

# Ensemble Heterogeneous Feature Selection (EHFS) and Heterogeneous Ensemble Classifier with VOTE (HECV) for Automatic Detection of Snore Sounds

E. Jeslin Renjith, A. Christy

**Abstract---** Obstructive Sleep Apnea (OSA) is generally considered as a sleep co-related breathing difficulty with some important well known disabling indication. This research work tends to differentiate the condition of OSA victims. The significant work is to propose an innovative Heterogeneous Ensemble Classifier with Velum Oropharyngeal Tongue Epiglottis (HECV) method with a high level of effectiveness. In the formulated study, data recording, features extraction, multi-feature selection, classification and performance evaluation are the five stages of processing. At the initial stage, the noises represented in the audio signals which were eliminated through Adaptive Fuzzy Median Filter (AFMF) algorithm. After that Crest Factor, original Frequency, Spectral Frequency Features, Subband Energy Ratio, Mel-Scale Frequency Cepstral Coefficients (MFCC), Empirical Mode Decomposition (EMD) Features, and Wavelet Energy Features are collected from the noise suppressed audio signals and inputs are fed into Ensemble Heterogeneous Feature Selection (EHFS) technique. EHFS algorithm fuses the outputs of filter and wrapper oriented feature selection methods. These identified features are classified using HECV technique which renders a good classification outputs by validation without the regard of subjects. The outputs show that the formulated HECV approach gives better performance in snore detection when co-related with other classifiers.

**Index Terms---** Obstructive Sleep Apnea (OSA), Heterogeneous Ensemble Classifier with Velum Oropharyngeal Tongue Epiglottis (HECV), Velum Oropharyngeal Tongue Epiglottis (VOTE), Multi-feature selection algorithm, and Ensemble Heterogeneous Feature Selection (EHFS).

## I. INTRODUCTION

Obstructive Sleep Apnea (OSA) is generally a sleep co-related breathing difficulty with highly significant disabling indication. OSA has partial or full damage of the upper airway at the sleep inducing noisy breathing and heavy snoring. Snoring is considered as a general indication of OSA, influencing the OSA patients higher than 80% [1]. Occupational and motor vehicle accidents, cardiovascular disease, and stroke are at a greater risk of OSA victims [2]. Several medical researchers stated that till 82% of men and women of 93% with balanced to heavy OSA stay undiagnosed [3].

OSA is considered to be a continuous total or partial damage of the upper airway at the time of sleep [4]. Meanwhile, the research on acoustic analysis of snoring sounds targeted on the connection between maximum level of OSA and various types of snoring sound features like

intensity [5-6] 9,10, power spectral measures [7-8] 11,12, bi-spectral and non-linear measures [9] 13, formant frequencies [10] 14 and temporal features [11] 15. After all, a very small priority was given on the effects of variations in the upper airway during the time of sleep on the features of snoring sounds.

In the diagnosis of OSA polysomnography (PSG) is considered to be the best standard, which involves a multi-channel recording of parameters that are physiological at the time of natural sleep [12]. In many cases, the recording of PSG is done in the sleep laboratory. The portable devices are used in the screening process of Cardio-respiratory that comes along with less recording channels that are used as a matter of choice or as added measures. In the recent years, there are many methods and applications have been surveyed and created for the purpose of monitoring the sleep. [13]. The accuracy of diagnostics of the listed methods have still not been checked and proved in the area of clinical studies.

PSG and cardio respiratory screening provide a reliable diagnosis as towards the type and severity of the sleep related breathing disorder. However, it is of very limited use to identify its underlying mechanisms. A diagnostic procedure that has been established for the evaluation of obstruction and vibration locations and mechanisms in the UA is Drug Induced Sleep Endoscopy (DISE).

DISE is increasingly used by sleep surgeons and appreciated as a useful tool to identify the location of vibration and obstruction. However, it has a number of disadvantages: DISE is cost intensive, as it requires the attendance of personnel and appropriate equipment for safe administration and monitoring of sedation, as well as endoscopic equipment. Further, it is time consuming, a DISE investigation typically requires more than 20 min overall. Also, it cannot be performed during natural sleep, as the introduction of the endoscope would cause the patient to wake up. Another method to identify the location of obstruction in the upper airway is multi-channel pressure measurement [14]. Here, a thin tube with multiple pressure sensors is introduced into the upper airway. An advantage of this method is that, it can be used in natural sleep. However, the tube within the upper airway is not accepted by every patient.

Acoustic analysis could be a method to find out vibration mechanisms within the upper airway, which is easier for doctors and patients.

Manuscript received September 16, 2019.

E. Jeslin Renjith, Research Scholar, Bharathiar University, Coimbatore. (e-mail: Jeslin.renjith@gmail.com)

Dr.A. Christy, Research Supervisor, Bharathiar University, Coimbatore. (e-mail: ac.christy@gmail.com)

# ENSEMBLE HETEROGENEOUS FEATURE SELECTION (EHFS) AND HETEROGENEOUS ENSEMBLE CLASSIFIER WITH VOTE (HECV) FOR AUTOMATIC DETECTION OF SNORE SOUNDS

Minimum numbers of studies tend to pertain in the area of determining the location and vibration forms and obstruction in the upper airway from the acoustic properties of snore sounds. The studies specified above are concentrated on calculating certain well-selected acoustic features for their sensitivity to the anatomical mechanisms of snoring sound generation or upper airway obstruction. The comparisons and the outcome are based on statistical analysis, and basic signal observation. The application of multi-feature analysis methods to this problem has been proposed in this work; however, advanced signal processing methods, and machine learning models have not yet been used for this purpose. This work intends to distinguish the severity of OSA patients. The major aim of this work is to introduce a new Heterogeneous Ensemble Classifier with VOTE (HECV) approach with higher efficiency. In this study, the proposed system has five processing stages namely data recording, features extraction, feature selection, classification and performance evaluation.

