

Apneic Events Detection Using Different Features of Airflow Signals

FATMA ZEHRA GOGUS*, AND GÜLAY TEZEL*

RECEIVED ON 11.08.2018 ACCEPTED ON 30.10.2018

ABSTRACT

Apneic-event based sleep disorders are very common and affect greatly the daily life of people. However, diagnosis of these disorders by detecting apneic events are very difficult. Studies show that analyzes of airflow signals are effective in diagnosis of apneic-event based sleep disorders. According to these studies, diagnosis can be performed by detecting the apneic episodes of the airflow signals. This work deals with detection of apneic episodes on airflow signals belonging to Apnea-ECG (Electrocardiogram) and MIT (Massachusetts Institute of Technology) BIH (Bastons's Beth Isreal Hospital) databases. In order to accomplish this task, three representative feature sets namely classic feature set, amplitude feature set and descriptive model feature set were created. The performance of these feature sets were evaluated individually and in combination with the aid of the random forest classifier to detect apneic episodes. Moreover, effective features were selected by OneR Attribute Eval Feature Selection Algorithm to obtain higher performance. Selected 28 features for Apnea-ECG database and 31 features for MIT-BIH database from 54 features were applied to classifier to compare achievements. As a result, the highest classification accuracies were obtained with the usage of effective features as 96.21% for Apnea-ECG database and 92.23% for MIT-BIH database. Kappa values are also quite good (91.80 and 81.96%) and support the classification accuracies for both databases, too. The results of the study are quite promising for determining apneic events on a minute-by-minute basis.

Key Words: Apneic Event Detection, Feature Extraction, Classification, OneR Attribute Eval Feature Selection, Random Forest.

1. INTRODUCTION

SBD (Sleep Breathing Disorders) generally include OSA (Obstructive Sleep Apnea) which may affect approximately 2% of females and 4% of males [1]. Events occurred during night as called apneic events. These events include apnea and hypopnea and they are characterized with cessation of airflow for at least 10 seconds [2]. In the day time following the night in which apnea and hypopnea events occur, subjects

experienced situations such as sleepiness, tiredness, carelessness, low concentration [3]. These situations can cause to traffic and work accidents, depression, impaired learning etc. Moreover, apnea-induced sleep disorders can trigger heart disease, cardiovascular disfunction, hypertension and myocardial infarction [2-3]. Therefore, diagnosing and treatment of SBD is important. In a clinical environment, diagnosing is generally made by

Authors E-Mail: (fzgogus@ktun.edu.tr, gtezel@ktun.edu.tr)

* Department of Computer Engineering, Konya Technical University, 42075 Konya, Turkey.

polysomnography. Through polysomnography apneic events are specified and number of events are counted [4]. Calculation of apneic events number is important for diagnosis but the identification of events and determination of time intervals at which events occur are also important especially for treatment [4]. For determination of time intervals, apneic episodes of signal must be detected and separated from the normal episodes. In this way, both events can be identified and the time when events occur can be determined.

This study focuses on the detection of apneic episodes on a minute-by-minute basis. So, we can also learn the minutes at which apneic events occur. Although several signals have been used in many studies [5-8], airflow signals were selected for this study and detection of apneic episodes was made on the airflow signals. Because, these signals give the primary indication of apneic events [3]. Detection processes generally contain analysis of airflow signals and classification of the signals according to their specific characteristics. Therefore, characteristics of signals must be well defined. In order to define the signal characteristics, various feature sets including classical feature set, amplitude feature set and descriptive model feature set were produced in this study. And then these sets were used with a classifier RF (Random Forest) to detect the apneic events. In the study, it was aimed how successfully the apneic episodes were detected and

which feature sets or features were more effective in this success.

2. MATERIALS AND METHOD

In this study experiments were performed with nasal airflow signals obtained from two separate databases, Apnea-ECG and MIT-BIH Polysomnographic [9-10]. These databases can be accessed on the PyhsioNet website [11]. Using records obtained from these databases, this study was realized in four distinct stages; preprocessing, feature extraction, feature selection and classification. The block diagram of the study is shown in Fig. 1.

2.1 DATABASE

2.1.1 Apnea_ECG database

This database is described in Penzel et. al. [9]. Data were recorded in Philips University in Marburg, Germany. In this database, since only 5 airflow signals contain the apneic events. These 5 signals were selected for this study.

Reference annotation file associated with each signal was created by a sleep expert to indicate the presence or absence of apnea during 1 minute. Each minute is labeled as 'A' when apnea was in progress at the beginning of the associated minute, otherwise this minute is labelled as 'N' [2].

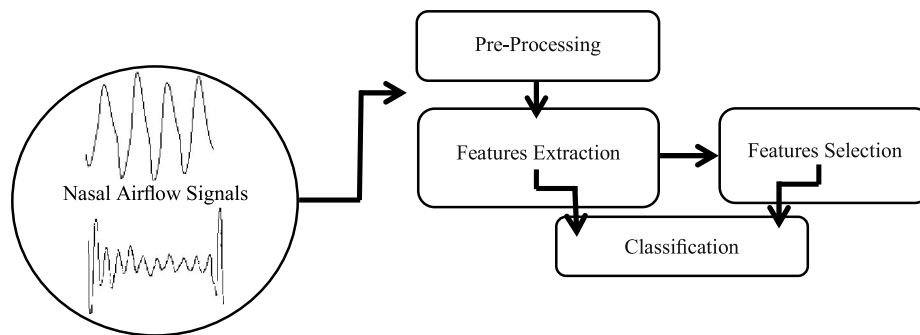


FIG. 1. BLOCK DIAGRAM OF STUDY

In this database, apneas were associated with 90% drops in airflow. Also, minutes containing hypopneas (defined as intermittent drops in airflow below 50%, accompanied by drops in oxygen saturation of at least 4%, and followed by compensating hyperventilation) were scored as minutes containing apnea [11].

