

IDENTIFICATION OF OBSTRUCTIVE SLEEP APNEA USING ARTIFICIAL NEURAL NETWORKS AND WAVELET PACKET DECOMPOSITION OF THE HRV SIGNAL

Sarah Qasim Ali and Abdulnasir Hossen *

Department of Electrical and Computer Engineering, College of Engineering, Sultan Qaboos University
Oman, P. O. Box 33, Al-Khoudh, 123 Muscat, Oman.

Abstract: The advancement of telecommunication technologies has provided us with new promising alternatives for remote diagnosis and possible treatment suggestions for patients of diverse health disorders, among which is the ability to identify Obstructive Sleep Apnea (OSA) syndrome by means of Electrocardiograph (ECG) signal analysis. In this paper, the standard spectral bands' powers and statistical interval-based parameters of the Heart Rate Variability (HRV) signal were considered as a form of features for classifying the Sultan Qaboos University Hospital (SQUH) database for OSA syndrome into 4 different levels. Wavelet packet analysis was applied to obtain and estimate the standard frequency bands of the HRV signal. Further, the single perceptron neural network, the feedforward with back-propagation neural network and the probabilistic neural network have been implemented in the classification task. The classification between normal subjects versus severe OSA patients achieved 95% accuracy with the probabilistic neural network. While the classification between normal subjects versus mild OSA subjects reached accuracy of 95% also. When grouping mild, moderate and severe OSA subjects in one group compared to normal subjects as a second group, the classification with the feedforward network achieved an accuracy of 87.5%. Finally, when classifying subjects directly into one of the four classes (normal or mild or moderate or severe), a 77.5% accuracy was achieved with the feedforward network.

Keywords: Sleep Apnea; Identification; Classification; HRV; Wavelet Packet Decomposition; Neural Networks.

تحديد انقطاع التنفس الانسدادي أثناء النوم باستخدام الشبكات العصبية الاصطناعية وتحليل حزمة الأطوال الموجية لإشارة HRV

سارة قاسم علي وعبدالناصر حسين *

المخلص: أتاح لنا التقدم العلمي في مجال تقنيات الاتصال بدائلًا جديدة واعدة للتشخيص عن بعد واقتراحات علاجية محتملة لمرضى مختلف الاضطرابات الصحية. ومن ضمن هذه التقنيات القدرة على كشف مرضى متلازمة انقطاع التنفس الانسدادي أثناء النوم عن طريق تحليل إشارة التخطيط الكهربائي للقلب (ECG). اعتبرت في هذه الدراسة قوى النطاقات الطيفية القياسية والإحصائيات المقسمة على فترات زمنية لإشارة تغير معدل نبضات القلب (HRV) سمة من سمات تصنيف وتقسيم قواعد بيانات مستشفى جامعة السلطان قابوس المتعلقة بمتلازمة انقطاع التنفس الانسدادي الى اربع مستويات. تم استخدام حزمة الأطوال الموجية لتقدير نطاقات التردد القياسية لإشارة HRV التي بنيت عليها الدراسة. وقد تم تصميم كل من الشبكة العصبية الحسية الوحيدة، و الشبكة العصبية ذات التغذية الأمامية والانتشار الخلفي، والشبكة العصبية الإحتمالية في مهمة التصنيف. حقق الإختبار دقة 95% في التمييز بين الأصحاء والمرضى من ذوي الأعراض الحادة باستخدام نموذج الشبكة العصبية الإحتمالية، ودقة 95% في التفرقة بين الأشخاص الأصحاء والأشخاص المصابين بالمتلازمة ويعانون من أعراض بسيطة. أما عند تشكيل عينة تشمل مرضى يعانون من أعراض مختلفة الشدة (بسيطة ومتوسطة وشديدة) وعينة أخرى من أشخاص أصحاء فإن نموذج الشبكة العصبية ذات التغذية الأمامية والانتشار الخلفي تمكنت من التفريق بين العينتين بدقة بلغت 87.5%، وتمكن نموذج الشبكة العصبية ذات التغذية الأمامية والانتشار الخلفي من التمييز بين الأنواع الأربعة بدقة بلغت 77.5%.

الكلمات المفتاحية: متلازمة انقطاع التنفس أثناء النوم؛ الكشف؛ التصنيف؛ معدل تغير نبضات القلب؛ تحليل حزمة الأطوال الموجية؛ الشبكات العصبية الاصطناعية.

*Corresponding author's e-mail: abhossen@squ.edu.om



1. INTRODUCTION

A sleep disorder occurs when the pattern of sleep is interrupted repeatedly during sleep (Kumar V 2008). Lack of sleep results in abnormalities in functions of the brain leading to cognitive impairment, changes in mood, low productivity, daytime sleepiness and abnormal hormonal rhythms (Wilson S 2016). Sleep apnea is a chronic disease that affects the health and productivity of individuals (National Heart, Lung and Blood Institute 2016) since it causes abnormal sleep pattern. Obstructive Sleep Apnea (OSA) is the most common type of sleep apnea followed by Central Sleep Apnea (CSA) and Mixed Sleep Apnea (MSA). OSA affects 3~4% and 2% of middle-aged men and women respectively (Lee W *et al.* 2008). Unlike CSA which results from heart failure or brain disorders, where the brain fails to control breathing leading to cessation of all respiratory airflow and movements, OSA results from a repeated process of complete or partial collapse in the upper airways of the respiratory system ranging from few seconds (minimum 10 sec) to minutes despite the ongoing brain efforts for the body to breath. The OSA events may occur more than 30 times and up to 100 times per hour. MSA on the other hand, is a mixture of CSA and OSA in the same individual. (American Academy of Sleep Medicine 2001; National Sleep Foundation 2016). OSA has been related to some serious co-morbidity such as cardiovascular diseases, arrhythmia, strokes, obesity, depression, certain types of hypertension and type 2 diabetes mellitus (Global Leaders in Sleep and Respiratory Medicine 2013; Xie W *et al.* 2014). There are several screening methods used for OSA detection to find evidence of its presence in patients for further evaluation. These methods depend on psychometric and physical evaluations during the routine health check-ups. Polysomnography (PSG) sleep study is the gold standard test for sleep apnea diagnosis. This test requires the patient to sleep in a sleep laboratory while attached using several electrodes to many devices for different biometric measures carried out by qualified sleep physicians overnight. The severity of sleep apnea is commonly determined by an Apnea-Hypopnea Index (AHI) which represents the number of obstructive, central, mixed and hypopnea episodes occurring during an hour of sleep (American Academy of Sleep Medicine 2001). If the AHI ranges between 0-5 apneic episodes during an hour of study time or sleep then the condition is considered normal. An index of 5-15 is considered mild while an index of 15 – 30 is considered moderate and if the index is 30 or above, the subject is considered to have a severe degree of sleep apnea (Global Leaders in Sleep and Respiratory Medicine 2013). Electrocardiography (ECG) is a method used to measure the electrical activity of the heart by placing electrodes on different parts of the body (WebMD 2016). A normal sinus rhythm reflects the normal activity of the heart while

