Sleep Apnea detection using Deep Learning Methodologies

MSc Research Project
Data Analytics

Gunjit Jain
Student ID: x20251432

School of Computing
National College of Ireland

Supervisor: Mr. Vladimir Milosavljevic

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Gunjit Jain
Student ID:	x20251432
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Mr. Vladimir Milosavljevic
Submission Due Date:	19/09/2022
Project Title:	Sleep Apnea detection using Deep Learning Methodologies
Word Count:	XXX
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	19th September 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Sleep Apnea detection using Deep Learning Methodologies

Gunjit Jain
x20251432

Abstract

One of the most common sleep disorders in the present world is Sleep apnea, which is a condition in which a person stops breathing while in sleep for 10 seconds up to a minute. This happens due to the blockage in the upper airway in the throat muscles. Polysomnography has been the gold standard for the detection of sleep apnea, where a person needs to visit a sleep clinic and sleep the whole night under the supervision of a sleep expert. Numerous sensors are attached to the body to capture the readings. Hence the process is very expensive, time taking, and intrusive. The readings include heart rate variability, SpO2, and other physiological data. This research aims to detect sleep apnea at home using smartwatches. The ECG and SpO2 readings have been analyzed using a spectogram. The models used in this study are CNN and hybrid neural network DenseNet121 + CNN for classifying whether a person is suffering from Sleep Apnea or not. The findings of this research can be used to detect sleep apnea at home.

Keywords: Sleep Apnea, Obstructive Sleep Apnea, Convolutional Neural Network, Deep Learning, DenseNet121, Spectogram

1 Introduction

The process of sleeping is crucial to human existence. It significantly affects one's overall quality of life as well as physical and mental health. Lack of focus, memory lapses, energy loss, lethargy, weariness, and emotional instability are just a few of the symptoms caused by sleep disorders. Diabetes, depression, and high blood pressure are the three most typical adverse consequences of sleep apnea. A sleep disorder termed sleep apnea (SA) is defined by short breathing pauses when a person is asleep. Around the world, 3-7% of adult men and 2-5% of adult women have OSA., with the global prevalence ranging from 3 to 5 percent (Adnan et al.; 2022). It happens because of complete closure of the airway near the throat (apnea) or partial obstruction of this airway, which reduces the quantity of oxygen the body absorbs by 50%. (i.e. hypopnea). A person's breathing usually stops during sleep for 10 seconds or longer, however, it occasionally does so for up to a minute.

The World Health Organization claims that there are 200 million people who suffer from sleep apnea globally. The prevalence of the condition fluctuates between 2 and 4 percent. In addition to being common, sleep apnea is dangerous since it can result in a number of life-threatening problems. Arousals are frequent, leading to ineffective sleep and ongoing sleep loss, which causes daytime tiredness, irritation, fatigue, and poor attention, as well as poorer cognitive capacity. Commonly, these symptoms are followed by even more serious consequences, such as job conflicts, and traffic accidents. Other issues caused by sleep apnea include hypertension, heart attacks, morning headaches, irregular heartbeats, stroke, shortened life expectancy, etc.

Obstructive Sleep Apnea (OSA) and Central Sleep Apnea(CSA) are two major forms of sleep-disordered breathing (SDB). However, OSA makes up more than 80% of all cases of SBD, making it the most prevalent kind; CSA is far less common. Even when there is still an attempt to breathe, OSA is caused by airflow restriction because of airway obstruction or upper airway collapse in the throat muscles for a minimum of 10 seconds or more (Nguyen et al.; 2014). In CSA, instead of the airway blockage, breathing stops due to insufficient respiratory effort because the brain fails to instruct the muscles to breathe (Randerath; 2022). Hypopnea is when the airflow is partially obstructed while asleep and the individual takes shallow breaths for ten seconds or longer. The airflow is 30% less during hypopnea than it is during normal breathing. Obstructive hypopnea, central hypopnea, and mixed hypopnea are other subtypes of hypopnea.

Identifying and diagnosing sleep apnea is primarily determined by overnight polysomnography (PSG) in specialized sleep laboratories (Nguyen et al.; 2014). Various physiological markers, including cardiovascular function, respiration, sleep status, and oxygen saturation are examined during polysomnography. It includes non-invasive ways to collect physiological information such as ECG (Electrocardiogram), SpO2 (Pulse Oximetry), nasal airflow, EEG (Electroencephalogram), EMG, and abdominal and thoracic movements. Skilled sleep technicians then classify each segment of the collected signals as apnea or non-apnea using a standard reference. With these annotations, the Apnea-Hypopnea-Index (AHI) is calculated to classify the severity of the condition as severe, moderate, or mild based on the occurrence of apnea and hypopnea events in an hour. To manually diagnose, doctors must invest a lot of time in monitoring and analyzing such data. PSG is a pricey and time-consuming treatment, nevertheless, because of the pain of the electrodes and the vast amount of information needed.

In order to minimize the pain of the PSG procedure, extensive research has been done on how to diagnose sleep apnea at home utilizing portable instruments. Studies have been done to determine which physiological signal may be utilized to identify sleep apnea using wearable technology at home. The most encouraging outcomes are frequently displayed by heart rate variability and blood oxygen readings. This study aims to simplify sleep apnea diagnosis by applying deep learning models. People don't realize they have sleep apnea until someone in their family recognizes it, so approximately 80% of cases go undiagnosed. Despite identifying the problem, they don't want to visit a sleep facility and spend many days sleeping through the night to assess the complexity of the problem.

The goal of this study is to streamline the process of sleep apnea detection so that consumers can check for it at home with their smartwatches. The ECG and SpO2 data, both of which are signals, are employed in this study. Most of the smartwatches in the market today are capable of reading this data in real-time. This data can be used to effectively analyse sleep apnea at home.

The research question is the problem statement that the study seeks to solve. *'How effectively can heart rate and SpO2 data be utilized to predict Sleep Apnea using a smartwatch?'* is the research question for this research.