2.1.2 MIT-BIH Polysomnographic Database

The MIT-BIH Polysomnographic Database consists of multiple physiologic signal recordings in sleep. Recordings were gathered in BIH Sleep Laboratory for evaluation of chronic OSA syndrome [10]. Also, database include annotation file associated with signals. File contains apnea information and sleep stage. Annotation was made by sleep experts according to the presence or absence of apneic events during 30 second periods [11]. Because rules of apneic event scoring was not specified for this database, it was accepted general apnea, hypopnea scoring rules as 90% drops in airflow signal for apnea, 50 or 70% drop in airflow signal for hypopnea [12].

2.2 Pre-Processing

Proper pre-processing can cause to be obtained of better results in all signal processing area. Filtering and segmentation processes constituted as pre-processes of our study.

In the various studies carried out by airflow signal, the filters whose frequency range is varying from 0.01 and 15Hz have been used. Bandpass filter with cut-off frequency of 0.01-0.15 was used in study of Koley and Dey [3]. Diaz et. al. [4] and Huang et. al. [13] used 0.05-5Hz bandpass filter. Selvaraj and Narasimhan [14] applied lowpass filter with frequency of 0.7 Hz. As can be seen, there is no definite frequency range for these signals in

literature studies. The reason for this can be explained by Varady et. al. [15] as that airflow signals are specific to the applied sensor and they can change during measurements because of sensor or patient movements. Therefore, different filters with various frequency range were tried during the study. As a consequence of the experiments, frequency values were determined that would allow the apnea region to appear more clearly on the signal. While fourth-order Butterworth bandpass filter with cut-off frequency of 0.1-0.15Hz were applied to airflow signals of Apnea-ECG database, second-order Butterworth bandpass filter with cut-off frequency of 0.01-0.5Hz were used for airflows of MIT-BIH database. Fig. 2 shows airflow signals which are filtered by bandpass filter with various frequency values.

As seen from Fig. 2, areas in the red circle are apneic events and they are seen more clearly in determined frequency ranges.

Generally, it is expected that the airflow signals will show a sinusoidal pattern. The period of these signals represents respiratory cycle and amplitudes are always normalized between -1 and 1 [8]. As a usual, all the signals used in our work were normalized between -1 and 1.

After the filtering and normalization processes, both Apnea-ECG database airflow signals and MIT-BIH database airflow signals were segmented into 1-min episodes. As mentioned before, reference annotation file associated with each signal of Apnea-ECG database were created by a sleep expert to indicate the presence or absence of apnea during 1 minute [2,9]. Same annotations were made by sleep experts according to the presence or absence of apneic events during 30 sec periods by MIT-BIH polysomnographic database. For Apnea-ECG

database, if airflow signals were segmented into 30 sec segments, we cannot decide that how many seconds of apneic events are included in the segment and which segment will be called apneic without sleep experts. Therefore, we can only use 1 minute annotations for this database. However, we can convert 30 seconds annotations of MIT-BIH polysomnographic database to 1 minute annotations. If one of the 30 seconds of the 1-minute segments is named apneic, we can assume that this minute is apneic. If none of the 30 seconds of the 1-minute segments include apneic events, this minute is called normal. For these reasons, air flow signals were segmented at 1-minute episode.

Reference annotation file of MIT-BIH database were rearranged according to 1 minute. Segmented minutes of two databases were evaluated according to whether they contained apneic event or not.

For MIT database, although some episodes included the air flow cessation, they were ignored because the patient is awake at that time. Airflow cessations are not valid for apnea scoring clinically when the patient is awake.

At the end of the pre-processing stage, 2513 and 1699 episodes with 1-minute length were created from Apnea-ECG database and MIT-BIH database, respectively.

2.3 Feature Extraction

When apneic events occur at night, airflow signals exhibit different characteristics. In order to separation of apneic episodes from normal, changing characteristics must be specified. In this study, classic features set, amplitude features set and descriptive model features set were created to define different characteristics of the airflow signals for both databases separately. Subsequently, features that were more sensitive to apneic events were identified.

2.3.1 Classic Features Set (Set-1)

Set-1, generally include descriptive statistic features is called as classic because the features in this group are commonly used in various signal processing applications [3,16-17]. It enable to us a brief knowledge about data and their distributions. The 1-minute airflow signal episodes generally include both apneic event

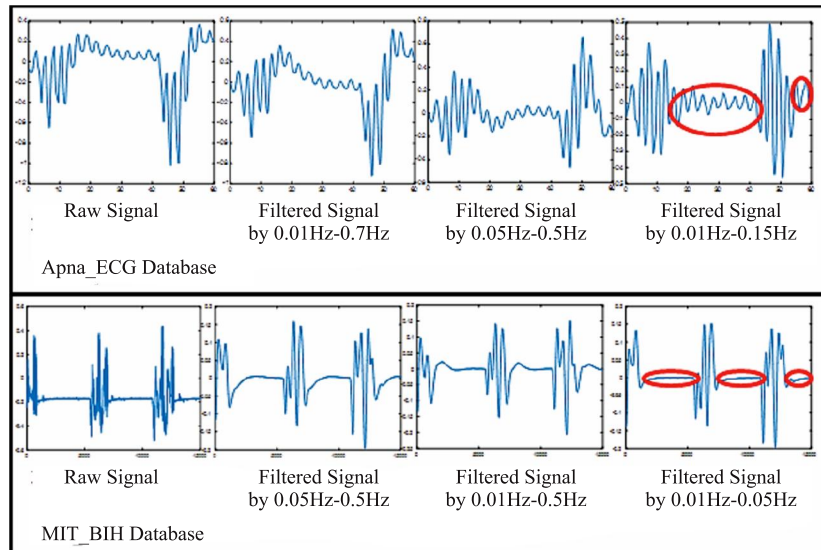


FIG. 2. AIRFLOW SIGNALS FILTERED BY DIFFERENT FREQUENCY RANGE

patterns and normal patterns. If there are no apneic events in the episodes, a more periodic structure is observed without any major changes. This case facilitates the separation of normal and apneic episodes. In this study, 22 features were extracted to find out the changes within 1-minute episodes. Features in this data set are described in Table 1. The “//” sign in the tables represents the distinction between features. For example, S9 represents the mean absolute deviation, and S10 represents the median absolute deviation.

