Apnea Detection in a Contactless Multisensor System using Deep Learning Algorithm

A. Pomprapa1, M. S. Sayani1, T. Anwar2, A. Stollenwerk3, S. Kowalewski3, P. von Platen1, S. Leonhardt1

1Chair for Medical Information Technology, RWTH Aachen University, Aachen, Germany
2Universiti Teknologi Petronas, Perak, Malaysia
3Informatik 11 Embedded Software, RWTH Aachen University, Aachen, Germany

Contact: pomprapa@hia.rwth-aachen.de

Introduction

Obstructive sleep apnea (OSA) is an ordinary breathing disorder, causing a temporary blockage of the upper airway during sleep. It usually stops breathing for 10 seconds or longer and the apnea episodes may regularly happen from 5 to 100 times an hour. The severity criteria can be assessed by Apnea-Hypopnea Index (AHI) or Respiratory Disturbance Index (RDI) in three main categories: mild (5-15 events an hour), moderate (15-30 events an hour), and severe (more than 30 events an hour) [1]. The signs include loud snoring, restless sleep, or daytime sleepiness. If it remains untreated, OSA can lead to many possible serious health complications, such as depression, mood disorders, and heart disease. It can also bring about drowsiness with the uprising risk of accidents while driving and working.

In this work, a prototype of contactless multisensor chair was set up in order to detect vital parameters and a deep learning algorithm based on a hybrid model of convolutional neural network (CNN) and recurrent neural network (RNN) was applied for the detection of apnea in a sitting position, simulated scenarios such as in a pilot seat, in a seat of truck driver or in any types of passenger seat, where a big data can be collected during a flight or a journey, processed in real-time for prewarning or alarms, and transmitted to a cloud storage at a medical center. The main objective of this work is to set up a new software architecture for data acquisition and vital parameter extraction and to demonstrate the integration of deep learning algorithm for the detection of apnea.

To realise this system for the digital lifestyle, a modified \(\lambda\) architecture is proposed with a compromise between real-time processing and restricted computational resources. This software structure offers the benefits not only to a real-time stream processing but also a batch or a large volume of collected data processing. Typically, this approach tries to balance throughput, fault-tolerance, and latency (or time interval) in batch processing.

This contribution is organized as follows. In the section of “Materials and Methods”, it begins with a hardware configuration for data acquisition based on multiple devices for the monitoring of heart and lung activities. It also provides a software architecture using \(\lambda\) architecture as well as a deep learning algorithm. The results of our experiments are then given in a subsequent section, namely “Results”, followed by a section of “Discussion”. The article ends in the last section with the conclusions and future works.

Materials and Methods

Hardware Configuration

A hardware system consists of a number of unique contactless sensors mainly produced in the Chair of Medical Information Technology (MedIT), including a magnetic induction (MI) sensor for the monitoring of breathing [2], a cECG system for the monitoring of heart activity [3], a smart photoplethysmography (PPG) [4], ballistocardiography (BCG) [5], and a thermal camera (Seek Thermal Compact) with a micro-USB connector. The additional analog/digital (A/D) converter was set up by using NI USB-6259 as an interface for cECG and BCG measurements with a sampling frequency of 1,000 Hz. All devices are then connected to a FUJITSU desktop computer (Intel\textsuperscript{TM} Core\textsuperscript{TM} i5-3470 CPU @3.20 GHz, 4 cores with 14 GB random access memory (RAM)).

Data Flow with System Interface

All signals are acquired through different communication ports and socket. Signal processing and feature extraction were carried out in all channels and passed directly to Kafka producer. The producer consists of a pool of buffer space, which can hold records that have not been transmitted to the server. Apache Kafka is able to handle these big data with very low time interval or latency with a range in milliseconds. The data flow of system interface is shown in Figure 1.

Figure 1: Data flow with system interface in a contactless multisensor system [6].
However, in this article, we mainly focus on the signal acquired from a MI sensor, where respiratory rate (RR) is extracted and the preprocessed signals are used for the detection of apnea.

**Software Architecture**

To process and analyse big data, the state-of-the-art architecture [7] is proposed as demonstrated in Figure 2. The software is programmed by Java programming language for data acquisition. With the proposed architecture, Apache Kafka and Apache Spark worked underground to manage the data in two main layers, namely speed layer and batch layer. The speed layer supports stream processing where fast data is being processed in real-time. While batch layer handles all stored records for batch or large collection processing. The deep learning model can use both data from speed layer and batch layer for apnea detection. With limited computational resources, the full functionality of deep learning algorithm in the batch layer had not been implemented in this work. The visualization of real-time data and apnea detection was set up using Kibana with the aid of AngularJS, which integrated into a serving layer. In our configuration, Elasticsearch was used as a search engine or a master database in this work.

**Figure 2:** λ architecture for apnea detection using a deep learning algorithm [6].

**Deep Learning Model for Apnea Detection**

In this work, two different structures of deep learning models have been implemented, namely convolutional neural network (CNN) [8] and a hybrid model of CNN and long short-term memory (LSTM). This was implemented under a Java-based framework using Deeplearning4j, which is a deep learning library for integration, vectorisation, modeling, and evaluation. Since the performance of the hybrid model is significantly better than the proposed CNN model, only the hybrid deep learning model is presented in this article. LSTM can be classified as a recurrent neural network (RNN) [9]. It offers a feedback structure of neural network whilst CNN provides a feedforward neural network. LSTM is explicitly designed to prevent the long-term dependency problem. However, it might have a convergence issue during model training. The designed hybrid deep learning model with 9 layers is given in Figure 3. It contains 3 CNN layers, 3 max subpooling layers, 1 dense layer, 1 LSTM layer and 1 last RNN output layer.