The research question has been addressed through multiple tasks, hence the research goal has been achieved. To accomplish this research's objective, data collection, data preparation, including signal processing, and feature extraction are key activities. The Apnea-ECG database was used for performing these activities. Since the ECG and SpO2 data are signals, signal processing must be done in order to extract their characteristics from them. The baseline CNN model and a hybrid neural network consisting of DenseNet121 + CNN are used for further categorization. The effectiveness and applicability of each of the suggested sleep apnea categorization models will next be evaluated.

The signals present in the form of records are shown in Figure 1 below. These signals were converted into spectogram using Fast Fourier Transformation and then deep learning models are applied to it. This methodology has never been used to detect sleep apnea in the past works accessed so far. By utilizing these techniques, this research aims to close the detection gap for sleep apnea at home. The field of sleep apnea diagnosis at home can benefit from this research's use of these very recent deep learning methods.

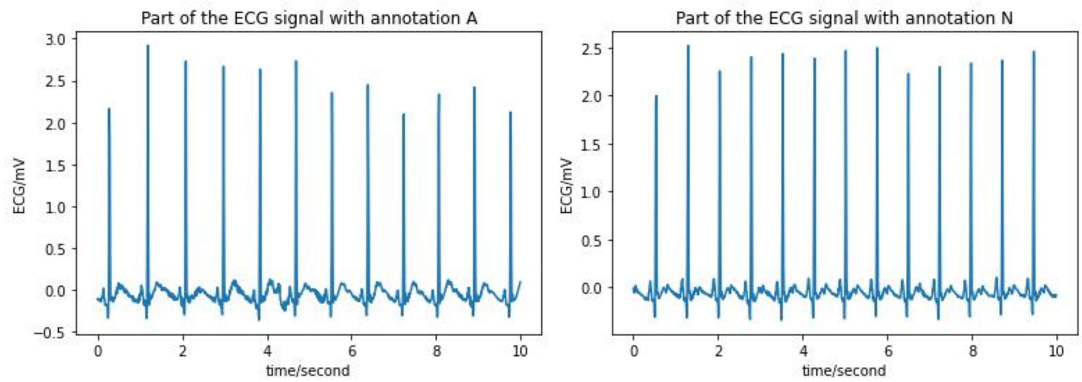

Figure 1: Sample of ECG signal for both annotations

2 Related Work

This part critically analyzes previous research on sleep apnea that has taken place in recent years. The medical industry has used a variety of deep learning and machine learning techniques to improve the speed, accuracy, and convenience of illness identification and treatment. The gold standard for detecting sleep apnea up until this point has been polysomnography, which is costly, busy, and involves human involvement. As a result, numerous researchers have worked to simplify this procedure so that it may be completed at home.

2.1 Heart rate variability and ECG signals for detection

Various types of studies have made substantial use of single-channel ECG data to identify sleep apnea. Additionally, it is captured during polysomnography. Heart rate variability was utilized by (Schrader et al.; 2000) to identify Obstructive Sleep Apnea (OSA) using ECG data from the PhysioNet database. Using spectral components and other derived variables, four separate frequency bands were classified using linear discriminant functions. Finally, Heart rate variability, which was once thought to be an efficient approach to diagnosing Sleep Apnea, was used to detect Obstructive Sleep Apnea (OSA). The accuracy of the ECG data using various signal processing algorithms and autoregressive models was 92.0% for learning set and 88.31% for test set. Due to the fact that this approach works with R-R time series data of ECG signals, it is challenging to implement for smartwatch data.

To classify OSA more accurately, a somewhat different technique has been utilized in various publications. Different signal processing techniques have been utilized to deconstruct the signal and extract characteristics that will be used for classification in the pre-processing stage. A method for diagnosing sleep apnea using single-lead noisy ECG data is presented by (Fatimah et al.; 2020). The Fourier decomposition method (FDM) was applied to break down the provided ECG data. With the necessary cut-off frequencies, this divided the signal into a collection of Fourier intrinsic band functions (FIBFs). Following the removal of baseline wander noise and power-line interference, each FIBF was then cleaned up, and properties including entropy and mean absolute deviation were retrieved and supplied to the classifier. But Fourier Transform (FT) is utilized for stationary signals while ECG is a continuous signal, (Qatmh et al.; 2022) employ Discrete Wavelet Transform (DWT) instead. DWT is effective for transitory signals as well. After the first breakdown of the ECG data, statistical characteristics were collected, and ANN was utilized to categorize. The validation performance of the MATLAB-based neural pattern recognition (NPR) tool was 92.3 percent after 19 iterations. Considering that only 1 hidden layer and 7 neurons were employed, this was incredibly successful. This demonstrates that identifying sleep apnea can benefit from a small amount of pre-processing using a less sophisticated neural network.

In a research, (Sharma et al.; 2019) conducted an automated OSA diagnosis which was developed based on short-duration ECG records, utilizing wavelet-based analysis as a practical tool for spectral localization of non-stationary data, including ECG signals, which were analyzed using wavelets.. The ideal two-band filter bank (FB) approach was used to divide the ECG signal into wavelet-frequency bands. In order to test the 35-fold cross-validation method, several classifiers were employed. Compared to standard state-of-the-art methods, this provided a decent level of accuracy. For single-lead ECG, (Feng et al. 2021) attempted to add unsupervised feature learning because, typically, feature sets for medically related data are insufficient and seldom well annotated. In this study, a SA detection method based on the time-dependent cost-sensitive (TDCS) classification model and the frequential stacked sparse auto-encoder (FSSAE) was proposed. For better classification performance, the TDCS classification model combines the MetaCost algorithm with Hidden Markov models (HMMs). The FSSAE employs an unsupervised learning strategy to extract feature sets. Traditional feature engineering methods take a very long time and require a lot of prior knowledge to choose the best feature sets. The

AUC value for the TDCS classifier was 0.8596. It is rather obvious from these studies that an ECG is a reliable predictor of finding sleep apnea.