2.3.2 Amplitude Features (Set-2)

According to AASM criteria [12] and previous studies, amplitudes of airflow signals [3,15] or peaks of the signals [13] vary during apneic events. Events are called as apnea, if there is a decrease in the peak or amplitude of signal by

$\geq 90\%$ [12]. AASM guidelines [12] define events as hypopnea if there is a decrease in the peak or amplitude of signal by $\geq 50\%$ or by $\geq 30\%$. This decrease should be accompanied by ≥ 3 or ≥ 4 oxygen desaturation. It is also necessary that these reductions last for at least 10 seconds. Based on this definitions of apneic events, Set-2 was created. Table 2 shows the features of Set-2.

In order to extract amplitude features, firstly, peak and trough points and corresponding times of these points in the airflow signal were determined for any apneic episode as shown in Fig. 3. Then, features between A1-A15 were calculated using these peaks and trough points. Amplitude was calculated as shown in Equation (1). Where $p(i)$ and $t(i)$ are i th peak and i th trough, respectively.

$$A[i] = p[i] - t[i] \quad (1)$$

TABLE 1. SET-1: CLASSIC FEATURES

Feature Number	Description of Features
S1	Energy
S2	Minimum
S3	Maximum
S4	Mean
S5	Standard Deviation
S6	Variation
S7	Skewness
S8	Kurtosis
S9 // S10	Mean Absolute Value // Median Absolute Value
S11	Root Mean SquareLevel
S12	Peak Magnitude to RMS Ratio
S13 / S14	Hjorth Parameters : Mobility / Complexity
S15	Sum of Absolute Signal Value
S16	Sum of Absolute 1st Order Derivative Signal Value
S17 // S18 // S19	Mean // Standard Deviation // Variation of Absolute 1st Order Derivative Signal Value
S20	Mean of Signal Envelope
S21	Standard Deviation of Signal Envelope
S22	Mean of Signal Envelope to Standard Deviation of Signal Envelope ratio

Since hypopnea is expressed by a 30 or 50% amplitude and a peak reduction, A2, A3, A5, A6, A8 and A9 features were used to define these apneic events. In the same way, A4, A7, A12 and A13 features was used to detect apneic events called apnea. When we set the baseline as the mean of the highest 20% of the peaks or signal

amplitudes, peaks values and troughs values less than 50% of baselines, duration between the peaks values higher than 50% of baselines and amplitude values were obtained as in Fig. 4. For example, A3, A6, A10 and A11 were computed using these important points and intervals.

TABLE 2. SET-2: AMPLITUDE FEATURES

Features Number	Description of Features
A1	Average Amplitude
A2	The total number of signals peak values less than 70% of the baseline peak
A3	The total number of signals peak values less than 50% of the baseline
A4	The total number of signals peak values less than 10% of the baseline (The baseline peak is the mean value of the highest 20% of the signal peaks)
A5	The total number of signals amplitude values less than 70% of the baseline amplitude
A6	The total number of signals amplitude values less than 50% of the baseline
A7	The total number of signals amplitude values less than 10% of the baseline (The baseline amplitude is the mean value of the highest 20% of the signal amplitudes)
A8 // A9	Mean // Maximum of time between peaks higher than 70% of baseline peak
A10 // A11	Mean // Maximum of time between peaks higher than 50% of baseline peak
A12 // A13	Mean // Maximum of time between peaks higher than 10% of baseline peak
A14	Average of absolute differences between two successive amplitudes of 60 sec signal
A15	Average of absolute differences between two successive mean values of 60 sec signal over 4 second interval

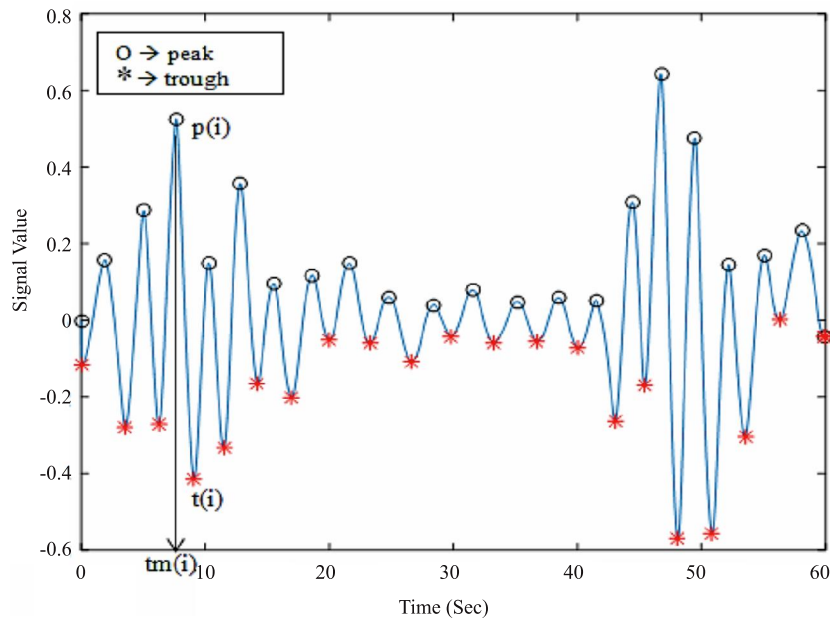


FIG. 3. PEAK, TROUGH POINTS AND TIME OF POINTS

2.3.3 Descriptive Model Features (Set-3)

Descriptive models give information related to dataset by simulating them. This dataset consists of 17 features named as M1-M17 extracted from parametric normal (Gaussian) model, non-parametric kernel model, histogram model and clustering model of every 1-minute signal episode.

