ABSTRACT

Sleep is indispensable to human beings and it plays a vital role in the strengthening of immune system and vital functions. Even though awareness on the benefits of sleep is little among people, it is a fact that one-third of human life is spent in sleep. Sleep is characterized as Rapid Eye Movement (REM) sleep and Non-REM (NREM) sleep. REM and NREM sleep alternates in a cyclic manner, that leads to good quality of sleep. When this cycle is disrupted due to pathological or psychological difficulty, quality of sleep gets affected and leads to health disorders. Generally, sleep studies are conducted using Polysomnography (PSG) equipment at sleep laboratories in hospitals. Polysomnography test is a complex procedure that involves nocturnal recording of various bio signals and requires a person to stay at hospital. Sleep apnea is a sleep breathing disorder that is clinically described as the occurrence when breathing is paused for more than 10 seconds and this occurs for more than five times during sleep. The major types of sleep apnea are Obstructive, Central and Mixed. Obstructive Sleep Apnea (OSA) is a commonly occurring sleep breathing disorder that occurs due to collapse of upper airway muscles.

The goal of this research work is to develop a methodology for screening of Obstructive Sleep Apnea through the analysis of Heart Rate Variability of Electrocardiogram (ECG) signals during sleep. The objectives of this research work are to generate Heart Rate Variability (HRV) signals from sleep ECG, characterize the HRV signals with their features and develop computational methods to identify the occurrence of OSA using those features. ECG signals are taken from Apnea-ECG records of PhysioNet database. The records consist of ECG signals from three categories of Sleep Apnea: Severe, Mild and Moderate. Record durations vary from 7 to 8 hours of nocturnal sleep. Annotations are provided in the database for every one

minute of sleep record as normal condition or under apnea condition. A smaller part of this research work aims at developing an embedded heart rate variability logger to aid in sleep studies.

The proposed methodology consists of four modules: Heart rate variability extraction module, Feature extraction module, Feature selection module and Classification module. In the Heart rate variability extraction module, HRV time series is derived by detecting successive ‘R’ peaks of sleep ECG signal. Two approaches are considered for detection of ‘R’ peaks: Pan-Tompkins algorithm and a Shanon Energy based Hilbert Transform algorithm. These two algorithms are evaluated using the sleep records and their annotations provided in the database.

In the Feature extraction module, various time domain, frequency domain and statistical features are derived from the HRV signals. The feature set selected for this research work consists of 22 features that comprises of Mean of RR intervals, Standard Deviation, Median, Mean Absolute Deviation of RR intervals, pNN50, pNN20, pNN30, RMSSD, Short term variability, Long term variability, Ratio of short term to long term variability, Low frequency power, High frequency power, Low to high frequency ratio, Largest Lyapunov Exponent, Hjorth parameters, Inter Quartile Range, Kurtosis, Skewness, Detrended Fluctuation Analysis and Sample Entropy.

The function of Classification module is to classify the HRV segments as under normal sleep condition or apnea condition by utilizing the derived HRV features. Classifiers proposed in this work are k Nearest Neighbour (kNN), Support Vector Machines (SVM) with Polynomial, Pearson Coefficient and Radial Basis Function kernels, J48 and Random Forest classifiers. All the proposed classifiers are supervised classifiers. Hence, feature sets along with their target labels are provided to the during the learning phase. The performance of classifiers are evaluated using three

approaches : Evaluation using training set, 10 fold Cross Validation (CV) and 60:40 split percentage testing.

Feature selection module consists of dimensionality reduction approaches to reduce the dimension of the proposed feature sets. In this work, Correlation based Feature Selection method (CFS) and Principal Component Analysis (PCA) are applied for dimensionality reduction. CFS finds a subset of original features while PCA creates a new set of reduced features from the original features. These approaches identify a smaller number of features that have the highest variance towards their target labels. Finally, the reduced feature sets are applied on to the classifiers and evaluated for their performance. Among the proposed classifiers, kNN classified the HRV signal with an accuracy of 100% during training set evaluation and achieved a classification accuracy of 90 to 95% during CV and split percentage testing.

The proposed HRV logger was tested with ECG signals from database and ECG signals recorded in real time. It achieved a maximum sensitivity of 97.7% and specificity of 95.56% in ‘R’ peak detection. The average recorded percentage error of detection was 1.8 %. From literature, it is found that several studies had been conducted on identifying OSA from the Apnea-ECG database with a classification accuracy of 70% to 100% using various classification algorithms and features. But, limited number of studies have been conducted on feature dimensionality reduction for this database. Hence, this research work is significant in proposing a combined methodology of Shannon Energy based generation of HRV time series, feature reduction using Correlation based Feature selection method and use of kNN classifier for the screening of OSA.

These findings provide new information on the dimensionality

reduction approach in screening OSA through ECG signals, which might aid in developing a robust Computer Assisted Diagnosis model for diagnosing OSA in the future. This method might be used for conducting longitudinal studies to evaluate Cardio Vascular disorders and Neural disorders.