2.2 Detection using SpO2 measurements and respiration signals

Since the ECG signal varies with breathing, SpO2 data may be found in both signal form and the ECG signal collected via PSG. Some domain expertise is needed to use these signals to extract features. However, a deep belief network has been used by (Mostafa et al.; 2017) to extract the characteristics without domain-specific information. For the creation of the deep belief network, Restricted Boltzmann Machines autoencoders with two layers and a soft-max encoder with one layer were used. Since SpO2 is the most convenient to record at home, it is predominantly employed in this investigation. Deep Belief Network, a form of unsupervised learning, was utilized to automatically identify characteristics from the raw data (DBN). The databases for the UCD and Apnea-ECG were resampled at 1 Hz. Finally, a separate index that is the average of accuracy, specificity, and sensitivity was employed to measure the outcomes.

Therefore, the main lesson to be learned from this study is that SpO2 is a crucial element to take into account for the identification of sleep apnea. Additionally, feature extraction using a neural network is possible without significant signal processing on the SpO2 signal data.

2.3 Detection utilising SpO2 and ECG information

(Ghandeharioun; 2021) has employed a combination of ECG and SpO2 data to identify sleep apnea since ECG and SpO2 are equally good factors to predicting sleep apnea. They have employed a mutual information analysis in addition to deep learning or machine learning approaches to choose SpO2 and ECG characteristics. This study employs mathematical relations and transformations before training the classifier since using deep learning techniques on the raw data requires a lot of computing power and high-quality data storage. They used one private database in addition to two public databases. SpO2 from the UCD database and ECG signals from the Apnea-ECG database was the key features taken into account.

(Van Steenkiste et al.; 2018) has demonstrated that different respiratory signals which are derived from the ECG can be utilized to automatically diagnose sleep apnea because PSG is a time-consuming method. In this work, machine learning techniques were used to compare respiratory signals obtained from ECG recordings, including abdominal respiratory belts, total relative chest volume, thoracic respiratory belts, and a few others. To extract pertinent respiratory events, all the data were run through a low-pass filter having a cutoff frequency of 0.75 Hz. All the models were validated using a 5-fold cross-validation procedure. The study's final finding was that while hypopnea is more difficult to detect automatically, OSA and CSA are comparatively easier. Additionally, a tiny sensor on the patient's chest is used to produce a bio-impedance signal, which works effectively. This suggests that the ECG and SpO2 in combination can be a useful tool for identifying Sleep Apnea.

2.4 Detection using data from a smartwatch

Numerous research has attempted to use wearable devices to detect sleep apnea as the PSG is a very complex procedure. To detect the same at home, numerous technologies have been developed. Today's smartwatches come equipped with a variety of sensors, including SpO2 readers, heart rate monitors, accelerometers, etc. thanks to technological advancements. As a result, (Chen et al.; 2021) attempted to identify various Sleep Apnea categories using a smartwatch. The smartwatch's accelerometer record a person's signal when they are sleeping. The smartwatch accelerometer records chest, arm, and wrist motions during breathing and while the user is sleeping. This data is subtle and a bit noisy. They have attempted to mitigate this by denoising the signals using a variety of data calibration approaches. These signals' troughs and peaks were retained because they represent the cycle of respiration and potential apnea occurrences. When an apnea event occurs, the oxygen level drops, and the person then takes a quick breath, which raises the signal of the accelerometer. The proposed technique has a number of limitations. Vibration can be used to inform the person to wear the smartwatch properly if it is being worn loosely, which prevents the movements from being properly caught. Furthermore, it was discovered that the closer the smartwatch is to the wrist, the more precise the respiratory signal is. The data was analyzed using a variety of machine learning approaches, however, the Adaptive Boost algorithms and the Random Forest produced the best results. The reported F1 score was 0.9649.

In order to improve the accuracy of sleep and sleep stage prediction, (Walch et al.; 2019) demonstrated for the first time how to apply the mathematical methods which are accepted and widely accessible to analyze heart rate and raw acceleration data from a common wearable device. Along with the PSG data, they also captured heart rate information and acceleration from the Apple watches. The heart rate data was filtered, smoothed, and interpolated to provide a value once per second in order to emphasize high-change phases. As classification algorithms, Random Forest, KNN, logistic regression, and multilayer perceptron (MLP) were first employed. The results of this study were validated using cross-validation including Monte-Carlo and leave-one-out cross-validation. A unique approach to screening OSA utilizing sound screening and a deep learning model was put out by (Romero et al.; 2022). It analyzes smartphone recordings of sleep breathing made at home. This method produced results that were almost as accurate as SpO2-based detection, although the system based on acoustic often had more specificity. Due to the possibility that the desaturation was unrelated to sleep apnea, more false positives were produced by the SpO2-based system than the acoustic-based system, despite its high sensitivity. This study suggests a simple, do-it-at-home solution, but it has several shortcomings. The sound recorded by the system may contain ambient noise and/or may not have been fully captured owing to the subject's shifting posture during sleep. From these studies, it was demonstrated that various data that may be obtained from a wristwatch or a smartphone can be utilized intelligently to identify sleep apnea at home more accurately than other cutting-edge methods.

2.5 Detection through conventional techniques and other physiological factors

The human body undergoes several alterations during a sleep apnea episode. The chest and abdomen may move, and there may be adjustments to the SpO2, heart rate, airflow and other factors. In the example below, CNN is used to automatically identify and extract characteristics from the original data (Cen et al.; 2018). Signals for rib-cage, oronasal airflow, blood oxygen saturation, and abdominal motions were up-sampled before being adjusted with a mean unit variance of 0. To detect OSA events in this instance, a 1-second annotation was employed rather than a 1-minute annotation, since annotations of 1-minute limit both the temporal precision and the amount of data available for model training. This study also included a technique known as leave-one-out cross-validation, according to which the classification model was trained using the remaining acquisitions while one acquisition from all the acquisitions was selected each time for testing. This resulted in an average accuracy of 79.61% when combined with deep learning.

