ear. The O2-M1 means the occipital electrode on the right side of the brain is referenced to the left pre-auricular electrode near the left ear. One of the findings is that the heart rate began to increase after the oxygen desaturation event happened in time \( t_3 \).

### 2.2 Apnea/Hypopnea Event Detection

The flat signals do not mean that the signal is very flat because the up and down can be seen according to minute, second and hour view. The baseline amplitude for the apnea is the preceding 3 to 6 breaths in which time segment before and after the event is represented as \( t_1 \) and \( t_3 \) respectively, and the amplitude before and after the apnea event is represented as \( A_1 \) and \( A_2 \) respectively in which \( t_1 < t_2 < t_3 \). If the maximum relative amplitude for the apnea threshold is set to 10 percent, it means that \( A_2 / A_1 \) should be less than 10 percent and the apnea event is detected if the amplitude of the event is twenty percent or below of preceding data according to Fig. 7. For the hypopnea event the maximum threshold is less than 70 percent according to AASM manual scoring 2017.

The following Fig. 8, illustrates that the minimum duration of the event that can be identified as apnea or hypopnea is ten seconds. It is highlighted in red bounding box in which \( T \) is equal to ten seconds.

### 2.3 Deep Convolutional Neural Network

The simplified version of the convolutional neural network is portrayed in Fig. 9 in which the first layer is the input layer, the last layer is the output layer which can be identified whether certain segment from the sensor has apnea or not, and the layers between the first and last layer are the hidden layers.

The following Tab. 1 represents the summary of the design of the deep convolutional neural network in which ten layers are used. Moreover, it includes three two-dimensional convolutional layers and three max pooling layers. After running the deep convolutional neural network model for each empirical study, there will be 32,614 trainable parameters for the convolutional neural network and 794 samples are used for training and 3,178 samples are used for
3 PERFORMANCE EVALUATION

It can be seen that five empirical studies are made using the accuracy metrics of the proposed deep convolutional neural network model using fifty patients and each empirical study includes 3,700 to 4,000 sample data of ten patients. Each empirical study is illustrated in two figures: one for accuracy and the other for loss. The cross entropy is used for calculating the loss rate. The overall performance of the proposed model is 91.3085 percent. The comparison between proposed convolutional neural network model and the other models can also be found in this performance evaluation section. It contains the comparison different classifier such as linear discriminant analysis (LDA), support vector machine (SVM), bagging representation tree, artificial neural network using SPO2 signal and the proposed convolutional neural network of deep learning model.

The summary of the comparison of five empirical studies regarding training prediction loss can be seen in Fig. 11 in which scenario 1, scenario 2, scenario 3, scenario 4 and scenario 5 are illustrated using blue, orange, green, purple and cyan trend line respectively. In this paper, the total number of patients that participated in the detection of sleep apnea is fifty patients and subject specific scenario validation is used with the split rate of 0.2. Each of the scenarios had ten different patients of different age, BMI, weight, AHI index. All of the trend lines relating to training loss were going down until the number of twenty iterations and at the same time, they fluctuated and were unstable until twenty iterations. They became more stable after 20 iterations. The stable condition for training prediction loss was between 20 iterations to 100 iterations.

<table>
<thead>
<tr>
<th>Table 1 Deep Convolutional Neural Network Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer (type)</td>
</tr>
<tr>
<td>conv2d 1 (Conv2D)</td>
</tr>
<tr>
<td>max_pooling2d 1</td>
</tr>
<tr>
<td>conv2d 2 (Conv2D)</td>
</tr>
<tr>
<td>max_pooling2d 2</td>
</tr>
<tr>
<td>conv2d 3 (Conv2D)</td>
</tr>
<tr>
<td>max_pooling2d 3</td>
</tr>
<tr>
<td>flatten 1 (Flatten)</td>
</tr>
<tr>
<td>dense 1 (Dense)</td>
</tr>
<tr>
<td>dropout 1 (Dropout)</td>
</tr>
<tr>
<td>dense 2 (Dense)</td>
</tr>
<tr>
<td>Total params: 32,614</td>
</tr>
<tr>
<td>Trainable params: 32,614</td>
</tr>
<tr>
<td>Non-trainable params: 0</td>
</tr>
<tr>
<td>Train on 794 samples, validate on 3178 samples</td>
</tr>
</tbody>
</table>

According to Tab. 2, the SPO2 signals are used for the comparison between the proposed model and other classifiers. The proposed model outperformed the other models with the accuracy of 91.3085 percent with the split
rate of 0.2 in which the training data is 20 percent (794 samples) and testing data is 80 percent (3178 data samples).

The following Tab. 2 illustrates the comparison between different types of classifier such as LDA, SVM, Bagging representation tree, artificial neural network for sleep apnea detection. It can be seen that deep convolutional neural network has the accuracy of 91.3085 percent. In other word, the model can predict very well for 913085 data out of 100,000 data. The other model use only a few percentage data for testing and a huge percentage of data for training. However, the twenty percent of data are used for the training and 80 percent of data are used for testing in order to avoid overfitting in the proposed model for detection of sleep apnea. Actually, the twenty percent used for training is not small enough for the proposed model because of 8 hours of sleep recording for each patient. Fifty patients involve in this five empirical studies. The other model uses the standard database with a few amount of data, however proposed model use own database after collecting data of fifty patients who have sleep apnea with the help of health care professionals from the sleep lab.

4 CONCLUSION

As a summary, the preparation and segmentation the data model is a crucial step before feeding into the deep learning. The experiment of deep learning using deep CNN (convolutional neural network) model is tested using SPO2 using subject validation with the split rate of 0.2 in order to avoiding overfitting problem. The overall accuracy for the entire sample using subject validation is 91.3085 in which the trend line for testing is better than the trend line for training

Acknowledgement

This work was supported by Thailand’s Education Hub for Southern Region of ASEAN Countries (TEH-AC), Prince of Songkla University Graduate Studies Grant as well as by Department of Computer Engineering, Prince of Songkla University. The authors would like to express the gratitude for our supervisor, co-supervisor and sleep technicians from Songklanagarind Hospital, Thailand for guiding in my research work and supporting access to the respiratory signals acquired in Sleep Disorders Service and Research Center.

Remark

The article was orally presented at the 23rd International Computer Science and Engineering Conference (ICSEC2019).

5 REFERENCES


International Conference on Information Processing in Sensor Networks (IPSN), 179-190. https://doi.org/10.1145/2461381.2461405


Authors’ contacts:

Hnin Thiri CHAW
Dept of Computer Engineering,
Prince of Songkla University,
90112 Hat Yai, Thailand
hninthirichaw87@gmail.com

Sinchai KAMOLPHIWONG
Dept of Computer Engineering,
Prince of Songkla University,
90112 Hat Yai, Thailand
ksinchai@psu.ac.th

Krongthong WONGSIRIRANG
Dept. of Otolaryngology Head Neck Surgery,
Songklanagarind Hospital,
90112 Hat Yai, Thailand
golf@hotmail.com

266 TECHNICAL JOURNAL 13, 4(2019), 261-266