Nowadays, probability distribution plays an important role in various research areas [18]. The probability distribution is a mathematical function and it gives the probability of each value of the variable or gives the probability that the variable falls in a particular interval [19]. The most frequently used models, the parametric normal (Gaussian) model and non-parametric kernel model, were preferred in this study because data can be described efficiently by a model consisting of probabilistic distributions for facilitating analysis and classification [20-23]. In addition to these models, the histogram model was used since histograms give graphical representation

of data distribution. These three models were applied to peak values of every 1-minute signal episodes. Fig. 5 summarizes this process. The upper part of the Fig. 5 shows the normal, histogram and kernel model of any apneic episode. The lower part of the Fig. 5 shows the kernel functions for both normal and apneic episodes. Also calculation of M11 and M12 features mentioned in Table 3 is seen from the lower part of Fig. 5. As seen from the lower parts of the Fig. 5, when normal episodes were considered, the peak values corresponding to maximum function values are higher than those of apneic episodes. In addition it can be said that when compared to normal episodes the area between -0.2 and 0.2 belongs to apneic episodes is greater .

Moreover, clustering model analysis helps to distinguish data samples to subgroups called as cluster. Samples in every cluster represents similar characteristics. Therefore, determination of cluster features can provide meaningful information for detection of apneic episodes. In this study, the samples of peaks for every 1-minute signal episode

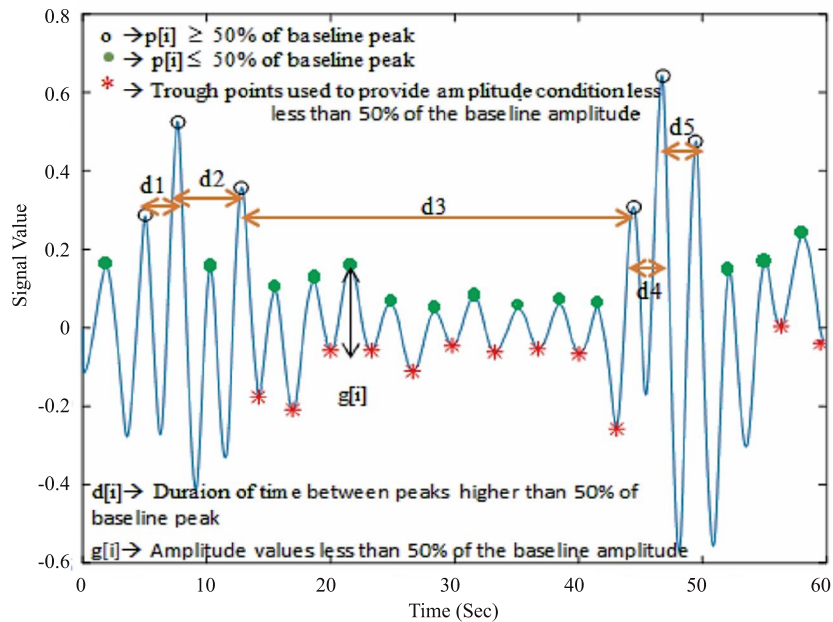


FIG. 4. CALCULATION OF FEATURES A3, A6, A10, A11

were separated into 3 clusters and absolute mean of every cluster was calculated. The minimum of these three mean values was selected as a feature (M17). First cluster contains high positive peak values. Second cluster is consisted of low positive signal values. Hence, this cluster represents the value of apneic events. Values forming the negative parts of the sinusoidal signal pattern are in the third cluster. This cluster does not important for this study. Table 3 shows the features created by descriptive models.

2.4 Feature Selection

Determination of best features that represents the data better than others increases the discrimination of apneic episodes from normals [3]. In this study, 54 features were extracted in different categories and best features must be determined to obtain the best performance. For this purpose, OneR Attribute Eval Feature Selection Algorithm which calculates the value of a features using the OneR classification algorithm was used [24]. The OneR