(Zhang et al.; 2013) provided a system for real-time detection and categorization of sleep apnea. To identify and cure sleep apnea, they have developed a smart-pillow architecture in conjunction with a pulse oximeter. A pillow modification algorithm with real-time feedback is created to alter the pillow's shape. The level of blood oxygen is checked once again after adjustment to evaluate the pillow's efficacy and, as a result, an appropriate modification strategy was identified. In the final test with 40 patients to determine the system's effectiveness, both the frequency and duration of sleep apnea episodes were decreased by more than 50%. The readings of the oxygen desaturation from the pulse oximeter present inside the smart-pillow system and the readings of polysomnography (PSG) were categorized using an SVM classifier to evaluate the system's accuracy. This classification accuracy hovered around 90%. This was a wise decision in the area of sleep apnea diagnosis because the study not only focused on identifying sleep apnea but also on its recovery. Nocturnal oximetry is used by (Marcos et al.; 2009) to identify OSA. Here, SaO2 values were employed rather than SpO2 readings. This study aims to evaluate the diagnostic efficacy of conventional statistical pattern recognition methods for OSA. When using nonlinear and spectral input features, the signals of SaO2 were chosen as the basis. These properties were assessed using classifiers based on Logistic regression and K-nearest neighbors, as well as classifiers based on linear and quadratic discriminant analysis. This study showed that the ability of SaO2 data to diagnose OSA is improved by the use of the latest classification algorithms and signal processing methods.

Respiratory instances are manually evaluated overnight for sleep apnea diagnosis. This procedure takes a long time and costs money. (Nikkonen et al.; 2021) used data from nasal pressure-airflow, thorax respiratory effort, thermistor airflow, and peripheral blood oxygen saturation to train an LSTM neural network for scoring respiratory events automatically. As recurrent neural network architecture enables the categorization or labeling of each sample point in a sequence, hence this design was chosen for this study. The LSTM architecture was used because it is best suited to handle both short and long sequences while maintaining important details throughout the process. This approach evaluates each respiratory occurrence so that humans can examine the automated scoring if necessary. The validation between the manual scoring and the automated scoring was carried out epoch by epoch. This was accurate to 88.9%. All of these researches identified sleep apnea using other indicators in addition to the conventional SpO2 and ECG signals. These investigations also make it clear that sleep apnea may be detected at home by observing respiration abnormalities in all forms.

3 Methodology

This section outlines the strategy that will be used to analyze the data and get insightful information from it. Additionally, it outlines the technologies and techniques that will be employed throughout the project in order to provide the intended outcomes and respond to the submitted research question. In this study, knowledge will be extracted from the selected database using the knowledge discovery in databases (KDD) technique. The decision to utilize KDD instead of the Cross-Industry Standard Data Mining technique (CRISP-DM) was made because, in CRISP-DM, the project deployment is important in terms of business application which is not the case with KDD. This is in line with our study.

The selection of the right data for the intended research is the first stage in this methodology. The following phase is data preparation and transformation, which involves cleaning and standardizing the data before feeding it to the model. The cleansed data is then subjected to data mining models. The results are examined in the last phase to assess the model's effectiveness.

3.1 Tools used and process flow

The technologies and process flow that are utilized to detect sleep apnea using SpO2 and ECG readings are listed in this section. The wfdb module for Python has been used to extract these signals from a dat file that contains digitized ECG signals along with the apn files containing the annotations. The deep learning algorithms are trained using the Google Colab platform. The procedure to be used for this research is shown in Figure 2 below.

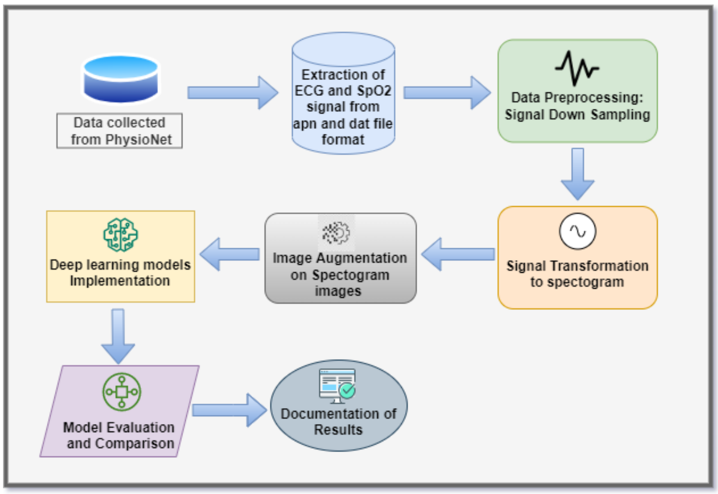

Figure 2: Flow diagram for the process followed

3.2 Data Selection

The dataset used in this study is a public dataset namely Apnea-ECG Database (Penzel et al.; 2000), which is accessible via PhysioNet (Goldberger et al.; 2000). Eight of the 70 people in this database had SpO2 signal data along with the ECG data. The length of the readings for each patient last between seven and ten hours. In this dataset, 13 individuals are considered normal with an AHI 5, 6 individuals are considered mildly apneic with an AHI 30, and 13 individuals are considered severely apneic with an AHI ¿30. Both the ECG and SpO2 data from this database have been used in this study. This data was extracted from PSG in the dat file format along with annotation files and header files having the name and format of associated signal files. Each 60-second block of the signal data is annotated as normal (N) or atrial premature contraction (A).

3.3 Data Preprocessing and Transformation

In this stage, the data is filtered and processed by getting rid of unnecessary readings because the data captured from the devices have noise and errors. Lack of data preprocessing might result in undesired outcomes because data sets may contain inaccurate content or distorted readings. The sampling frequency for the Apnea-ECG database is 100 Hz. The first minute is always annotated as 'N', hence it is not taken into consideration.

The cleaned and processed data is converted in accordance with the project's specifications in this step. To extract characteristics from the signals, signal processing or signal decomposition is done in this step. After decomposition of the ECG signal, it is transformed into spectrogram using Fast Fourier Transform (FFT). These spectograms are converted and saved as images to be fed to the deep learning model. These images were augmented and the whole dataset was split into train and test sets in the ratio of 80% and 20%.