pumping blood to perform the sympathetic and parasympathetic activities (UCDavis Health System 2016). A typical ECG signal is produced when the heart chambers contract and expand to pump oxygenated-blood throughout the body and circulate the desaturated blood to the lungs. Fig. 1 (a) shows a typical ECG signal (Sharma S. *et al.* 2019).

Heart rate (HR) is a simple measurement that indicates the average number of heart beats during a certain time period (usually, a minute). A low HR reflects resting status while high HR indicates stress or exertion (Moore J 2016). Heart Rate Variability (HRV) on the other hand is a measure of the time variability in milliseconds between consecutive beats or correspondingly in the instantaneous HR. In other words, variation analysis of instantaneous HR versus time axis. HRV is sometimes called the R-R interval (RRI) analysis, where R is the peak point of the QRS complex in the ECG wave, or the Inter-Beat-Interval (IBI) analysis. When the individual is at rest, high HRV is favorable while low HRV is observed at an active or stressed state. HRV has been used as a measurement to assess overall cardiac health and reflect the state of the Autonomic Nervous System (ANS) activities (Hamilton G. *et al.* 2019).

The ANS is the involuntary division of the nervous system and consists of autonomic neurons that conduct impulses from the central nervous system (brain and/or spinal cord) to glands, smooth muscles and cardiac muscles (DanTest Clinicians Team 2016). The role of the ANS is to continuously fine-tune the functions of organs and organs systems to maintain internal stability and balance. ANS has two main components called the Sympathetic and Parasympathetic Nervous Systems (SNS and PSNS respectively). The SNS triggers the fight or flight response leading to increased heart rate, blood pressure and sweating, and pupil dilation etc. On the other hand, the PNS complements the operations

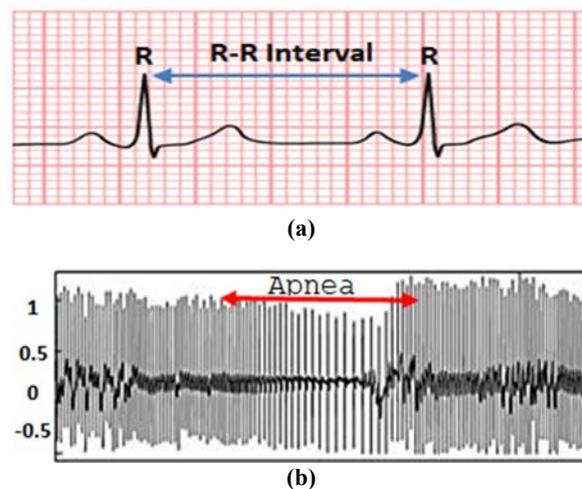


Figure 1: (a) Normal ECG signal.
(b) ECG signal at Apnea Episode.

performed by the SNS and triggers the rest and digest response where the opposite behavior occurs. When the airway is partially or completely obstructed; the heart rate changes and hence ECG signal alters. When the oxygen level decreases in the body during sleep apnea event the heart cells receives less oxygen and hence the heart rate is reduced and the R-R interval increases between consecutive beats as shown in Fig.1 (b) (Sharma S. *et al.* 2019).

This alters the brain's sleep and wakes the brain for immediate action. The brain responds by sending strong tones to the respiratory system to increase breathing speed. The later increases the heart rate suddenly and hence increases the blood pressure in order to pump more blood to compensate for the lack of oxygen.

In frequency-domain analysis, signals can be represented in a graph that shows how much (energy) of the signal lies within given frequency bands over a range of frequencies. The well-known Fourier methods such as the Fast Fourier Transform (FFT) implementation of the Discrete Fourier Transform (DFT) are usually used for identifying the available spectral content for both stationary and non-stationary signals (Polikar R 2011). However, for non-stationary signals, the frequency content varies with time and hence DFT based methods fail to provide the time-related information at which those frequencies occur. Moreover, it can only reflect the frequencies that are present in the signal but not when they were present. Since most of the physiological signals like the ECG, etc. are non-stationary signals; time-frequency analysis such as the Short-Time Fourier Transform (STFT) and Discrete Wavelet Transform (DWT) are used as alternatives to the Fourier analysis when estimating the available PSD content (Polikar R 2011). The classification features used in this research depend on the Power Spectral Density (PSD) at different frequency levels estimated by implementing the Discrete Wavelet Packet Decomposition (DWPD) method. This is to overcome the resolution related problems of the STFT. The discrete wavelet decomposition utilizes various mother wavelets of different scales to be able to adapt to fast and slow changes in the analyzed signal (Polikar R 2011). The wavelet decomposition method is implemented using filter banks (Misiti M *et al.* 1996). A set of high pass and low pass filters allow the signal to be decomposed reaching a certain decomposition level in which the signal can be further analyzed, de-noised or compressed (Misiti M *et al.* 1996). The PSD calculation is done by mathematical modulation to the filters output coefficients (Sysel P *et al.* 2008). The DWT allows only the low-frequency components of the low pass filter to be analyzed to further levels as they are thought to be the ones that carry important information (Misiti M *et al.* 1996). However, DWPD allows both outcomes of the filters (low and high-frequency) to be further decomposed. The later emphasizes the outliers, edges and transient signals

which are crucial to tackle the OSA episodes. Hence, the DWPD leads to finding more desirable features for classification applications.