## II. LITEATURE REVIEW

Khandoker et al [15] brings out the utilization of Support Vector Machines (SVMs) for automatic recognition of victims with OSAS categories (+/-) with the help of features collected from night-time ECG recordings, and co-relates its efficiency with other classifiers. In order to recognize OSAS+/- subjects, the features are gathered from wavelet decomposition of heart rate variability (HRV) and ECG-derived respiration (EDR) signals of the entire records are given as inputs to train the SVM classifier. The optimal SVM parameter set is then evaluated by using a leave-one-out procedure. In comparison, classification performance of K-nearest neighbour, probabilistic neural network, and linear discriminant classifiers on test data was lower. These results, therefore, indicate major potential in handling SVM in ECG-based screening and can aid sleep specialists in the starting evaluation of patients with suspected OSAS.

Liu et al [16] used a 3D model of the human head to identify the source of snoring sounds; however, the model was computationally intensive and highly dependent on having an accurate 3D model of the head for every individual. Ng et al [17] developed an electrical equivalent model of the upper airway to investigate the effects of upper airway narrowing on spectral features of snoring sounds.

Beeton et al [18] investigated characteristics of the probability density function by testing the sensitivity of the statistical moments to amplitude patterns in the snoring acoustic signal. Snoring sounds from 15 patients were recorded whilst performing sleep nasendoscopy evaluation. Also demonstrated, using a fuzzy 2-means clustering method, that a combination of the statistical dimensionless moment coefficients of skewness and kurtosis could discriminate between pure palatal and non-palatal snoring subject groups.

Kim et al [19] aimed to identify acoustic biomarkers indicative of the severity of sleep disordered breathing (SDB) by analyzing the breathing sounds collected from a large number of subjects during entire overnight sleep. Next the audio features were collected and a set of features that vary between different SDB severity groups was identified as a potential acoustic biomarker. The classification tasks

were performed using several machine learning techniques in order to validate the acoustic biomarker. This study signify that any device with a microphone, such as a smartphone could be potentially utilized as a screening tool for detecting SDB.

Jin et al [20] study was to systematically determine the accuracy of acoustic analysis of snoring in the diagnosis of OSA using a meta-analysis. A standardized review and meta-analysis of sensitivity, specificity, and other measures of accuracy of acoustic analysis of snoring in the diagnosis of OSA were organized. The median of apneahypopnea index threshold was 10 events/h, range: 5–15 or 10–15 if above-mentioned suggestion is accepted. Acoustic analysis of snoring is a relatively correct but not a capable method for diagnosing OSA. There is an immediate need for laborious studies involving large samples and single snore event tests with an efficacy criterion that reflects the appropriate features of snoring acoustics for OSA diagnosis.

Alshaer et al [21] study was to determine the accuracy of BresoDx in distinguishing patients with predominantly central sleep apnea (CSA) from those with predominantly OSA. Acoustic features of the BresoDx recording for respiratory events were extracted including turbulence, frequency content, and snoring distribution. Acoustic features and demographics were used to train an artificial intelligence algorithm using PSG scoring as the reference. Data from all subjects except one were used for training the algorithm and the left-out subject used for validation. This process was repeated iteratively so that each subject was used for validation once. A subject was classified as having CSA if > 50% of events were central and OSA otherwise.

Benavides et al [22] use of speech analysis is proposed as an alternative or complement to existing screening methods. A set of voice features that could be related to apnoea are defined, based on previous results from other authors and own analysis. These features are examined first in isolation and then in combination to calculate their discriminative power to categorize voices as corresponding to apnea patients and healthy subjects. As a result of the analysis, a Linear Discriminant Analysis (LDA) was defined including a subset of eight features. This model was tested on a separate database containing 20 healthy and 20 apnea subjects yielding a sensitivity of 85% and a specificity of 75%, with a F1-measure of 81%.

Mesquita et al [23] propose a new methodology for identifying two distinct types of snores: the so-called non-regular and regular snores. The analysis is performed at a particular time interval in between regular snore sounds that occur in short segments throughout the night. The features that are obtained after the analysis of time interval between regular snores attained an accuracy rate of 88.2% on classification. Provides promising results and puts forward a valuable aid for the early screening of subjects suspected of having Sleep Apnea Hypopnea Syndrome (SAHS).

Fiz et al [24] used a novel automatic snoring detection and analysis system to screen snoring during full-night polysomnography to measure whether the acoustic features of snores vary in relation to the apnea-hypopnea index

(AHI) and to categorize subjects according to their AHI. Automatic snore intensity and frequency checking and analysis could be a capable tool for screening OSA patients, sensitively improving the managing of this pathology.

### III. PROPOSED WORK

The major aim of this work is to introduce a new Heterogeneous Ensemble Classifier with VOTE (HECV) approach with higher efficiency. The initial stage of the

work the collected audio signals recorded using Drug-Induced Sleep Endoscopy(DISE) and these signals were digitized by the frequency rate of 44.1 kHz, Pulse Code Modulation (PCM), and down-sampled to 16 kHz per each sample, which is the lowest sampling rate of the audio recorder. Secondly the noises presented in the audio signals were removed by Adaptive Fuzzy Median Filter (AFMF) algorithm.