3.4 Data Mining and Evaluation

After being converted, the data is analyzed using data mining techniques to extract important information to do the classification for apnea or non-apnea. The methodologies that are suggested in this study include CNN and DenseNet121 + CNN which is a hybrid neural network.

Sleep Apnea detection involves the classification of signals to determine whether a particular portion of the reading is an apnea minute or non-apnea minute. Finally, accuracy is used to assess the model output. Confusion matrix, F-Score, and other matrices are also taken into account while assessing the model. However, accuracy is one of the key assessment measures in this study since it aims to properly categorize individuals with sleep apnea.

4 Design Specification

To identify sleep apnea, several studies have used deep learning techniques like ANN, CNN, and Deep Belief Network (DBN) as well as machine learning methods like SVM, Random Forest, and KNN. These algorithms have produced useful outcomes. While the characteristics from the original data that will be input to the machine learning algorithm must be accurately estimated. This calls for enough domain expertise in the area of study. However, including more features might occasionally have a negative impact on the model's performance. Due to this, several researchers have worked to identify the greatest attributes in relation to the issue.

Unsupervised learning-based deep learning algorithms may automatically extract characteristics from the raw data, in contrast to feature-based approaches. Therefore, in addition to its frequent application in computer vision and speech recognition, deep learning may also be employed efficiently in sleep analysis. Deep learning models are used in this study for the above-mentioned reasons.

Local minima is an issue in common sleep apnea detection algorithms like MLP and CNN, however, they are mitigated in DBN due to the pretraining step. (Mostafa et al.; 2017) analyzes basic ECG data to identify sleep apnea using an advanced Deep Belief Network. DBNs may be thought of as stacks of Restricted Boltzmann Machines that train the network with a greedy method. Layer-by-layer learning and fine-tuning of the generating weights occur in a DBN. In DBN, each of the RBN layers learns the entire input, in contrast to CNN, where each layer first learns a straightforward pattern before being concatenated. This facilitates the network's autonomous learning of characteristics from the raw data. This concept provided the motivation for using deep learning techniques to automatically learn different features without having a lot of domain expertise.

4.1 Deep Recurrent Neural Network

For large time-series data, Deep Recurrent Neural Networks (DRNN) have shown to be successful. The efficiency of several machine learning, as well as deep learning-based models in the identification of sleep apnea from the data collected from single-lead ECG, has been compared by (Bahrami and Forouzanfar; 2022). The top machine learning and deep learning algorithms for ECG signals were assembled and put to the test in this study. These hybrid models produced better results than the separate models. The hybrid algorithms ZFNet-GRU, VGG16-LSTM, and ZFNet-BiLSTM had the greatest results across all evaluation criteria. LSTM, BiLSTM, and GRU are DRNNs among these ZFNet, while VGG16 is a CNN-based method. With this inspiration, the combination of different DRNN like DenseNet121 + CNN and baseline CNN is used in order to get desired results.

5 Implementation

The Apnea ECG dataset used for this study contains annotations for all the readings. Each 60-second block of the annotation data is labeled as normal (N) or atrial premature contraction (A). For running and executing the model, the Google Colab platform was used which was associated with GPU Tesla T4. The whole dataset was uploaded to google drive and the drive mount feature was used in the colab platform. This data is read using python's waveform-database package (WFDB). All the necessary libraries including Tensorflow, Keras, SkLearn, and wfdb were imported at the beginning. The dataset was explored to find the annotations and the signal associated with it. These signals were divided into 1-minute parts and converted into spectogram as the data contains annotations for each minute. The spectogram sample of the above signal is shown in the Figure 3 below.

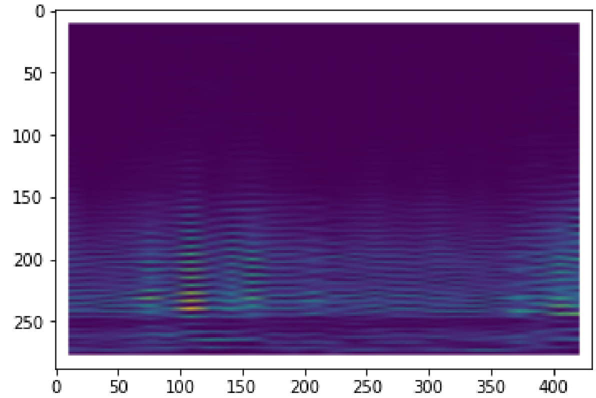

Figure 3: Spectogram for the signal

These components have been visualized to represent each class more precisely using dimension reduction techniques like principle component analysis (PCA). Hence, these spectograms were then converted into Eigenvalues which can be reshaped into a matrix. The primary components have been considered for each class that makes up 70% of the variability. These principle components formed using eigenvalues are shown in Figure 4 below.

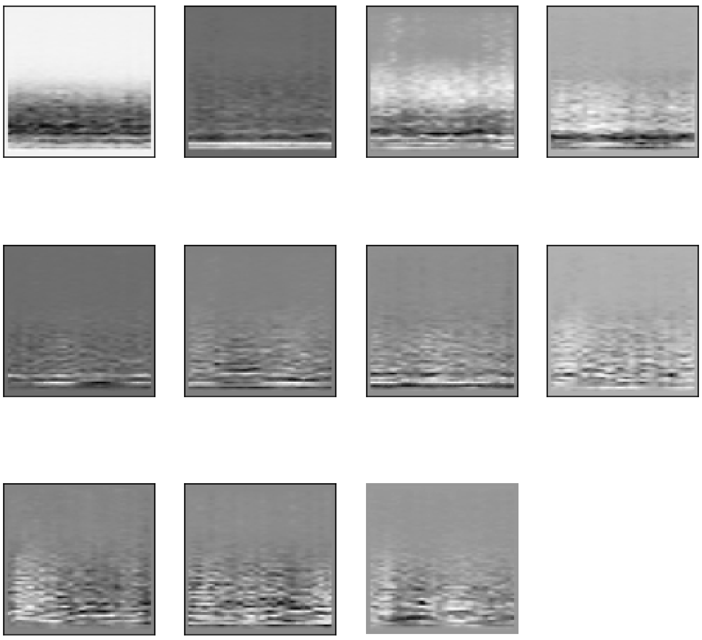

Figure 4: EigenImages for principle components

The Figure 5 shows the average eigen images for both Apnea and non-apnea type along with the difference between normal and apnea patient averages.