In time domain analysis, a signal instance's real values are visualized. A time domain graph shows how a signal changes with time. OSA episodes can be analyzed by observing the cyclic length variability of the HRV (a.k.a. RRI) signal. The term NN is used sometimes in place of the RR to emphasize that normal beats are being processed. There exist multiple well-established features that are normally used to analyze the beat-to-beat intervals (Mietus J *et al.* 2002). These features include: the Root Mean Square of Successive Differences (RMSSD), the Standard Deviation of Successive Differences (SDSD), the Standard Deviation (Std.) of entire RRI signal, the mean of the entire RRI signal, the NNx family measures which include (NN of $x \leq 50$: NN50, NN30, NN20...etc.), and finally the pNNx family measure.

2. ECG DATABASE

It has been suggested by some researchers, that the uniqueness of the data sets affects the classification results of the different proposed methods, hence similar results cannot be obtained using different datasets (Lado M *et al.* 2011). In this research the ECG signals were collected from the Sleep Laboratory of the Physiology Department of the Sultan Qaboos University Hospital (SQUH) while performing PSG studies for 80 subjects. These records were obtained from 20 normal subjects ($0 < \text{AHI} < 5$), 20 mild subjects ($5 < \text{AHI} < 15$), 20 moderate subjects ($15 < \text{AHI} < 30$) and 20 severe subjects ($\text{AHI} \geq 30$). The records were divided into two groups: a training set and a test set; where both of the sets comprise of 10 normal ECG signals and 10 Apneic ECG signals from each of the mentioned groups (total of 40 signals per set). Features were extracted from both sets and the training set is used to train the neural networks for classification purpose while the test set was used to check the classification performance.

3. RRI EXTRACTION

The HRV signal is generated by finding the R-to-R Intervals (RRI) from the original ECG signals as declared in Fig. 1 (a). The proposed method in (Al Ghunaimi B 2003) was used to generate the RRI data. In order to accurately identify those R peaks, QRS detection is to be carried. The QRS detector that was used in (Al Ghunaimi B 2003) is a part of the Physionet tools available in the Physionet website and is based on the Pan-Tompkins Algorithm. Intervals corresponding to Normal-to-Normal peaks are extracted. The generated RRI data could contain false intervals, missed intervals and/or ectopic intervals.

False RR intervals were removed by setting the lower and upper limits of its values to ones found in normal subjects which are typically around 400-2000 milliseconds. Removing of outliers was achieved using a 41-points Moving Average Filter (MAF). Re-sampling at 1 Hz and substituting of missed peaks were then achieved by simple linear interpolation implemented by MATLAB (Al Ghunaimi B 2003). Re-sampling and estimation of missed value are intended to generate an equally spaced RRI data and preserve the temporal sequence that is necessary for the frequency domain analysis. At this point it can be assumed that a clean RRI data sampled at 1 Hz containing no missed or outlier values was generated.

4. SPECTRAL FEATURES

In this research, the RRI signals were decomposed using discrete wavelet packet decomposition up to 9 levels in order to define the VLF band precisely according to the standard definition (VLF starts at 0.0033 Hz). At the ninth level, the signal would have been decomposed into 512 frequency bands including the low-frequency and high-frequency bands.

The PSD summation of these bands provides the total power spectrum of the signal, while the PSD summation by grouping according to certain frequency bands would provide us the desired features for the analysis. The discrete RRI signals are sampled at $F_s=1$ Hz and the maximum spectral frequency of RRI is found by $F_s/2=0.5$ Hz. Using DWPD, the 0.5 Hz is decomposed into 512 bands ($9 \text{ levels} = 2^9$), while each band covers $0.5/512 = 0.0009765625$ Hz. Therefore, the standard spectral bands of the decomposed RRI are covered as shown in Fig. 2. In addition, the spectral values near zero Hz have the most energy content that dominates other spectral values. Therefore, we intend to exclude those spectral values below 0.0033 Hz; hence, the first three bands of the decomposed RRI signal were ignored and the PSD of the VLF was estimated starting from the 4th-band at 0.00390625 Hz up to 0.040039063 Hz.

Fig. 3 shows the spectral powers summation of the different bands defining each feature. Furthermore, three other features extracted from the power ratios of the VLF, LF and HF features are calculated to form a total of six spectral features. These ratios are described as in Fig. 4.

In Fig. 5, the power of the HF band is sketched for both normal and severe subjects for the training set. The values of the HF band at normal subjects are higher than those of severe subjects. However, some severe subjects have powers that overlap with normal subjects resulting in difficulties to classify and hence a combination of features are to be investigated (The University of Nottingham 2017).

5. STATISTICAL FEATURES

In this work, multiple features were calculated from the RRI signals in time domain using MATLAB software. These features include the Root Mean Square of Successive Differences (RMSSD) which is calculated by finding the square root of the mean of the successive differences between adjacent NNs. The Standard Deviation of Successive Differences (SDSD) where the Standard Deviation of Average NN Intervals (SDANN) is calculated over a short period of 5-minutes or so to measures the change in heart rate due to cycles longer than five minutes. The Standard Deviation (Std.) of entire RRI signal, the mean of the entire RRI signal, The NNx family measures which include (NN of $x \leq 50$: NN50, NN30, NN20 etc.), where it represents the number of pairs of successive NN's that differ by more than $x=50, 30, 20 \dots$ etc. milliseconds and finally the pNNx family measure which includes the NNx measure divided by the total number of NNs of the signal. To sum up, the features are: the RMSSD, pNN15, pNN20, pNN30, pNN50, NN15, NN20, NN30, NN50, SDANN and Standard deviation of the entire signal which equate to eleven features.