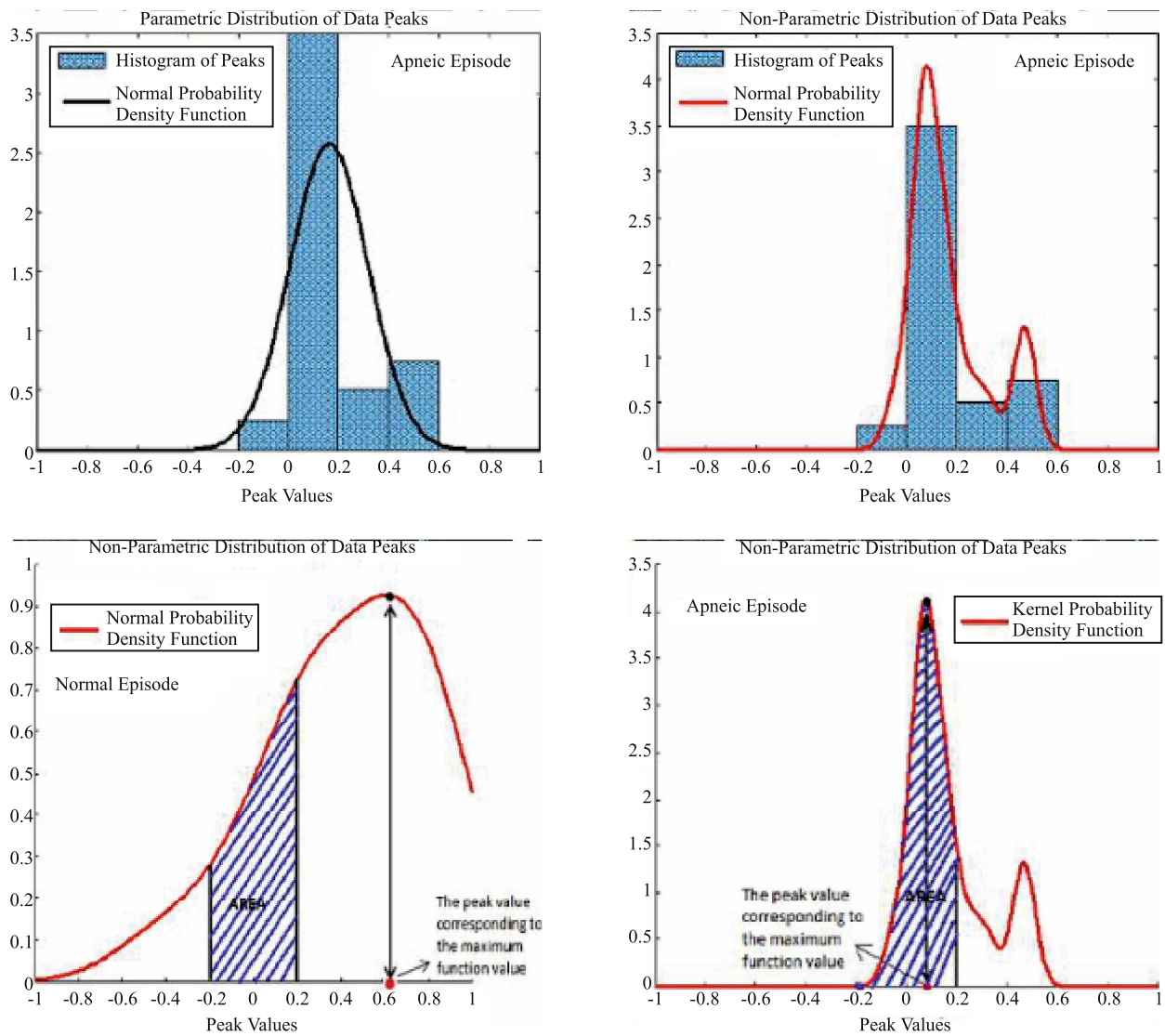


FIG. 5 . NORMAL, KERNEL, NON-PARAMETRIC AND HISTOGRAM DISTRIBUTIONS OF DATA AND PEAKS

classification algorithm ranks effectiveness of each individual feature and choose the top few to use [25]. In this study, attribute selection process was carried out with 10 fold cross-validation and average rank were obtained for every features. Firstly top 5 features were evaluated by a classifier. Then the properties were added sequentially according to the effectiveness values respectively, and the classification process was performed each time. Finally, best feature sets were determined so that it gives maximum classification result.

2.5 Classification

The classification process was carried out to detect apneic episodes of entire airflow signals using RF classifier in Weka 3.9 [26]

The RF algorithm was suggested by Brierman [27]. It is an ensemble of decision tree classifiers [1]. The RF has a lot of advantages such as rapidness, robustness to noise and outliers, resistance to overfit [1,28]. Moreover, in the algorithm, there are very few parameters to be determined such as number of features (m) to be used for each node

and the number of trees (N) to be created [29]. Also, in the literature, RF has shown successful results for two-class problems [28-30]. Due to these positive properties, the RF classification algorithm has been found suitable for this study and it was preferred to classify the 1-minute airflow signal episodes.

CA (Classification Accuracy), Prec (Precision), recall, kappa statistic and area under of ROC curve measures were used to evaluate performance of classifier and study [31].

3. EXPERIMENTAL RESULTS

In this study, it was aimed to detect apneic episodes of all airflow signals with RF using two different databases. In order to achieve this aim, pre-processing, feature extraction, feature selection and classification stages were realized.

In the pre-processing stage, filtering and segmentation processes were carried out. After the pre-processing stages, 2513 episodes from Apnea-ECG database [9] and 1699 episodes from the MIT-BIH database [10] were

TABLE 3. SET-3: DESCRIPTIVE MODEL FEATURES

Features Number	Description of Features
M1 // M2 // M3 // M4	Mean // Median // Standard deviation // Variance of Normal (Gaussian) probability density function belongs to peaks value
M5	The peak value corresponding to the maximum of normal density function value
M6	The area between -0.2 and 0.2 in the normal density function
M7 // M8 // M9 // M10	Mean // Median // Standard deviation // Variance of Kernel probability density function belongs to peaks value
M11	The peak value corresponding to the maximum of kernel density function value
M12	The area between -0.2 and 0.2 in the kernel density function
M13 // M14 // M15 // M16	Mean // Median // Standard deviation // Variance of histogram of peaks
M17	Minimum of mean peak values belongs to each cluster

created in 1-minute length. In the feature extraction stage, 54 features in 3 different categories were built for each database separately. These features were grouped as classic features, amplitude features and descriptive model features. After the creation of these feature sets, normalization process were applied to the features.

In the classification stage, representative features extracted from 1-minute signal episodes were classified as two class, either apneic or not. Apneic class includes apnea and hypopnea events because these events have not been evaluated separately.

Classification process using RF was repeated 50 times for every feature set with 10 fold cross validation. The maximum of 50 classification results were accepted as the classification result of that set. Number of tree (N) is set to 100 for RF classifier. The number of features (m) to be used for each node was left at the default value. Firstly, Set-1, Set-2 and Set-3 were classified separately. Following these classifications, features of sets were merged and classification was carried out with all 54 features. Finally, OneR attribute eval algorithm was applied to 54 features of Apnea-ECG database. Thus, effective features were specified. Feature selection process was not reapplied on the airflow signals obtained from MIT-BIH database. Same effective fetures were used for this database and classification was performed again with them.

Table 4 presents the classification results of Set-1, Set-2, Set-3, all features and selected 28 features for Apnea-ECG database, separately. It was seen that from the results, classification result of Set-1 are slightly better than both of Set-2 and Set-3. The lowest classification results were obtained with descriptive model features (Set-3). It was also observed that obtained

classification accuracy, precision, recall, kappa and ROC values for all 54 features of Apnea-ECG database are better than the results of each three individual classifications. As shown in Table 4, classification of selected 28 effective fetures produces the best classification accuracy as 96.21%. Moreover, precision, recall and kappa values with selected features are the best among all classification results.