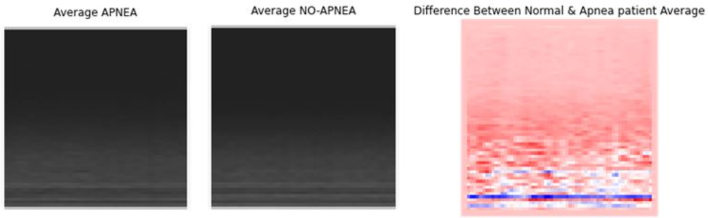

Figure 5: Average of Eigen images and their difference

5.1 Model 1: Baseline Convolutional Neural Network

Four convolutional layers, each followed by a batch normalization layer and a max-pooling layer, are designed as a baseline model. All convolution layers employ a 3 X 3 kernel size with padding one to maintain the height and breadth of their input feature maps, just like VGG blocks. In the convolution layer, there are 32 filters and channels. Due to the fact that this is a binary classification problem, the output layer must have one neuron with sigmoid activation to output the likelihood of one class versus the other. In order to fasten learning and enhance generalization, Batch normalization is frequently used in each convolutional layer. The model summary is shown in Figure 6

```
In [15]: print(model.summary())

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 224, 224, 32)      896

 batch_normalization (BatchN  (None, 224, 224, 32)     128
 ormalization)

 max_pooling2d (MaxPooling2D  (None, 112, 112, 32)     0
 )

 conv2d_1 (Conv2D)           (None, 112, 112, 64)      18496

 batch_normalization_1 (Batc  (None, 112, 112, 64)     256
 hNormalization)

 max_pooling2d_1 (MaxPooling  (None, 56, 56, 64)       0
 2D)

 conv2d_2 (Conv2D)           (None, 56, 56, 128)       73856

 batch_normalization_2 (Batc  (None, 56, 56, 128)      512
 hNormalization)

 max_pooling2d_2 (MaxPooling  (None, 28, 28, 128)      0
 2D)

 conv2d_3 (Conv2D)           (None, 28, 28, 128)       147584

 batch_normalization_3 (Batc  (None, 28, 28, 128)      512
 hNormalization)

 max_pooling2d_3 (MaxPooling  (None, 14, 14, 128)      0
 2D)

 global_average_pooling2d (G  (None, 128)              0
 lobalAveragePooling2D)

 dense (Dense)               (None, 1)                 129

=================================================================
Total params: 242,369
Trainable params: 241,665
Non-trainable params: 704
```

Figure 6: Summary of baseline Convolutional Neural Network

This model's training accuracy rises steadily over time until it reaches 89%, while the validation accuracy ranges between 72–83%. The training loss continues to decrease linearly until it approaches almost zero but the validation loss stalls after only 13 epochs after reaching its minimum. The charts for accuracy and loss are displayed in Figure 7

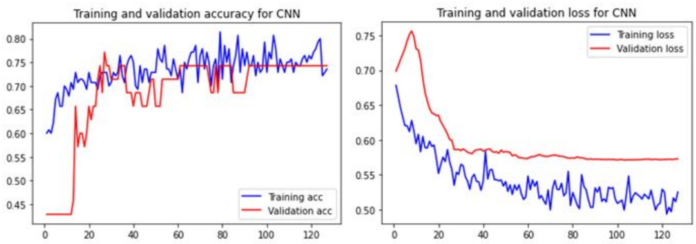

Figure 7: Results of baseline CNN

These plots showed a little overfitting which is due to fewer training samples in comparison to parameters. Following this same model has been applied to the images post data augmentation. The results of the baseline CNN with augmented images are shown Figure 8

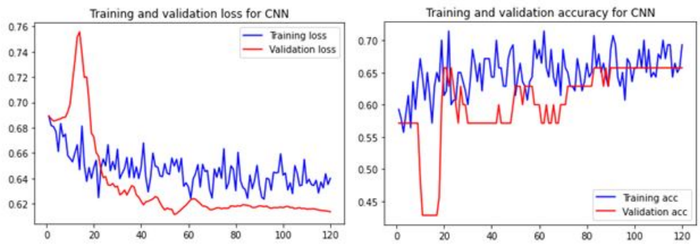

Figure 8: Results of baseline CNN with augmentation

The learning curves of the model with augmented images are not better than in the previous experiment as data augmentation adds some noise to the data in this case. Hence the model without data augmentation is considered.

5.2 Model 2: DenseNet121 + CNN

In this model, a Convolutional Neural network is designed and added on top of the DenseNet121 network. The pretrained DenseNet121 model was used along with the designed CNN model to get better results. The designed CNN model contains 2 Conv2D layers and these two layers are concatenated in the next step. The CNN model description is shown in Figure 9

The results of this model are shown in Figure 10. The problem of overfitting has been addressed to a great extent in this model.

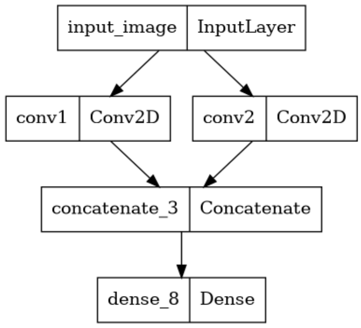

Figure 9: Hybrid CNN model

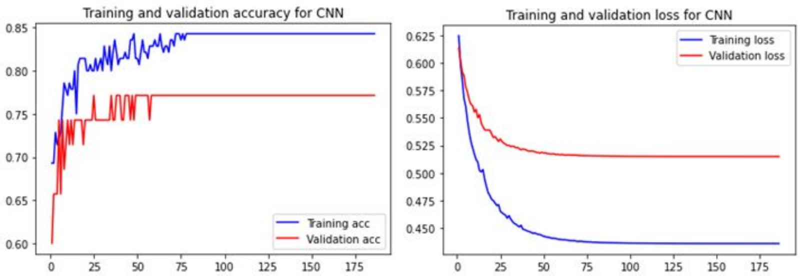

Figure 10: Learning curves for Hybrid CNN model

After a little fine-tuning of the model by freezing and unfreezing some of the layers, the accuracy on the train data came to a little less than 1.00 while the test accuracy came out to be around 0.73.