In order to observe a relation between time domain and frequency domain features, the values of RMSSD, NN50, and pNN50 features were normalized by dividing them to the total power of the RRI signal as shown in Fig. 6. Similar behavior of Fig. 5 was observed where the power of the high-frequency band for severe apnea level subjects decrease and the same features' values increase for normal subjects. This allows us to use these features in a similar analogy as spectral features for representing the parasympathetic activity of the Autonomous Nervous System for instance. Of course, variations are also present in some of the OSA patients where their spectral or time domain features exhibit similar behavior as normal subjects making it more difficult to differentiate between the cases using a single feature.

6. NEURAL NETWORKS

The Artificial Neural Networks (ANNs) were designed based on the rudimentary understanding of the biological nervous system back in the 1950s to help solving and computing any arithmetic or logical statement (Hagan M *et al.* 1996). The ANNs are a collection of computational units called neurons composed of inputs with weights, biases and transfer functions to perform the thresholding act and produce an output. The three networks used in this work are of supervised learning type; this is because we already have information about the subject's original condition (normal, mild, moderate, and severe). In supervised learning, the training set consists of

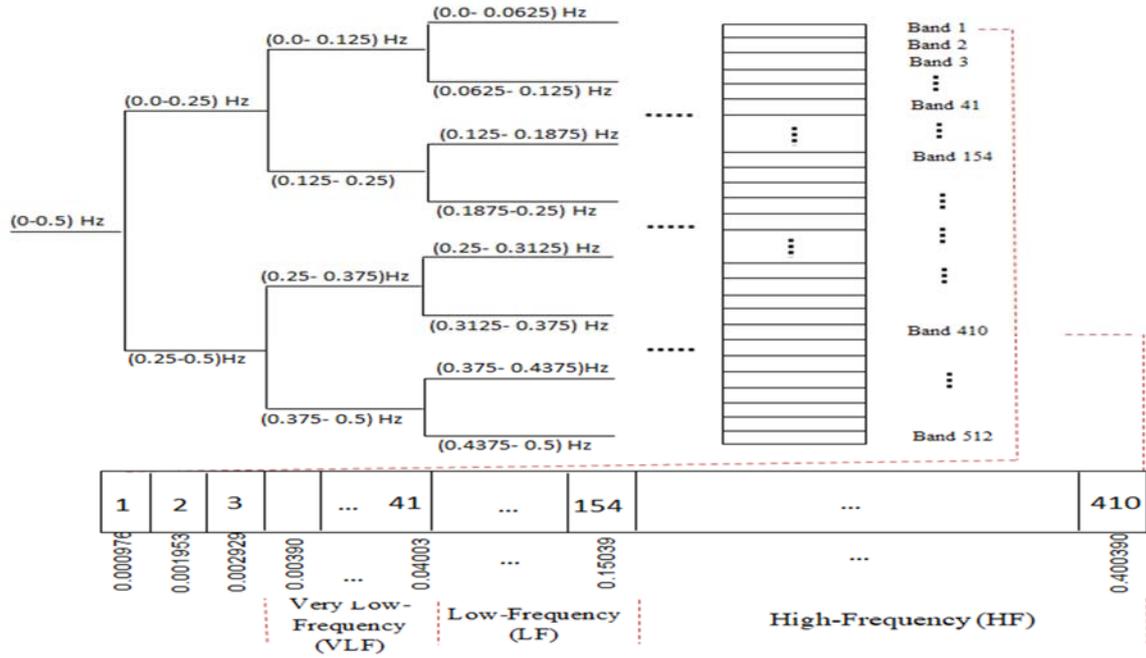


Figure 2. Frequency bands of RRI packet decomposition into 9 levels.

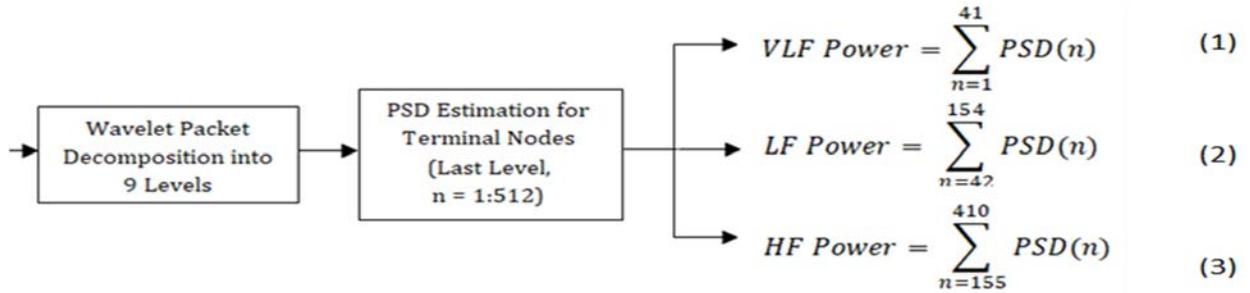


Figure 3. Calculation of the three main features.

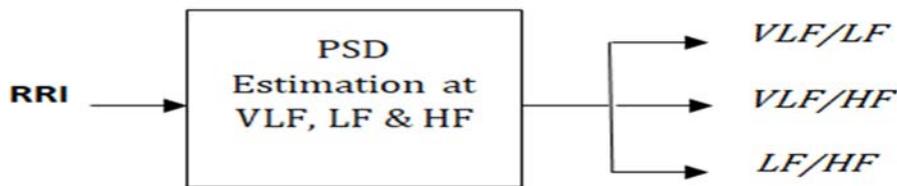


Figure 4. Main three spectral ratio features.