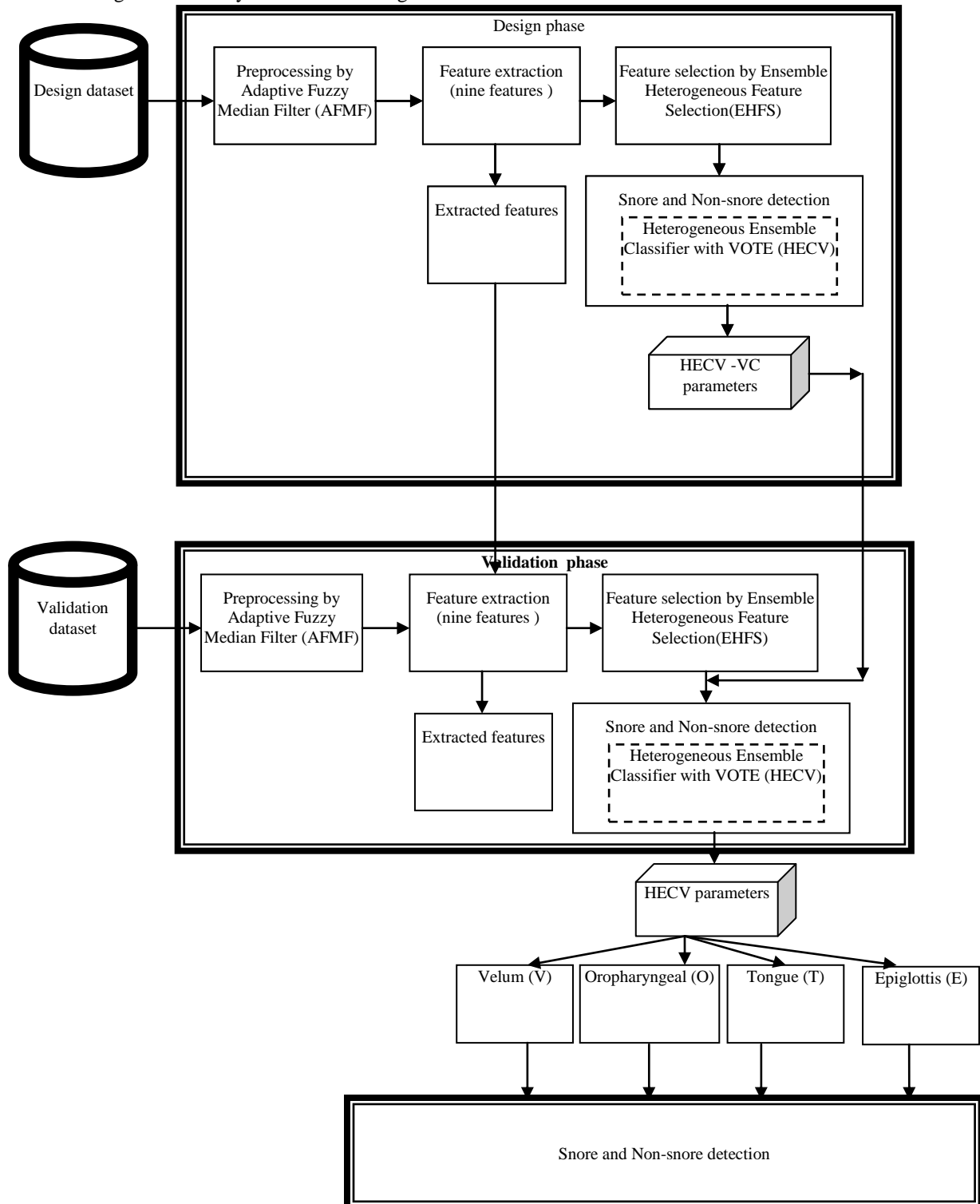


Figure 1: Architecture diagram

## ENSEMBLE HETEROGENEOUS FEATURE SELECTION (EHFS) AND HETEROGENEOUS ENSEMBLE CLASSIFIER WITH VOTE (HECV) FOR AUTOMATIC DETECTION OF SNORE SOUNDS

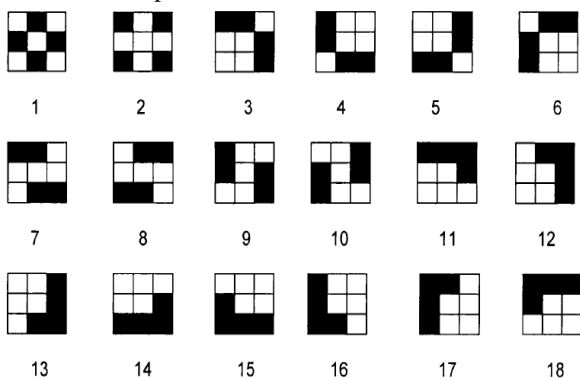
Some of the features like Crest Factor, original Frequency, Spectral Frequency Features, Subband Energy Ratio, Mel-Scale Frequency Cepstral Coefficients (MFCC), Empirical Mode Decomposition (EMD) Features, and Wavelet Energy Features have been extracted from the noise suppressed audio signals and input into Ensemble Heterogeneous Feature Selection (EHFS) approach. These selected features are classified using HECV approach which shows better classification results by subject independent validation. The proposed HECV approach showed the better classification results snore sounds of 40 male patients has been recorded by using DISE, and classified by using ENT experts depending on the VOTE classification [25]. The overall architecture of the proposed EHFS and HECV approach is illustrated in Figure 1.

### Pre-processing

For design and validations phases, the collected audio signals recorded in the DISE were digitized by the frequency rate of 44.1 kHz, Pulse Code Modulation (PCM), and down-sampled to 16 kHz per each sample, which is the lowest sampling rate of the audio recorder. These signals were noise removed based on the adaptive noise suppression which follows the procedure of AFMF. Reduction of noise in audio signals is one of the most basic audio signals. In recent years, fuzzy logic-based filters have shown to provide efficient noise filtering. The fuzzy filters can be

1. Fuzzification of weighted mean filters,
2. Fuzzy combinations of the outputs of several classical sub filters or
3. "Pure" fuzzy filters, in which no classical filter operators are used, but the filter output is directly determined by a set of rules.

A new fuzzy-based filter is designed to down-sampled to 16 kHz per each sample and is able to operate in real-time at snore sounds higher than normal sounds. The filter is based on the evaluation of fuzzy similarities between frequency ranges within a local processing window. Eighteen 4-sound templates are represented in the following figure 2. The outputs of a set of sub filters are then combined to provide the final filter output.