6 Evaluation

Accuracy is the ability of a classification model to accurately classify the data. In medical situations, it is crucial that an individual who is unwell be accurately classified; but, if some individuals who are not ill are also recorded as positive, there may be a problem. Therefore, it may be claimed that accuracy is more significant in medical instances.

As many researchers in the past have used accuracy to show the effectiveness of various models. Therefore, accuracy becomes crucial to be considered as an evaluation factor. Additionally, a measure that included sensitivity, specificity, and accuracy was employed by (Mostafa et al.; 2017). This suggests that each of the aforementioned indicators is equally crucial for the measurement and detection of sleep apnea.

6.1 Baseline Convolutional Neural Network

The Baseline Convolutional Neural Network designed for this research gave an overall accuracy of 73% while the F1 score was 0.75. The other evaluation metrics are shown in Figure 11. The confusion matrix shown in Figure 12 indicates that around 70% are truly classified as Apnea and around 74% were truly classified as Non-Apnea.

	precision	recall	f1-score	support
0.0	0.68	0.71	0.69	24
1.0	0.77	0.74	0.75	31
accuracy			0.73	55
macro avg	0.72	0.73	0.72	55
weighted avg	0.73	0.73	0.73	55

Figure 11: Evaluation of Baseline CNN

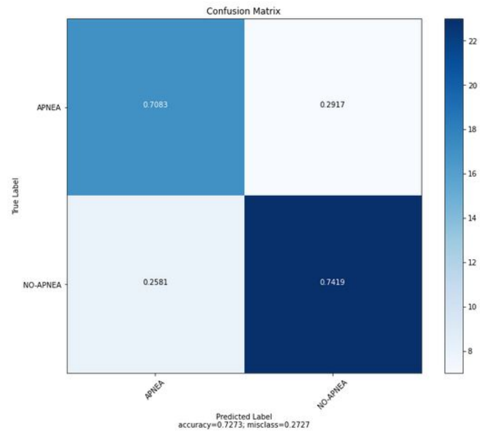

Figure 12: Confusion Matrix for Baseline CNN

6.2 Baseline Convolutional Neural Network after Data Augmentation

In this case, the data augmentation technique was applied to the data before feeding it to the CNN model. But the accuracy decreased to 69% as data augmentation added a little noise to the signals data. Hence, the above model was considered for final evaluation.

6.3 DenseNet121 + CNN

The Hybrid Convolutional Neural Network having DenseNet121+CNN, designed for this research, gave the training accuracy of around 85% but this model worked exceptionally after fine-tuning it. Hence, the fine-tuned model was considered for primary evaluation.

6.4 DenseNet121 + CNN after fine-tuning

The Hybrid CNN was fine-tuned by freezing and unfreezing a few layers. The result after fine-tuning was better than previous results. Here the training accuracy was around 99% and the validation accuracy was around 77%. The overall accuracy of this model after fine-tuning was 81.82% while the F1 score was 0.84. The other evaluation metrics are shown in Figure 13. The confusion matrix shown in Figure 14 indicates that 73.08% are truly classified as Apnea and 89.66% were truly classified as Non-Apnea.

	precision	recall	f1-score	support
0.0	0.86	0.73	0.79	26
1.0	0.79	0.90	0.84	29
accuracy			0.82	55
macro avg	0.83	0.81	0.82	55
weighted avg	0.82	0.82	0.82	55

Figure 13: Evaluation of Hybrid CNN with DenseNet121

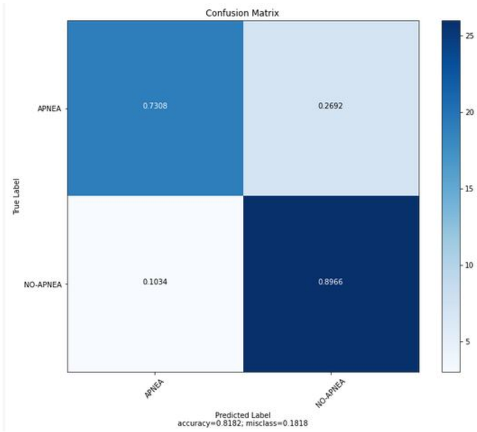

Figure 14: Confusion Matrix for Hybrid CNN with DenseNet121

6.5 Discussion

The best model in this study was found to be DenseNet121+CNN when fine-tuned. The Baseline CNN model designed for this research also gave good accuracy. The results of both the models are compared in the Table 1. The problem of local minima occurs in most of the ECG signal data as found in the previous research. This problem has been tackled using DRNN.

Table 1: Comparison of Models

Model	**Acucuracy**	**F1-score**	**Precision**	**Recall**
Baseline CNN	0.73	0.75	0.77	0.74
DenseNet121 + CNN	0.81	0.84	0.79	0.90

These results can be improved by considering more spectogram images along with higher resolution while feeding to the Neural network. In previous research, spectogram were not used for classifying sleep apnea. Moreover, only a few researchers used both ECG and SpO2 data together for this task. The dataset used for this research has recordings for SpO2 data along with ECG data only for 8 persons among 70 persons. In the future, further research can be done with different datasets having sufficient SpO2 data along with ECG data to get better results.