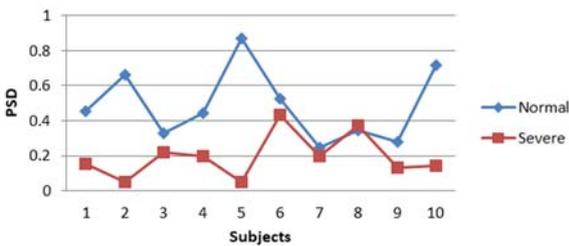


Figure 5. Power spectral density (HF band) between normal and severe subjects.

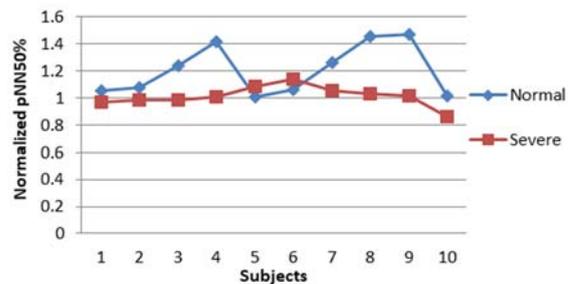


Figure 6. Normalized pNN50 % values for normal and severe subjects.

multiple training examples each is a pair of input (feature) and the desired output (target). The supervised learning algorithm then tries to process and analyze the training examples to produce an inferred function (classification boundary region), which can be used to map new examples. If the network training and the learning algorithm are performed well, an optimal scenario would be generated in which the network will generalize and be able to determine the class label (target) of unseen instances correctly (Mohri M *et al.* 2012). This section introduces the three artificial neural networks' models used in this research and describes their learning algorithms and transfer functions briefly.

The perceptron network refers to multiple-inputs single-neuron network. It is considered the building block of a more complicated network called the Feed forward Neural Network. Usually the single-layer perceptron is used for binary and linear classification. The feed forward network is one of the most commonly used artificial neural networks and is typically composed of multiple perceptrons aligned in layers. It is considered the first type of ANN's that was used to solve non-linear problems where the perceptron has failed and is considered powerful networks that can almost approximate any function (Hagan M *et al.* 1996). Probabilistic Neural Networks (PNN's) are widely used classification and pattern recognition. When employed for classification problems, the class probability of input is estimated and the class with the highest probability is selected as the output class. This means that the input belongs to the class that provides the highest probability when introduced to that input. The design of a PNN is straightforward and extremely less complicated than other multilayer networks. This is because it does not depend on weights learning and hence does not require training. A PNN network generalizes well and is guaranteed to converge to a Bayesian classifier (simple well-studied conditional probabilistic classifier) providing the correct probability when presented with enough training data samples. A PNN consists of several sub-networks, each of which is a Parzen window pdf estimator for each of the classes. The inputs are the set of measurements/features and are used as centers for the radial basis (Gaussian functions) of the second layer. The third layer performs an average operation of the outputs from the second layer for each class and a final voting is performed by the third layer selecting the largest value to determine the associated class label.

In this research, the perceptron network is composed of one neuron of multiple inputs for one output classification, and two neurons for four outputs classification embedded with hard limit transfer function. The feedforward network is composed of two hidden layers each of five neurons (total 10) of the tangent-sigmoid transfer function and an output layer of one or two neurons according to the number of outputs (similar to perceptron layer) with pure

linear transfer function. The probabilistic network is composed of input neurons changing according to the number of inputs with radial basis transfer function. Figs. 7 and 8 show the connections of the layers for the feed forward and probabilistic neural networks.

The performance of the training step and testing step of the neural networks are investigated by calculating well-known performance metrics such as the specificity, sensitivity and accuracy. The specificity reflects the number of accurately diagnosed healthy subjects while the sensitivity reflects the number of the accurately identified patients. The accuracy is a measure of both of the correctly classified patients and normal among the total experiment set. In Fig. 9, the actual meaning for each performance metric can be observed while the following equations reveal how they are calculated.

$$\text{Specificity } (\%) = \frac{TN}{TN + FP} \cdot 100 \quad (1)$$

$$\text{Sensitivity } (\%) = \frac{TP}{TP + FN} \cdot 100 \quad (2)$$

$$\text{Accuracy } (\%) = \frac{TP + TN}{TP + FP + TN + FN} \cdot 100 \quad (3)$$

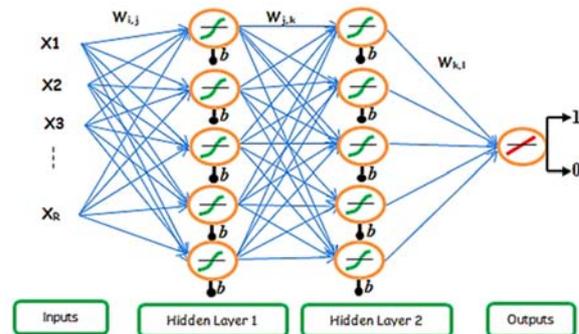


Figure 7. Feedforward neural network layer connections.

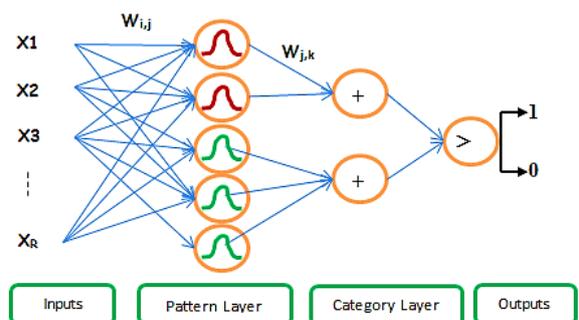


Figure 8. Probabilistic neural network layer connections.