Confusion matrices are also shown in Table 4. It can be seen from confusion matrices in Table 4, that total number of episodes with apneic and normal events are 1608 and 905, respectively. In the classification with selected 28 features, only 46 normal and 49 apneic episodes were classified wrongly. 2418 episodes were correctly classified. The number of misclassifications with the previous 4 feature sets is higher. Consequently, Table 4 illustrated that the best performance was obtained with the selected features, when all results in Table 4 were compared in terms of all evaluation criteria.

Classification results of Set-1, Set-2, Set-3, all features, selected 28 features and selected 31 features for MIT-BIH database with RF as shown in Table 5.

Unlike Apnea-ECG database, results of Set-2 features are slightly better than Set-1. Likewise, the lowest results were obtained with Set-3. When the features of three feature sets were merged for this database, the performance was better than individual results of feature sets, as in the other database. In addition, the same selected features used for Apnea-ECG database were also experimented for this database, initially. Although higher results were obtained than the individual classifications (Set-1, Set-2 and Set-3), lower results were acquired from classification with all features for MIT-BIH database. The reason of

classification result can be explained that, definition of apneic events have been made according to least

50% drop in airflow signal for Apnea-ECG database. The selected features A3, A4, A6, A7, A10, A13

TABLE 4. CLASSIFICATION RESULTS OF FEATURES SETS FOR APNEA-ECG DATABASE

Database : Apnea-ECG							
	CA	Confusion Matrix		Other Measurements			
		Normal	Apneic	Precision	Recall	Kappa	ROC
Set-1 Classic Features (22 Features)	95.62	848	57	94.1	93.7	90.49	98.3
		53	1555	96.5	96.7		
Set-2 Amplitude Features (15 Features)	95.54	849	56	93.8	93.8	90.33	98.1
		56	1552	96.5	96.5		
Set-3 Descriptive Model Features (17 Features)	95.22	841	64	93.8	92.9	89.62	98.5
		56	1552	96.0	96.5		
All Features Set-1,Set-2 and Set- 3 (54 Features)	95.98	854	51	94.5	94.4	91.28	99.1
		50	1558	96.8	96.9		
S1,S5,S6,S7,S8,S9,S10,S11 S13,S15,S16,S20,S21,S22 A1,A3,A4,A6,A7,A10,A13,A15 M5,M10,M11,M12,M13,M17 (28 Features)	96.21	859	46	94.6	94.9	91.80	99.1
		49	1559	97.1	97.0		

TABLE 5. CLASSIFICATION RESULTS OF FEATURES SETS FOR MIT-BIH DATABASE

Database : MIT-BIH							
	CA	Confusion Matrix		Other Measurements			
		Normal	Apneic	Precision	Recall	Kappa	ROC
Set-1 Classic Features (22 Features)	91.40	455	78	87.0	85.4	79.94	95.4
		68	1098	93.4	94.2		
Set-2 Amplitude Features (15 Features)	91.46	463	70	86.1	86.9	80.23	95.4
		75	1091	94.0	93.6		
Set-3 Descriptive Model Features (17 Features)	90.16	450	83	84.3	84.4	77.18	95.5
		84	1082	92.9	92.8		
All Features Set-1,Set-2 and Set- 3 (54 Features)	92.23	465	68	87.9	87.2	81.92	96.3
		64	1102	94.2	94.5		
S1,S5,S6,S7,S8,S9,S10,S11 S13,S15,S16,S20,S21,S22 A1,A3,A4,A6,A7,A10,A13,A15 M5,M10,M11,M12,M13,M17 (28 Features)	92.05	468	65	87.0	87.8	81.59	92.2
		70	1096	94.4	94.0		
Selected Features S1,S5,S6,S7,S8,S9,S10,S11 S13,S15,S16,S20,S21,S22 A1,A2,A3,A4,A5,A6,A7 A9,A10,A13,A15 M5,M10,M11,M12,M13,M17 (31 Features)	92.23	467	66	87.6	87.6	81.96	96.4
		66	1100	94.3	94.3		

support this expression. However, apneic event definition was not made for MIT-BIH database. According to AASM [12], 30% fall in the airflow signal can also describe hypopnea which is one of the apneic events. Therefore, it was added A2, A5 and A9 features in to the selected features and the number of selected features were reached 31 for MIT-BIH database. As a result of classification with 31 features, high classification accuracy was achieved as 92.23%. This accuracy value is the same as the classification accuracy of all 54 features. However, feature set with 31 features is always preferred because the number of features is less and the results are better in terms of Kappa and ROC.

As a result, the high classification accuracy for both databases were obtained with 96.21% and 92.23% values. Kappa values are quite good (91.80% and 81.96%) and support the CA values for both databases, too. If the features sets were compared, it can be said that there are no major differences between results of the classical feature set (Set-1) and the amplitude feature set (Set-2). However, when only descriptive model feature set was used, the achieved of RF was decreased. The study shows that the successful results were obtained when the features which reveal different characteristics of data were combined and effective features were selected from this combined dataset.

28 features including 14 classical features, 8 amplitude features and 6 descriptive model features were determined as effective features for Apnea-ECG database. For the MIT-BIH database, 31 features have been defined as effective features. These features contain 14 classical features, 11 amplitude features and 6 descriptive features.

Tables 6-7 show the average rank for every selected feature according to 10 fold cross-validation using OneR Attribute Eval Feature Selection Algorithm. Ranks of other features have not been shown because they are lower than those of selected features. According to Table 7, most effective features are amplitude features because 8 of the top 10 features are amplitude features for MIT-BIH database. Likewise, 6 of the top 10 features are amplitude features for Apnea-ECG database. When the top 20 features are considered for MIT-BIH database and Apnea-ECG database 11 and 9 of these features are amplitude features respectively, 6 of the features are descriptive model features for two database and 3 and 5 features are classical features for MIT and Apnea-ECG database respectively. So, it can be said that most effective features are amplitude features. Second-order effective features are descriptive model features and last effective features are classical features.