**Figure 2: Different 4-snore sound Templates**

Consider a small neighborhood of a center sounds  $X_o$ , in noise-free sounds. If assume that the sound details are no smaller than the processing window, the frequency ranges will be more or less constant in edge-free regions. This indicates that the intensities of the sounds inside this window are similar, except when the window contains an edge. As noise is added to the sounds, the sounds change

their original intensities and become less similar to each other. High similarities between the sounds in the window imply the presence of short-tailed noise, whereas low similarities indicate the existence of long-tailed noise. Different sub-filters are then activated to certain degrees, depending upon the similarities found in the window. When the similarities are low, a filter for long-tailed noise is activated. Here, use a simple median filter, since it has been widely used for the suppressing of long-tailed and frequencies while preventing blurring of edges. Hence, the basic idea in the Fuzzy-Based (FB) filter approach is to investigate whether the pixels in the local processing window are similar to each other. In order to check the similarity in different parts of the window, the frequencies are arranged into a set of 4- frequencies templates. The similarity between two frequencies are allowed to take values between 0 and 1, hence it is called fuzzy similarity. In this work, it is evaluated as a 1-D triangular membership function taking the intensity difference as input variable in the following equation (1)

$$\mu_{x_i=x_j} = \begin{cases} 1 - \frac{|x_i-x_j|}{\alpha} & |x_i-x_j|/\alpha \\ 0 & \text{else} \end{cases} \quad (1)$$

Generally, the fuzzy similarity can be modeled as a 2-D fuzzy relation surface. This allows for a membership function that varies with the two input parameters. In order to get the degree of similarity between the central frequency and a template, the similarities between the central frequency and each frequency in the template are calculated. Then, they are combined using an aggregation operator.

### Ensemble Heterogeneous Feature Selection (EHFS)

In ensemble learning, is performed by a collection of several classifiers [26] or algorithms is trained, and the final results of the ensemble is determined by combining the outputs of the single classifiers or algorithms, e.g. by means of using majority voting in the case of classification. There are two major steps in generating EHFS. The initial step includes generating a set of many feature selection algorithms, every giving their output, whereas the second step combines the results of the single feature selection algorithms. Combining the results of many FS algorithms has been performed by using weighted voting, e.g. in the case of obtaining a consensus feature ranking, or by including the mainly commonly chosen snore features in the case of obtaining a final snore features subset. The following feature selection algorithms are used in this work for snore features.

Filter algorithms be able to be measured as a pre-processing stage since they are independent on Heterogeneous Ensemble Classifier with VOTE (HECV).

Correlation based Feature Selection (CFS) is a simple filter algorithm that ranks snore sound feature subsets according to a correlation based heuristic accuracy function. The bias of the evaluation function is toward subsets that contain snore sound features that are highly correlated with the class and uncorrelated with each other.



Irrelevant snore sound features should be ignored because they will have low correlation with the class. Redundant features should be screened out as they will be highly correlated with one or more of the remaining snore sound features. The acceptance of a snore sound feature will depend on the extent to which it predicts classes in areas of the instance space not already predicted by other snore sound features. CFS's snore sound feature subset evaluation function is

$$M_{fs} = \frac{kr_{cf}}{\sqrt{k+k(k-1)r_{ff}}} \quad (2)$$

Where  $M_{fs}$  is the heuristic "merit" of a snore sound feature subset 'FS' containing k snore sound features,  $r_{cf}$  is the mean snore sound feature -class correlation ( $fs \in FS$ ) and  $r_{ff}$  is the average snore sound feature- snore sound feature inter-correlation. The numerator of this equation can be thought of as providing an indication of how predictive of the class a set of snore sound features is equal to each other; and the denominator of how much redundancy there is among the snore sound features.

#### Fast Correlated based Filter (FCBF)

The Fast Correlated-Based Filter (FCBF) method is based on Symmetrical Uncertainty (SU), which is defined as the ratio between the Information Gain (IG) and the Entropy (H) of two snore sound features, x and y:

$$SU(x, y) = 2 \frac{IG(\frac{x}{y})}{H(x)+H(y)} \quad (3)$$

where the information gain is defined as

$$IG(\frac{x}{y}) = H(y) + H(x) - H(x, y) \quad (4)$$

Being  $H(x)$  and  $H(x, y)$  the entropy and joint entropy, respectively. This method was designed for high-dimensionality data and has been shown to be effective in removing both irrelevant and redundant features.

Wrapper feature selection algorithm treats the HECV approach as a black box making it common. The selection of the snore sound features has the main improvement of considering the HECV approach; on the other hand this HECV approach provides improved detection results than filter approach. A frequently used wrapper approach is the so named Greedy Search strategy which increasingly creates the snore sound features set by considering or removing single snore sound features from an initial snore sound features subset. Greedy search has been divided into two types: Sequential Forward Selection (SFS) and Sequential Backward Selection (SBS).

Sequential Forward Selection(SFS) is the simplest greedy search algorithm. Starting from the empty set, sequentially add the snore sound feature  $fs_+$  that results in the highest objective function  $J(Y_k + FS)$  when combined with the snore sound features  $Y_k$  that have already been selected. SFS performs best when the optimal snore sound feature subset has a small number of snore sound features. More number of states can be potentially valuated, When the search is near the empty set. Towards the full set, the region examined by SFS is narrower since most of the snore sound features have already been selected. The search space is drawn like an ellipse to highlight the fact that there is less number of states towards the full or empty sets.

#### Algorithm

- Start with the empty set  $Y_0 = \{\emptyset\}$
- Select the next best snore sound feature  $FS_+ = \text{argmax}[J(Y_k + FS)], x \notin Y_k$
- Update  $Y_{k+1} = Y_k + X_+; k = k + 1$
- Goto 2

Sequential Backward Selection (SBS) starting from the full snore sound feature subset, sequentially remove the snore sound feature  $fs_-$  that results in the smallest decrease in the value of the objective function  $J(Y-fs_-)$ . Notice that removal of a feature may actually lead to an increase in the objective function  $J(Y_k-fs_-) > J(Y_k)$ . Such functions are said to be non-monotonic. SBS works best when the optimal snore sound feature subset has a large number of features, since SBS spends most of its time visiting large snore sound subsets.