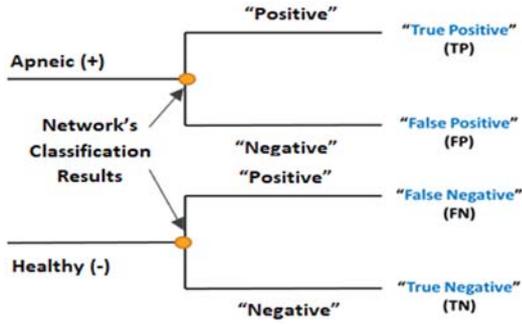


Figure 9. Networks performance metrics.

7. IMPLEMENTATION

In this work, four different versions of classification have been carried out using the perceptron network, the feedforward with back-propagation network and the probabilistic neural network. These classification versions include:

- Version 1: Normal vs. Severe Classification
- Version 2: Normal vs. Mild Classification
- Version 3: Normal vs. Patient (Mild, Moderate, Severe) Classification
- Version 4: Normal vs. Mild vs. Moderate vs. Severe Classification.

The discrete wavelet packet decomposition was implemented to extract power estimations of the RRI signals of SQUH ECG database at the classical frequency bands with Bi-orthogonal mother wavelet at 9 levels decomposition. Four different schemes of training and test data are selected:

- Scheme 1: Original first half data for trial: Original second half data for the test.
- Scheme 2: Original second half data for trial: Original first half data for the test.
- Scheme 3: Large set simulated from first-half data for trial: Original second half data for test.
- Scheme 4: Large set simulated from second-half data for trial: Original first half data for test.

The networks in scheme 1 and 2 are trained by the features of original data sets containing 10 subjects for each of the normal, mild, moderate and severe OSA states. While in scheme 3 and 4, the networks are trained using a large simulated set generated from the original data sets. The large training set was acquired by generating randomly, 1000 uniformly distributed feature sets between the maximum and minimum values of each feature from the original extracted spectral features. The networks testing and performance computation were carried by the remaining equivalent-size original data set of 10 subjects in each case.

The statistical time-domain features are used with classification version number 4 only as it is the most complicated classification version and hence investigation with different features may increase the chance of enhancing the performance.

8. RESULTS AND DISCUSSION

Figures 10 and 11 display the actual classification accuracies of every network performing version-1 classification of normal vs. severe OSA conditions and version-3 normal vs. patient OSA conditions respectively. The spectral features used are numbered in Figs. 11 and 12 as: (1. VLF, 2. LF, 3. HF, 4. VLF/LF, 5. VLF/HF, 6. LF/HF). The patient-class refers to a set including all the apneic levels of mild, moderate and severe subjects where the AHI index is exceeding 6 and up to 100 or more. The features were selected based on the highest training results they provided before testing the networks as well as for their lowest discrepancy among other schemes.

Figures 12 and 13 summarize the results of the networks at versions-2 and version-4 classification by emphasizing the spectral features generating the highest performances. It can be noticed that version-4 classification was the most difficult and the results are extremely poor. Herein, it can be concluded that spectral features alone are not enough for this complicated task. However, when using statistical time-domain features with the feedforward network; the result is extremely improved reaching an accuracy of 77.5% using two features only. The features are of pNN20 & pNN30 respectively. This shows that the statistical features extracted from the RRI signal at the time domain used with the feedforward network were able to generate a non-linear decision surface that helped classifying those subjects unlike the probability estimates or linear classification boundaries generated by the probabilistic and perceptron networks respectively. It can be observed that the feedforward with back-propagation neural network achieved high accuracies at different classification versions demonstrating the power and consistency of the network. Further, the specificity and sensitivity percentages are very close to each other at the different classification versions of the different networks which implies that only few unique cases were miss classified. Table 1, summarizes the results of the networks at the different classification versions by emphasizing the features generating the highest performances using scheme 1 of training. The classification performance between normal individuals and severe OSA patients was highest using the probabilistic network where the network was able to classify the subjects into either normal or severe conditions at 95% accuracy. This is an interesting result and shows how powerful the probabilistic network can be even when trained with a few sample points; it was still able to generalize well. Moreover, this accuracy was achieved when using a single spectral feature which is the very low frequency power feature. The test set included 20 subjects (10 normal and 10 severe) meaning that the network only missed the correct classification of one severe subject at a sensitivity of 90%.

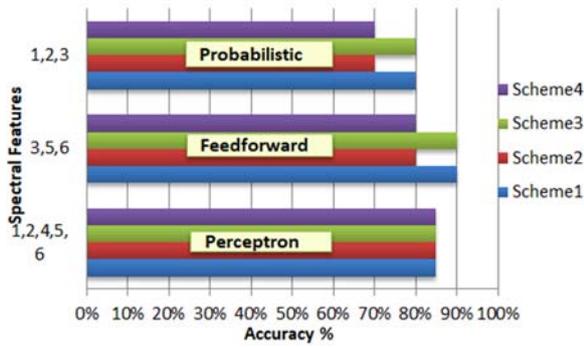


Figure 10. Networks Classification Results for Normal vs. Severe.

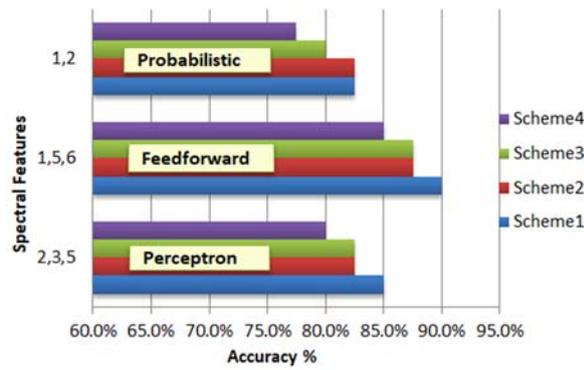


Figure 11. Networks classification results for normal vs. patient.