7 Conclusion and Future Work

This study focuses on the utilization of ECG signal data and SpO2 data to identify sleep apnea. The research question this study tried to address is "How effectively can heart rate and SpO2 data be utilized to predict Sleep Apnea using a smartwatch?". After extensive research, the features of the signals were extracted using the wfdb python package, and hybrid deep learning models such as Baseline CNN and DenseNet121+CNN are suggested to classify the ECG and SpO2 measurements obtained from polysomnography utilizing the Apnea-ECG database in order to predict the likelihood of sleep apnea. The model Baseline CNN gave an accuracy of 73% DenseNet121+CNN has given accuracy of 82%.This study, therefore, proposes improved deep learning models to diagnose sleep apnea at home using a smartwatch with the ultimate goal of simplifying the PSG procedure.

Some architectures require greater processing power and training time. Only 70 people's ECG readings were available in the database utilized for this study, and only 8 of those people also had SpO2 measurements. This is one of the study's limitations. The absence of data from smartwatches is another limitation of this study. The ECG and SpO2 data obtained during the PSG procedure are used in this investigation. In the future, larger datasets with more SpO2 measurements, ECG readings, and data recorded directly from the smartwatch can be used to obtain better results. The research's approach may be used to create specialized applications for smartwatches and smartphones that can detect sleep apnea in homes in real-time.

References

Adnan, M. M., Kassim, N. K., Mohd Rasid, N. M., Ibrahim, H. A., Ramli, R., Ghani, N. and Hanafi, M. H. (2022). Knowledge and attitude of obstructive sleep apnea among dental students in north east penisular malaysia, *Banglad. J. Med. Sci.* **21**(2): 384–390.

Bahrami, M. and Forouzanfar, M. (2022). Sleep apnea detection from single-lead ecg: A comprehensive analysis of machine learning and deep learning algorithms, *IEEE Transactions on Instrumentation and Measurement* **71**: 1–11.

Cen, L., Yu, Z. L., Kluge, T. and Ser, W. (2018). Automatic system for obstructive sleep apnea events detection using convolutional neural network, *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3975–3978.

Chen, X., Xiao, Y., Tang, Y., Fernandez-Mendoza, J. and Cao, G. (2021). Apneadetector: Detecting sleep apnea with smartwatches, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **5**(2).
URL: *https://doi.org/10.1145/3463514*

Fatimah, B., Singh, P., Singhal, A. and Pachori, R. B. (2020). Detection of apnea events from ecg segments using fourier decomposition method, *Biomedical Signal Processing and Control* **61**: 102005.
URL: *https://www.sciencedirect.com/science/article/pii/S1746809420301610*

Ghandeharioun, H. (2021). Online obstructive sleep apnea detection based on hybrid machine learning and classifier combination for home-based applications, *CoRR* **abs/2110.00660**.
URL: *https://arxiv.org/abs/2110.00660*

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K. and Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, *Circulation* **101**(23): E215–20.

Marcos, J. V., Hornero, R., Álvarez, D., del Campo, F. and Zamarrón, C. (2009). Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry, *Medical Engineering & Physics* **31**(8): 971–978.
URL: *https://www.sciencedirect.com/science/article/pii/S1350453309001246*

Mostafa, S. S., Mendonça, F., Morgado-Dias, F. and Ravelo-García, A. (2017). Spo2 based sleep apnea detection using deep learning, *2017 IEEE 21st International Conference on Intelligent Engineering Systems (INES)*, pp. 000091–000096.

Nguyen, H. D., Wilkins, B. A., Cheng, Q. and Benjamin, B. A. (2014). An online sleep apnea detection method based on recurrence quantification analysis, *IEEE Journal of Biomedical and Health Informatics* **18**(4): 1285–1293.

Nikkonen, S., Korkalainen, H., Leino, A., Myllymaa, S., Duce, B., Leppänen, T. and Töyräs, J. (2021). Automatic respiratory event scoring in obstructive sleep apnea using a long short-term memory neural network, *IEEE Journal of Biomedical and Health Informatics* **25**(8): 2917–2927.

Penzel, T., Moody, G., Mark, R., Goldberger, A. and Peter, J. (2000). The apnea-ecg database, *Computers in Cardiology 2000. Vol.27 (Cat. 00CH37163)*, pp. 255–258.

Qatmh, M., Bonny, T., Barneih, F., Alshaltone, O., Nasir, N., Al-Shabi, M. and Al-Shammaa, A. (2022). Sleep apnea detection based on ecg signals using discrete wavelet transform and artificial neural network, *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1–5.

Randerath, W. (2022). Positive airway pressure therapies in central sleep apnea, *in* S. M. Janes (ed.), *Encyclopedia of Respiratory Medicine (Second Edition)*, second edition edn, Academic Press, Oxford, pp. 181–196.
URL: *https://www.sciencedirect.com/science/article/pii/B9780081027233001141*

Romero, H. E., Ma, N., Brown, G. J. and Hill, E. A. (2022). Acoustic screening for obstructive sleep apnea in home environments based on deep neural networks, *IEEE Journal of Biomedical and Health Informatics* pp. 1–1.

Schrader, M., Zywietz, C., von Einem, V., Widiger, B. and Joseph, G. (2000). Detection of sleep apnea in single channel ecgs from the physionet data base, *Computers in Cardiology 2000. Vol.27 (Cat. 00CH37163)*, pp. 263–266.

Sharma, M., Raval, M. and Acharya, U. R. (2019). A new approach to identify obstructive sleep apnea using an optimal orthogonal wavelet filter bank with ecg signals.

Van Steenkiste, T., Groenendaal, W., Ruyssinck, J., Dreesen, P., Klerkx, S., Smeets, C., de Francisco, R., Deschrijver, D. and Dhaene, T. (2018). Systematic comparison of respiratory signals for the automated detection of sleep apnea, *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 449–452.

Walch, O., Huang, Y., Forger, D. and Goldstein, C. (2019). Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device, *Sleep* **42**.

Zhang, J., Zhang, Q., Wang, Y. and Qiu, C. (2013). A real-time auto-adjustable smart pillow system for sleep apnea detection and treatment, *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 179–190.