#### Algorithm

1. Start with the full set  $Y_0=FS$
2. Remove the worst feature  $FS_- = \text{argmax}[J(Y_k - FS)]; fs Y_k$
3. Update  $Y_{k+1} = Y_k - FS_-; k=k+1$
4. Goto 2

EFS algorithm fusion the results of all FS methods in combination.

#### HECV approach

HECV approach is help to classify the snore sounds in the upper airway. VOTE classifier which classifies the recordings into four major classes: the level of the Velum (V), the Oropharyngeal area consists of the palatine tonsils (O), the Tongue base (T), and the Epiglottis (E). During samples collection stage, Snoring sounds (SnS) with many vibration location or unknown base of vibration were removed from original records. From each integrated recording, three to five SnS which established no obstructive character, have been automatically chosen. From the 40 subjects, 11, 11, 8, and 10 subjects were classified into four major classes. Among one and five snoring events per class were extracted per each person. On the whole for implementation work we have 164 snoring events (41 episodes for each sensor category of SnS, length ranging from 0.728 to 2.495s with an average of 1.498 s). They segmented the episodes into distinct parts for additional feature extraction, EFS and HECV approach. Each segment has duration of 200 ms and neighbouring segments have been extending beyond of 50 %.

Ensemble methods [27-28] are learning algorithms that build a set of classifiers and then classify new data points by taking a weighted vote of their predictions. The first method for constructing ensembles manipulates the training examples to generate multiple hypotheses. The Learning algorithm is run more number of times, each time with a different subset of the training examples. The above stated mechanism particularly well for unsteady learning algorithms whose output classifier undergoes key changes with response to little changes in the training snore sound features.



# ENSEMBLE HETEROGENEOUS FEATURE SELECTION (EHFS) AND HETEROGENEOUS ENSEMBLE CLASSIFIER WITH VOTE (HECV) FOR AUTOMATIC DETECTION OF SNORE SOUNDS

Adaptive Neuro Fuzzy Inference System (ANFIS), and Support Vector Machine (SVM) algorithms are all stable and linear threshold algorithms are generally very stable.

## Adaptive Neuro Fuzzy Inference System (ANFIS)

ANFIS utilizes the hybrid-learning rule and manage complex decision-making or diagnosis systems. ANFIS has been proven to be an effective tool for tuning the membership functions of fuzzy inference systems. ANFIS is a uncomplicated data learning procedure that uses a fuzzy inference system model to transform a given input into a target output. This forecast involves membership functions, fuzzy logic operators and if-then rules. The five significant stages of processing in the operations involved in ANFIS are input fuzzification, application of fuzzy operators, application method, output aggregation, and defuzzification [30-31].

ANFIS uses “Initial knowledge representation into a group of constraints (network topology) to decrease the optimization search space”, from Fuzzy Systems and “adaptation of back propagation to structured network to computerize FC parametric tuning”, from Neural Networks, to get improved performance. The design goal of the fuzzy controller is to study and attain good performance in the presence of disturbances and uncertainties. The design of membership functions is done by the ANFIS batch learning technique, which amounts to tune a FIS with back propagation algorithm based on a collection of input-output data pairs. Generally, ANFIS is a multilayer feed forward network in which each node performs a particular function (node function) on incoming signals. For simplicity, we consider two inputs 'x' and 'y' and one output 'z'.

Rule 1: IF x is  $A_1$  and y is  $B_1$  THEN  $f_1 = P_1x + Q_1y + R_1$

Rule 2: IF x is  $A_2$  and y is  $B_2$  THEN  $f_2 = P_2x + Q_2y + R_2$

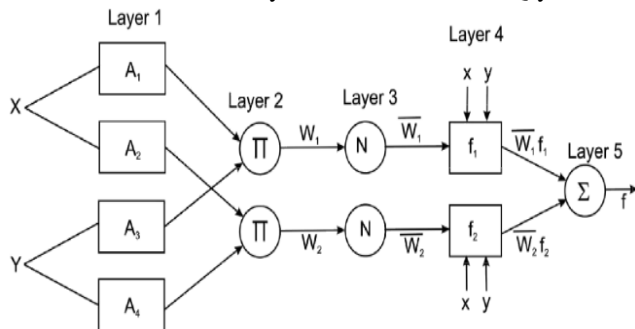


Figure 3: ANFIS Architecture

The ANFIS architecture is a five layer feed forward network as shown in Figure 3. An adaptive network is a MFFN in which every node does a specific operation on incoming signals also group of parameters focusing to this node. From node to node, the formulas for the node functions may vary and the option for every node function rely on the total input-output operation which the adaptive network is need to be done. An adaptive network only shows the direction of flow of signals between nodes. No weights are associated with the links.

To reflect different adaptive capabilities, use both circle and square nodes in an adaptive network. A square node (adaptive node) has parameters while a circle node (fixed node) has none. The group of parameters belonging to an adaptive network is the union of the

parameter group of every adaptive node. In order to achieve a desired input-output mapping, these parameters are updated according to given training data and a gradient-based learning procedure is used [32].

Layer 1: Every node in this layer is a square node with a node function (the membership value of the premise part)

$$O_i^1 = \mu_{A_i}(x) \quad (5)$$

Where, x is the input to the node i, and  $A_i$  is the linguistic label associated with this node function.