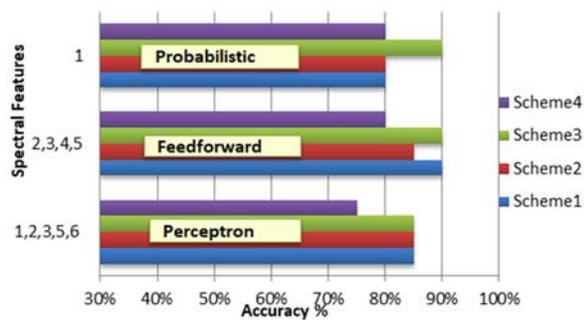


Figure 12. Networks classification results for normal vs. mild.

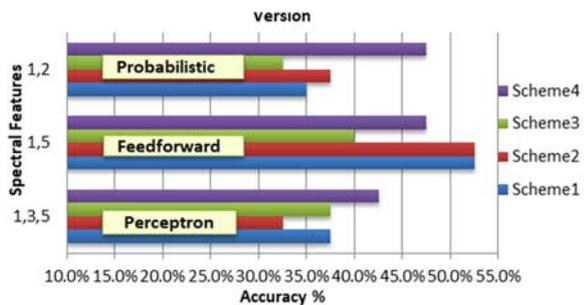
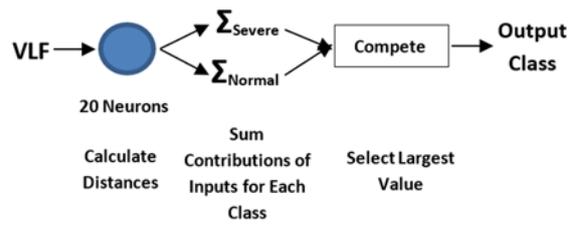
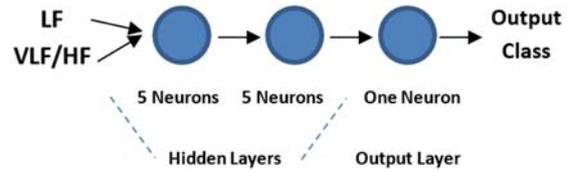


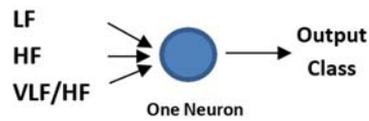
Figure 13. Networks classification results for normal vs. mild vs. moderate vs. severe.



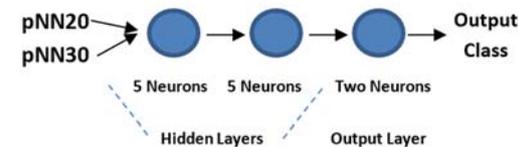
(a) Probabilistic Neural Network: Version-1: Normal vs. Severe Classification (95% Accuracy)



(b) Feed-Forward Neural Network: Version-2: Normal vs. Mild Classification (95% Accuracy)



(c) Perceptron Neural Network: Version-3: Normal vs. Patient Classification (85% Accuracy)



(d) Feedforward Neural Network: Version-4: Normal vs. Mild vs. Moderate vs. Severe Classification (77.5% Accuracy)

Figure 14. Neural Networks Topology for the Four Classification Versions.

The second classification version between normal and mild OSA patients witnessed its best performance at 95% accuracy by the feedforward network among all training schemes (original and large training sets). Hence, it can be considered a robust network algorithm, since many examples were introduced to the network that included different types of apneic episodes exhibiting different cessation occurrences and durations. Moreover, when comparing this result to other researches available in the literature, it is noticed that many of them had their highest results when dealing with small sets (~ 30 subjects where

~10 are normal and ~20 are severe). The latter makes our results acceptable and the employed database was efficient when used for classification. The third classification version between normal and OSA patients on the other hand, achieved performance accuracy of 87.5% using three different features (VLF+VLF/HF +LF/HF) by the feedforward network and achieved 85% accuracy by the perceptron network with the three different features (LF+ HF+ VLF/HF). Finally, it can be observed that the highest achieved test set accuracy of 77.5% in the fourth classification version between normal, mild, moderate and severe OSA subjects was accomplished by the feedforward network using a combination of two statistical features of pNN20 and pNN30 respectively.

The representations in Fig. 14 demonstrate each type of network topology with different features as inputs for the classification versions corresponding to the best accuracies listed in Table 1.

For the third version, the perceptron network is sketched oxygenated blood although its performance accuracy was 85% and less than that of the Feedforward network 87.5%, in order to show the topology of the perceptron network.

9. CONCLUSION

Obstructive sleep apnea (OSA) is a common breathing-related sleep disorder affecting individuals of different age groups, genders and origins. It is characterized by short-duration cessations in breathing during sleep due to the collapse of the upper airway. OSA is associated with major co-morbidities such as cardiovascular diseases, arrhythmias, strokes, obesity, depression, certain types of hypertension and diabetes. The golden and reliable standard test for the detection of OSA is a polysomnographic sleep study.

However, this test is time/labor-consuming, expensive and cumbersome. Analysis of a Heart Rate Variability (HRV) signal that is obtained from an Electrocardiograph (ECG) signal in time or frequency domain is an effective, non-invasive and promising method for the detection of OSA.

In this research, single perceptron, feedforward with back propagation and probabilistic artificial neural networks are investigated for their performance in classifying SQUH database subject's severity degree against four classification versions. The highest achieved accuracy of 95% was obtained when using VLF feature with the probabilistic neural network for normal vs. severe classification (version-1). The feedforward neural network achieved an accuracy of 95% as well when classifying normal versus mild OSA patients at a combination of LF and VLF/HF ratio features. In Version 3, the feedforward network achieves 87.5% accuracy using three features VLF and VLF/HF and LF/HF for normal vs. patient (including: mild, moderate and severe subjects in one group) classification. In the same version, the perceptron network achieved the highest performance accuracy of 85% using LF along with HF and VLF/HF ratio combination of features. Finally, for OSA severity degree classification (version-4) statistical time-domain features provided the highest accuracy of 77.5% when using a combination of pNN20 and pNN30 features with the feedforward neural network.