Most of selected classic features are statistical and they represents the changes that take place in the 1-minute signal episode. The features selected from amplitude feature set are associated with the AASM criteria. Selected 6 descriptive model features comprised of 1 feature related to normal density function, 3 features related to kernel density function and 2 histogram model features. The effective features used for both databases are the same except for the 3 added features (A2, A5 and A9) the MIT-BIH database. These used effective features make it possible to clearly identify the apneic events in both databases.

Literature studies, the results obtained from these studies and the comparison with this study are shown in Table 8.

It can be seen from Table 8, when this study has been compared with the studies [1,17,32] in literature, it generally outperformed the others. Only CA value of the study is less than CA value of the study performed by Avcý and

Akbas. Avcý and Akbas [1] used 8 airflow signals from Apnea-ECG database. However, only 5 airflow signal were used in this study. Three signals of Apnea-ECG database do not been apneic event. These signals is not important for our study. Therefore, 5 airflow signal were preferred by us. We think that the difference between the CAs is due to the number of used signals.

TABLE 6. APNEA-ECG DATABASE

No.	Average Rank	Features
1	88.51	S8
2.	87.52	A10
3	87.43	A3
4.	87.10	A6
5.	86.35	A1
6.	82.35	M11
7.	80.09	A4
8.	77.96	A7
9.	77.58	S22
10.	77.56	M13
11.	76.61	A13
12.	76.43	M5
13	75.24	A15
14.	75.09	M17
15.	73.56	S10
16.	72.33	M12
17.	68.58	M10
18.	68.13	S9
19.	67.78	A1
20.	67.71	S20
21.	67.08	S21
22.	65.47	S7
23.	65.23	S11
24.	65.13	S5
25.	64.73	S13
26.	64.26	S1
27.	64.24	S6
28.	63.26	S16

TABLE 7. MIT-BIH DATABASE

No.	Average Rank	Features
1	88.51	S8
2.	87.52	A10
3	87.43	A3
4.	87.10	A6
5.	86.35	A1
6.	85.71	A9 (Only MIT)
7.	84.90	A2 (Only MIT)
8.	82.89	A5 (Only MIT)
9.	82.35	M11
10.	80.09	A4
11.	77.96	A7
12.	77.58	S22
13	77.56	M13
14.	76.61	A13
15.	77.43	M5
16.	75.24	A15
17.	75.09	M17
18.	73.56	S10
19.	72.33	M12
20.	68.58	M10
21.	68.13	S9
22.	67.78	A1
23.	67.61	S20
24.	67.08	S21
25.	65.47	S7
26.	65.23	S11
27.	65.13	S5
28.	64.73	S13
29.	64.26	S1
30.	64.24	S6
31.	63.26	S16

TABLE 8. COMPARISON OF THIS STUDY AND LITERATURE STUDIES

Study Source	Database	CA	Recall/Sensitivity
Gil et. al.[32]	Apnea-ECG	-	95.3% (Apneic Events)
Avc? and Akbas [1]	Apnea-ECG	98.68%	-
This Study	Apnea-ECG	96.21%	97.0 % (Apneic Events)
Varady et. al. [17]	MIT-BIH	-	97.0% (Apnea) 78.7% (Hypopnea) Average sensitivity for apneic events: 87,85%
This Study	MIT-BIH	92.23%	94.3% (Apneic Events include apnea and hypopnea)

4. CONCLUSION

This study was aimed at the detection of the apneic events in 1 minute episodes obtained from airflow signals recorded during a night. In order to achieve this aim, 2 benchmark databases namely Apnea-ECG and MIT-BIH polysomnographic were used. During the study, different feature sets were created belonging to 1-minute signal episodes and it was investigated whether the signal episodes contained an apneic event or not using these feature sets with different characteristics. Also, effectiveness of features was evaluated on determination of apneic events. It is illustrated in the study that the combination of features and effective features representing the events significantly increases the performance of the work. When the study is evaluated according to results, it is promising for determining apneic events on a minute-by-minute basis.

ACKNOWLEDGEMENT

This study has been supported by Scientific Research Project, Selcuk University (Project Number: 2016- ÖYP-053).

REFERENCES

- [1] Avcı, C., and Akbaş, A., “Sleep Apnea Classification Based on Respiration Signals by Using Ensemble Methods”, *Journal of Bio-Medical Materials and Engineering*, No. 26, pp. 1703-1710, 2015.
- [2] Xie, B., Qiu, W., Minn, H., Tamil, L., and Nourani, M., “An Improved Approach for Real-Time Detection of Sleep Apnea”, *Proceedings of International Conference on Bio-inspired Systems and Signal Processing*, Rome, Italy, January 26-29, 2011.
- [3] Koley, B.L., and Dey, D., “Automatic Detection of Sleep Apnea and Hypopnea Events from Single Channel Measurement of Respiration Signal Employing Ensemble Binary SVM Classifiers”, *Journal of Measurement*, No. 46, pp. 2082-2092, 2013.
- [4] Diaz, J.A., Arancibia, J.M., Bassi, A., and Vivaldi, E.A., “Envelope Analysis of Airflow Signal to Improve Polysomnographic Assessment or Sleep Disordered Breathing”, *Journal of Sleep*, Volume 1, No. 37, pp. 199-208, 2014.
- [5] Vazquez, J.C., Tsai, W.H., Flemons, W.W., Masuda, A., Brant, R., Hajduk, E., Whitelaw, W.A., and Remmers, J.E., “Automated Analysis of Digital Oximetry in the Diagnosis of Obstructive Sleep Apnea”, *Journal of Thorax*, Volume 55, No. 4, pp. 302-307, 2000.
- [6] Khandoker, A.H., Gubbi, J., and Palaniswami, M., “Automated Scoring of Abstructive Sleep Apnea and Hypopnea Events Using Short-Term Electrocardiogram Recordings”, *Journal of IEEE Transactions on Information Technology in Biomedicine*, Volume 13, No. 6, pp. 1057-1067, 2009.
- [7] Xie, B., and Minn, H., “Real-Time Sleep Apnea Detection by Classifier Combination”, *Journal of IEEE Transactions on Information Technology in Biomedicine*, Volume 16, No. 3, pp. 469-477, 2012.