Layer 2: Every node in this layer is a circle node labelled  $\Pi$  which multiplies the incoming signals. Each node output represents the firing strength of a rule.

$$O_i^2 = \mu_{A_i}(x)\mu_{B_i}(y) \quad (6)$$

Layer 3: Every node in this layer is a circle node labeled N(normalization). The  $i^{\text{th}}$  node determines the ratio of the  $i^{\text{th}}$  rule's strength of firing to the over all firing strengths.

$$O_i^3 = \bar{W}_i = \frac{w_i}{w_1+w_2} \quad (7)$$

Layer 4: Every node in this layer is a square node with a node function

$$O_i^4 = \bar{W}_i f_i = \bar{W}_i(P_i x + Q_i y + R_i) \quad (8)$$

Layer 5: The single node in this layer is a circle node labeled  $\Sigma$  that computes the overall output as the summation of all incoming signals

$$O_i^5 = \text{results} \quad (9)$$

Equation (8) represents the overall output of the ANFIS, which is functionally equivalent to the fuzzy system[33].

## Support Vector Machine (SVM)

Let us consider a supervised binary classification problem. Let us assume that the training set consists of vectors from the  $n$ -dimensional feature space  $s_i$ . A target is associated to each vector  $y_i \in \{-1, +1\}$ . Let us assume that the two classes are linearly independent. This means that it is possible to find at least one hyperplane (linear surface) defined by a vector 'w' (normal to the hyper plane) and a 'b' bias that can separate the two classes without errors. The membership decision rule can be based on the function  $\text{sgn}[f(\mathbf{f}s)]$ , where  $f(\mathbf{f}s)$  is the discriminant function associated with the hyper plane and defined as

$$f(\mathbf{f}s) = \mathbf{w} \cdot \mathbf{f}s + \mathbf{b} \quad (10)$$

The SVM approach consists in finding the optimal hyper plane that maximizes the distance between the closest training sample and the separating hyper plane. It is possible to express this distance as equal to  $1/||w||$  with a simple rescaling of the hyper plane parameters w and b such that

$$y_i(\mathbf{w} \cdot \mathbf{f}s_i + \mathbf{b}) > 0 \quad (11)$$

The geometrical margin between the two classes is given by the quantity. The idea of margin is central in the SVM method, since it is a measure of its generalization ability. The larger the margin, the higher the expected generalization [33].

$$\min_{i=1, \dots, N} y_i(\mathbf{w} \cdot \mathbf{f}s_i + \mathbf{b}) \geq 1 \quad (12)$$

This classical linearly constrained optimization problem can be interpreted (using a Lagrangian formulation) into the following dual problem.



$$\begin{cases} \max_{i=1,\dots,N} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (f_{s_i}, f_{s_j}) \\ \text{subject to } \sum_{i=1}^N \alpha_i y_i = 0 \text{ and } \alpha_i \geq 0 \end{cases} \quad (13)$$

The Lagrange multipliers'  $\alpha_i$  expressed in (13) can be estimated using quadratic programming (QP) methods [33]. The discriminant function interrelated with the optimal hyperplane develops calculation depending both on the Lagrange multipliers and on the training samples, i.e.,

$$f(\mathbf{f}_s) = \sum_{i \in \mathbf{f}_s} \alpha_i y_i (\mathbf{f}_{s_i}, \mathbf{f}_s) + \mathbf{b} \quad (14)$$

Where 'FS' is the subset of training samples corresponding to the nonzero Lagrange multipliers  $\alpha_i$ . It is worth observing that the Lagrange multipliers efficiently weight each training sample according to its importance in determining the discriminant function. The training samples associated to nonzero weights are called support vectors. But it needs more running time complete the task, to solve this issue new variable is introduced that automatically select the weight values depending on the difference between the snore and non-snore sound features. To represent the weight values, difference between the two features values if found in the classifier, if the current feature is more different to the snore sound features then the weight is assigned to lesser than 0.5 if it is closer to snore sound detection than weight is assigned to greater than 0.8. These lie at a distance accurately equal to  $1/||w||$  from the optimal separating hyperplane.

$$\alpha_i = \text{diff}(\mathbf{f}_{s_i}, \mathbf{f}_s) = \begin{cases} \text{if } \text{diff}(\mathbf{f}_{s_i}, \mathbf{sn}) \text{ is lesser then } \alpha_i > 0.8 \\ \text{if } \text{diff}(\mathbf{f}_{s_i}, \mathbf{sn}) \text{ is higher then } \alpha_i < 0.5 \end{cases} \quad (15)$$

#### Majority voting

In each iteration 't' the learning algorithm is invoked to minimize the weighted error on the training set and it returns a hypothesis  $h_t$ . The weighted error of  $h_t$  is computed and applied to update the weights on the training examples. The effect of the change in weights is to place more weight on training examples that were misclassified by  $h_t$  and less weight on examples that were correctly classified. In subsequent iterations therefore AdaBoost constructs progressively more difficult learning problems. The final classifier,

$$h_f(\mathbf{f}_s) = \sum_{t=1}^T w_t h_t(\mathbf{f}_s) \quad (15)$$

is constructed by a weighted vote of the each classifier. Each classifier is weighted by  $w_t$  according to its accuracy on the weighted training snore sound feature subset that it was trained on.

## IV. RESULTS AND DISCUSSION

In this work, we analytically evaluate normally used multi-acoustic features for their detection accuracy on the classifiers of snore sounds depending on Logistic Regression (LR), k Nearest Neighbour (kNN), Ensemble Swarm Intelligent based VOTE Classification (ESIVC), Ensemble Convolutional Neural Network with VOTE Classification (ECNN-VC) approach and proposed HFCV approach. All experiments are implemented with the help of MATLAB R2012 software environment.