The results are considered promising since the networks only used a maximum of three features to provide such results. Some of the limitations of this work include the ECG database size, neural networks training processing-time especially the feedforward network, feature dimensions (combinations of inputs). Future research recommendations are to be on investigating deep learning neural networks.

Table 1. Highest classification results of the networks at training scheme 1.

Network Type	Classification Version	Features	Specificity	Sensitivity	Accuracy
Probabilistic	Normal vs. Severe	VLF	100%	90%	95%
Feedforward	Normal vs. Mild	LF+ VLF/HF	90%	100%	95%
Perceptron	Normal vs. Patient	LF+ HF+VLF/HF	80%	86.67%	85%
Feedforward	Normal vs. Patients	VLF+VLF/HF +LF/HF	90%	86.67%	87.5%
Feedforward	Normal vs. Mild vs. Moderate vs. Severe	pNN20+pNN30	80%	76.67%	77.5%

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

FUNDING

No funding was received for this project.

REFERENCES

- Al Ghunaimi B (2003), Statistical Signal Characterization and Sub-band Decomposition for Heart Rate Variability Analysis in Patients with Obstructive Sleep Apnea (Masters Dissertation, Sultan Qaboos University).
- American Academy of Sleep Medicine (2001), International classification of sleep disorders, revised: Diagnostic and coding manual. Chicago, Illinois: *American Academy of Sleep Medicine*.
- DanTest Clinicians Team, “ANS Balance Assessments”. (2016). Sited from: https://www.dantest.com/dtr_ans_overview.htm.
- Global Leaders in Sleep and Respiratory Medicine. (2013), Sleep Apnea Facts and Figures. ResMed Corp. San Diego, CA, USA.
- Hagan M. T. Demuth, H.B. Beale M.H., and De Jesús O (1996), Neural Network Design (Vol. 20). Boston: PWS publishing company.
- Hamilton G., Chai-Coetzer C., (2019), “Update on the Assessment and Investigation of Adult Obstructive Sleep Apnea” *Australian Journal of General Practice*, 48 (4): 176-181.
- Kumar V.M. (2008), Sleep and sleep disorders. *Indian Journal of Chest Diseases and Allied Sciences* 50(1): 129.
- Lado M.J. Vila X.A. Rodríguez-Liñares L. Méndez A. J. Olivieri D. N. and Félix P (2011), Detecting sleep apnea by heart rate variability analysis: assessing the validity of databases and algorithms. *Journal of Medical Systems* 35(4): 473-481.
- Lee, W., Nagubadi, S., Kryger, M. H., & Mokhlesi, B. (2008). Epidemiology of Obstructive Sleep Apnea: a population-based perspective. *Expert Review of Respiratory Medicine*, 2(3), 349–364.
- Mietus, J E; Peng, C.K.; Henry, I.; Goldsmith, R.L.; Goldberger, A.L. (2002). “The pNNx files: re-examining a widely used heart rate variability measure”. *Heart*. 88: 378–380.
- Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. M. (1996). *Matlab Wavelet Toolbox User’s Guide*. The MathWorks. Inc. Mass.
- Mohri M., Rostamizadeh A., Talwalkar A., (2012) *Foundations of Machine Learning*. The MIT Press ISBN 9780262018258
- Moore J. Heart Rate Variability Course (2016), “Heart Rate Variability vs. Heart Rate”, Sited from: <https://hrvcourse.com>.
- National Heart, Lung and Blood Institute. What Is Sleep Apnea?, (2012). Sited from: <https://www.nhlbi.nih.gov/health/health-topics/topics/sleepapnea>
- National Sleep Foundation. (2016). “Sleep Apnea”. Sited from: <https://sleepfoundation.org/sleep-disorders-problems/sleep-apnea>
- Polikar, R. (2011). *The Wavelet Tutorial Parts 1-3*. Second Edition. Sited from: https://cseweb.ucsd.edu/baden/Doc/wavelets/polikar_wavelets.pdf.
- Sharma S., Raval M., Acharya U.R., 2019, A new approach to identify obstructive sleep apnea using an optimal orthogonal wavelet filter bank with ECG signals’ *Informatics in Medicine Unlocked*, 16, 100170: 1-13.
- Sysel, P., & Misurec, J. (2008). Estimation of power spectral density using wavelet thresholding. In S. Kartalopoulos, A. Buikis, N. Mastorakis, & L. Vladareanu (Eds.), *WSEAS International Conference Proceedings Mathematics and Computers in Science and Engineering* (No. 7). World Scientific and Engineering Academy and Society
- The University of Nottingham, Practice Learning Resources: Cardiology Teaching Package. (2017). http://www.nottingham.ac.uk/nursing/practice/resources/cardiology/function/normal_duration.pp
- UCDavis Health System. ‘Heart Rate’, (2016). Sited from: www.ucdmc.ucdavis.edu/sportsmedicine/resources/heart_rate_description.html
- WebMD, ‘Electrocardiogram’, (2013). Sited from: [webmd.com/heartdisease/electrocardiogram/Wilson, S. \(2013\). *Sleep Disorders \(Oxford Psychiatry Library\)*. 2nd ed. Oxford, United Kingdom: Oxford University Press, USA, p.1](http://webmd.com/heartdisease/electrocardiogram/Wilson,S.(2013).SleepDisorders(OxfordPsychiatryLibrary).2nded.Oxford,UnitedKingdom:OxfordUniversityPress,USA,p.1)
- Xie, W., Zheng, F., & Song, X. (2014). Obstructive sleep apnea and serious adverse outcomes in patients with cardiovascular or cerebrovascular disease: a PRISMA-compliant systematic review and meta-analysis. *Medicine*, 93(29): 1-8.