**Figure 3:** Deep learning model based on the multi-layer hybrid architecture [6].

The MI input signals were divided into chunks of 2,000 points or 20 seconds. The classification was carried out in order to detect normal breathing or apnea from a machine perspective. In this particular system, Xavier weight initialisation method was defined with adaptive gradient algorithm [10] as an update strategy, negative log likelihood as a loss function, softmax activation function in the final RNN output layer with 0.01 learning rate.

The apnea data were collected mainly by a MI sensor from five healthy male subjects at the rest position and they were requested to mimic central apnea by holding their breathing for short periods irregularly. These generated the database of breathing patterns over seven hours. The starting and stopping periods of mimic apnea were labels as used as a supervised learning. 70% of the dataset (976 records) were used for training the deep learning model for detection of
apnea based on the MI sensor and 30% of the dataset (416 records) were applied for validation purpose.

Results

Data Acquisition

The primary signals were acquired from three main sensors i.e. photoplethysmography (PPG) sensor, magnetic induction (MI) sensor, and cECG system. All of these signals can be used for the extraction of respiratory signals [11], [12], [13]. The original signals were acquired and given in Figure 4. These signals can be preprocessed using appropriate filters or techniques for feature extraction. In this work, only a second-order Butterworth bandpass filter was applied to the processing of MI signals. The passband was designed in the range between 0.1 Hz and 1 Hz, corresponding to the breathing frequency of 6 bpm to 60 bpm.

Figure 4: A sample set of primary sensors acquired from MI, cECG, and PPG sensors for 30 s [6].

In this work, a peak detection algorithm was performed in order to estimate the respiratory rate (RR) based on a moving window of 10 s [14]. The estimation of RR was mainly for data representation and this feature was not used for training the hybrid deep learning model.

Training Performance of the Deep Learning Model

The overall 1,392 breathing records were applied for the evaluation of the proposed hybrid deep learning model. Each record contained the breathing data of 20 s. 70% of the dataset was used for training the CNNs and RNNs. The training was performed to update the model parameters, resulting in one cycle or an epoch. The repetitive training was activated one after another to update the new model parameters in order to achieve higher accuracy with no overfitting. In this particular work, the training was carried out for 34 epochs with the highest accuracy of 92% and the performance of the model is shown in Figure 5. The rest 30% of the dataset was utilized for validation purpose and its highest accuracy in the validation dataset was 90% at the 34th epoch. These two datasets were selected randomly.

Figure 5: Training performance of the model [6].

Visualisation of Apnea Detection

The data from the speed layer were stored in a database, handled by Elasticsearch. For visualisation of streaming data, Kibana dashboard was designed for real-time visualisation of all acquired signals and AngularJS was used as a platform for building web applications. In Figure 6, the upper curve showed the original MI signal for 60 s, which contained mimic apnea and the respiratory rate of 12 bpm was shown at the top right corner, which was computed based on the peak detection algorithm during feature extraction.

Figure 6: Visualisation of magnetic induction (MI) signal based on Kibana visualisation plugin for Elasticsearch and AngularJS platform [6].

The middle curve in Figure 6 represented a filtered MI signal based on the designed second-order Butterworth filter. The red part is highlighted using Java programming for the detection of apnea episode in real-time processing. Therefore, the system can effectively recognise the mimic apnea and it should be able to detect apnea in various scenarios for...
real-time warning and for investigation of historical records in batch processing.

**Discussion**

Due to the fact that typical respiratory rate for a healthy adult at rest is approximately 12-18 bpm, one breathing cycle thereby takes 3.3-5 s. Since a fixed 20-s window was implemented for the evaluation of apnea episode, 4-6 breaths are contributed to this timeframe. The detection of apnea based on this window size can be activated whenever there was no breathing cycle for 2-3 breaths. The adjustment of window size may be made for the notification of an apnea episode, however, the proposed window size was reasonable for the implementation of apnea detection.

Because there was no data available for the detection of respiratory patterns using a MI sensor, we, therefore, set up the measurement with a random mimic of apnea and only 1,393 breathing records were acquired and stored in the database. They included normal breathing as well as the mimic apnea. They were used for the training of the proposed deep learning model with the highest accuracy of 92% in the training dataset and 90% in the validation dataset. These accuracies should be improved by collecting a larger quantity of data.

Convolutional neural network (CNN) is a type of feed-forward artificial neural network and is usually applied to two-dimensional signals i.e. images. However, it can also be applied to time series like in this case. While long short-term memory (LSTM) network is a special kind of recurrent neural network (RNN) with a smart dynamical feedback structure to form short memory in the networks. Based on the dataset, the hybrid model of CNN and LSTM provided much better performance for the apnea detection than a CNN model. In this article, the hybrid model was therefore presented.

**Conclusions and Future Works**

A hybrid deep learning model based on CNN and LSTM has been implemented for the detection of apnea with the dataset measured from 5 healthy volunteers. The highest accuracy was 92% and 90% in the training dataset and in the validation dataset, respectively. With the aid of λ architecture, streaming data and batch data can be processed in real-time and apnea detection can be realised in the proposed contactless multisensor system, which can be used for an immediate alarm or long-term medical diagnostics.

There are many aspects in which this contactless multisensor system can be extended. The following points should be implemented in the future:

1. Classification of apnea into three main categories, namely mild, moderate, and severe
2. Comparison and evaluation of different deep learning models (i.e. CNN, LSTM, and a hybrid model of CNN and LSTM) with other classification algorithms such as random forests, decision trees, or nearest neighbors.
3. Implementation of robust sensor fusion for the detection and classification of apnea.

**References**