During samples collection stage, Snoring sounds (SnS) with many vibration location or unknown base of vibration were removed from original records. These SnS records have been then extracted from the audio signals and classified based on the ECNN- VC approach. From the 40

subjects, 11, 11, 8, and 10 subjects were classified into four major classes. On the whole for implementation work we have 164 snoring events (41 episodes for each sensor category of SnS, length ranging from 0.728 to 2.495s with an average of 1.498 s). They segmented the episodes into distinct parts for additional feature extraction, EFS and ECNN-VC approach. In order to evaluate the snore detection rate of all classifiers, apply the Unweighted Average Recall (UAR), described as follows:

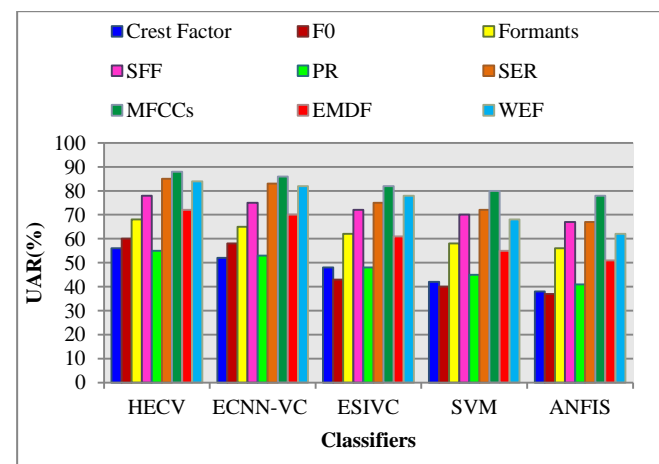
$$UAR = \frac{\sum_{class=1}^{N_{MC}} N_{class,correct} / N_{class,all}}{N_{MC}} \times 100\% \quad (16)$$

#### Without feature selection

The UAR results of three classifiers with different feature sets are shown in Table I.

**Table 1: UAR ([%]) Results Obtained With Nine Features and Five Classifiers without Feature Selection**

Features	UAR ([%])				
	HECV	ECNN-VC	ESIVC	SVM	ANFIS
<b>Crest Factor</b>	56	52	48	42	38
<b>F0</b>	60	58	43	40	37
<b>Formants</b>	68	65	62	58	56
<b>SFF</b>	78	75	72	70	67
<b>PR</b>	55	53	48	45	41
<b>SER</b>	85	83	75	72	67
<b>MFCCs</b>	88	86	82	80	78
<b>EMDF</b>	72	70	61	55	51
<b>WEF</b>	84	82	78	68	62
<b>ALL</b>	93	91	87	77	68
<b>Average</b>	71.77%	69.33%	63.22%	58.88%	55.22%



**Figure 4: UAR results vs. five classifiers without feature selection**

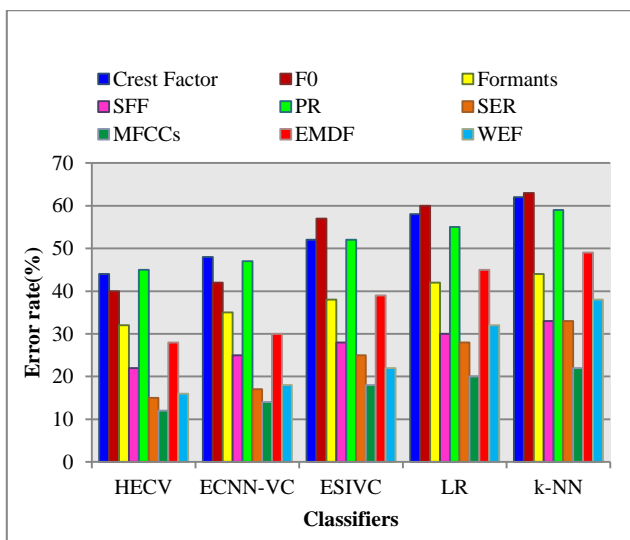
From the results it concludes that the proposed HECV approach with the purpose of MFCCs achieves the best snore sound detection results of 88%.

## ENSEMBLE HETEROGENEOUS FEATURE SELECTION (EHFS) AND HETEROGENEOUS ENSEMBLE CLASSIFIER WITH VOTE (HECV) FOR AUTOMATIC DETECTION OF SNORE SOUNDS

Among them nine of the different subsets, the novel SER performs provides next higher UAR of 85% when compared to other classifiers namely ANFIS classifier and Support Vector Machine (SVM) classifier. On the other hand, Crest Factor, F0, and PR800 don't give better detection results for all classifiers are illustrated in figure 2. From the results it concludes that the proposed HECV approach produces average accuracy results of 71.77% which is 2.44%, 8.55%, 12.88% and 16.55% higher when compared to ECNN-VC, ESIVC, SVM, and ANFIS classifiers respectively.

**Table 2: Error rate results obtained with nine features and four classifiers without feature selection**

Features	Error rate ([%])				
	HECV	ECNN-VC	ESIVC	SVM	ANFIS
<b>Crest Factor</b>	44	48	52	58	62
<b>F0</b>	40	42	57	60	63
<b>Formants</b>	32	35	38	42	44
<b>SFF</b>	22	25	28	30	33
<b>PR</b>	45	47	52	55	59
<b>SER</b>	15	17	25	28	33
<b>MFCCs</b>	12	14	18	20	22
<b>EMDF</b>	28	30	39	45	49
<b>WEF</b>	16	18	22	32	38
<b>ALL</b>	7	9	13	23	32
<b>Average</b>	28.22 %	30.66 %	36.77 %	41.11 %	44.77 %



**Figure 4: Error results vs. four classifiers without feature selection**

The error results of the five different classifiers is illustrated in the figure 4, it concludes that the proposed HECV approach produces lesser error results of 7% which is 2%, 6%, 16% and 25 % lesser when compared to other classifiers with features sets. It concludes that the proposed work performs better than the other classifiers shown in table II.