- [8] Bonillo, V.M., Estévez, D.A., Leal, A.F., and Pereira E.H., "Intelligent Approach for Analysis of Respiratory Signals and Oxygen Saturation in the Sleep Apnea/Hypopnea Syndrome", *Journal of The Open Medical Informatics*, Volume 8, pp. 1-19, 2014.
- [9] Penzel, T., Moody, G.B., Mark, R.G., Goldberger, A.L., and Peter, J.H., "The Apnea-ECG Database", *Journal of Computers in Cardiology*, Volume 27, pp. 255-258, 2000.
- [10] Ichimaru, Y., and Moody, G.B., "Development of the Polysomnographic Database on CD-ROM", *Journal of Psychiatry and Clinical Neurosciences*, Volume 53, No. 2, pp. 175-177, 1999.
- [11] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P., Ch., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C-K., and Stanley, H.E., "Physio Bank, Physio Toolkit, and Physio Net: Components of a New Research Resource for Complex Physiologic Signals", *Journal of Circulation*, Volume 101, No. 23, pp. 215-220, 2000.
- [12] Iber, C., Aneoli-Israel, S., Chesson, A.L., and Quan, S.F., "The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications", *American Academy of Sleep Medicine*, Westchester, 2007.
- [13] Huang, W., Guo, B., Shen Y., and Tang, X., "A Novel Method to Precisely Detect Apnea and Hypopnea Events by Airflow and Oximetry Signal", *Journal of Computers in Biology and Medicine*, No. 88, pp. 32-40, 2017.
- [14] Selvaraj, N., and Narasimhan, R., "Detection of Sleep Apnea on a Per-Second Basis Using Respiratory Signals", *Proceedings of IEEE Conference on Engineering in Medicine and Biology Society*, Osaka, Japan, July 3-7, 2013.
- [15] Várady, P., Micsik, T., Benedek, S., and Benyó, Z., "A Novel Method for the Detection of Apnea and Hypopnea Events in Respiration Signals", *Journal of IEEE Transactions on Biomedical Engineering*, Volume 49, No. 9, pp. 936-942, 2002.
- [16] Radha, M., Garcia-Molina, G., Poel, M., and Tononi, G., "Comparison of Feature and Classification Algorithms for Online Automatic Sleep Staging Based on a Single EEG Signal", *36th Annual International Conference on Engineering in Medicine and Biology Society*, Chicago, USA, August 26-30, 2014.
- [17] Yucelbas, S., and Yucelbas, C., "Pre-Estimation of Sleep Apnea Using Sleep ECG Signals", *International Conference on Theoretical and Applied Computer Science and Engineering*, Ankara, Turkey, November 10-11, 2017.
- [18] Zhanyu, M., "Non-Gaussian Statistical Models and Their Applications", Ph.D. Thesis, Stockholm, 2011.
- [19] Everitt, B.S., and Skrondal, A., "The Cambridge Dictionary of Statistics", Cambridge University Press, 2010.
- [20] Patel, J.K., and Read, C.B., "Handbook of the Normal Distribution, Series Statistics, Textbooks and Monographs", Marcel Dekker, 1996.
- [21] Nazarpour, K., Al-Timemy, A.H., Bugmann, G., and Jackson, A., "A Note on the Probability Distribution Function of the Surface Electromyogram Signal", *Journal of Brain Research Bulletin*, No. 90, pp. 88-91, 2013.
- [22] Parzen, E., "On Estimation of a Probability Density Function and Mode", *Journal of The Annals of Mathematical Statistics*, Volume 33, No. 3, pp. 1065-1076, 1962.
- [23] Rasmussen, C.E., and Williams, C.K.I., "Gaussian Processes for Machine Learning", The MIT Press, 2006.
- [24] Sik, M.S., "Data Mining and its Use in Cancer Early Diagnosis", Master Thesis, 2014.
- [25] Tripoliti, E., and Manis, G., "Feature Selection in HRV Analysis of Young and Elderly Subjects", *5th European International Conference on Federation for Medical and Biological Engineering*, Budapest, Hungary, September 14-18, 2011.
- [26] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I., "The Weka Data Mining Software: An Update", *ACM SIGKDD Explorations News Letter*, Volume 11, No. 1, pp.10-18, 2009.

- [27] Breiman, L., "Random Forests", *Journal of Machine Learning*, Volume 45, No. 1, pp. 5-32, 2001.
- [28] Akkoc, B., Arslan, A., and Kok, H., "Automatic Gender Determination from 3D Digital Maxillary Tooth Plaster Models Based on the Random Forest Algorithm and Discrete Cosine Transform", *Journal of Computer Methods and Programs in Biomedicine*, Volume 143, pp. 59-65, 2017.
- [29] Akkoc, B., Arslan, A., and Kok, H., "Gray Level Co-Occurrence and Random Forest Algorithm-Based Gender Determination with Maxillary Tooth Plaster Images", *Journal of Computers in Biology and Medicine*, Volume 73, pp. 102-107, 2016.
- [30] Koprinska, I., Poon, J., Clark, J., and Chan, J., "Learning to Classify e-Mail", *Journal of Information Science*, Volume 177, pp. 2167-2187, 2007.
- [31] Yucelbas, S., Yucelbas, C., Tezel, G., Ozsen, S., Kuccukturk, S., and Yosunkaya, S., "Pre-Determination of OSA Degree Using Morphological Features of the ECG Signal", *Journal of Expert Systems with Applications*, Volume 81, pp. 79-87, 2017.
- [32] Gil, E., Vergarab, J.M., and Laguna, P., "Detection of Decreases in the Amplitude Fluctuation of Pulse Photoplethysmography Signal as Indication of Obstructive Sleep Apnea Syndrome in Children", *Journal of Biomedical Signal Processing and Control*, Volume 3, No. 3, pp. 267-277, 2008.