### After feature selection results

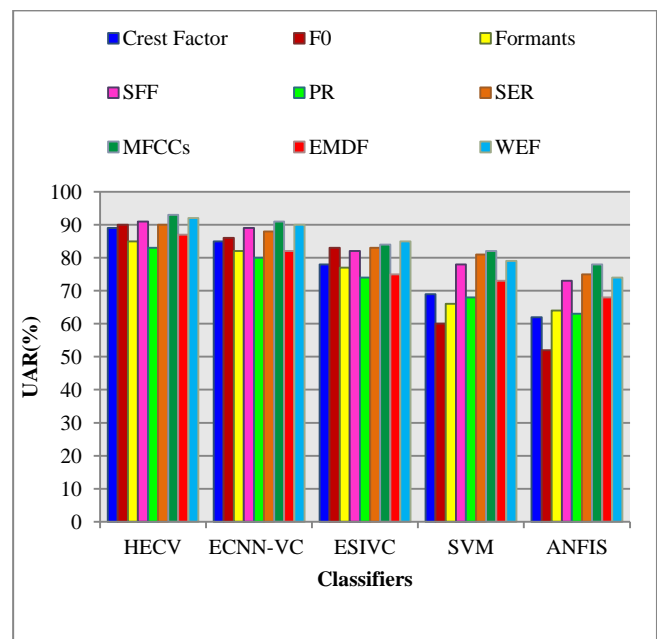
In this study, we use HEFS algorithm for the multi-dimensional feature selection. Particularly, for all feature sets, the average performance of the proposed HECV

approach considerably increases from 83% to 95%. For F0 (86), PR (80), and Crest Factor (85), improved after the completion of the HEFS discussed in table 3.

In the experiments results the proposed HECV approach achieve the UAR of 95% with the best combination of features.

**Table 3: UAR ([%]) results obtained with nine features and five classifiers after EHFS**

Features	UAR ([%])				
	HECV	ECNN-VC	ESIVC	SVM	ANFIS
<b>Crest Factor</b>	89	85	78	69	62
<b>F0</b>	90	86	83	60	52
<b>Formants</b>	85	82	77	66	64
<b>SFF</b>	91	89	82	78	73
<b>PR</b>	83	80	74	68	63
<b>SER</b>	90	88	83	81	75
<b>MFCCs</b>	93	91	84	82	78
<b>EMDF</b>	87	82	75	73	68
<b>WEF</b>	92	90	85	79	74
<b>ALL</b>	95	94	91	87	85
<b>Average</b>	88.88 %	85.88 %	80.11 %	72.88 %	67.66 %



**Figure 5: UAR results vs. four classifiers with EHFS**

Particularly, for all feature sets, the average performance of the proposed HECV approach produces 88.88% by HEFS is illustrated in the figure 5.

The average accuracy results of the proposed HECV approach is 88.88% which is 3%, 8.77%, 16% and 21.22% higher when compared to ECNN-VC, ESIVC, LR and k-NN classifiers respectively.



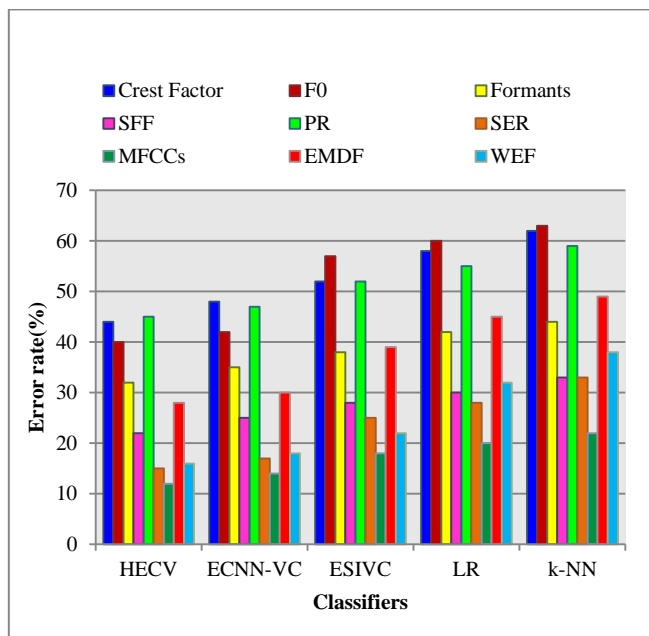


Figure 6: Error rate results vs. four classifiers with EHFS

The error results of the five different classifiers with EHFS is illustrated in the figure 6, it concludes that the proposed HECV approach produces lesser error results of 5% which is 1%, 4%, 8% and 10% lesser when compared to other HECV, ESIVC, SVM, and ANFIS classifiers respectively with all selected features sets. It concludes that the proposed HECV approach performs better than the other classifiers are discussed in table 4.

Table 4: Error rate ([%]) results obtained with nine features and five classifiers after EFS

Features	Error rate ([%])				
	HECV	ECNN-VC	ESIVC	SVM	ANFIS
Crest Factor	11	15	22	31	38
F0	10	14	17	40	48
Formants	15	18	23	34	36
SFF	9	11	18	22	27
PR	17	20	26	32	37
SER	10	12	17	19	25
MFCCs	7	9	16	18	22
EMDF	13	18	25	27	32
WEF	8	10	15	21	26
ALL	5	6	9	13	15
Average	11.11%	14.11%	19.88%	27.11%	32.33%

## V. CONCLUSION AND FUTURE WORK

Examining deviations in the upper airway reduction for the period of sleep is invasive and expensive. Because the snoring sounds are created by air instability and vibrations of the upper airway appropriate to its reduction. In the work we introduce an Ensemble Heterogeneous Feature Selection (EHFS) by combining results of filter and wrapper based feature selection methods. Filtering method follows the procedure of Correlation based Feature Selection (CFS), and Fast Correlated Based Filter (FCBF). A frequently used

wrapper approach is the so named Greedy Search strategy which has been divided into two types: Sequential Forward Selection (SFS) and Sequential Backward Selection (SBS). This EHFS introduces a subject-precise acoustic model of the upper airway in the direction of examine the impact of upper airway by snoring sounds of multi-features. Then classification is performed by Heterogeneous Ensemble Classifier with Velum Oropharyngeal Tongue Epiglottis (HECV). HECV algorithm fusion the results of all classifiers such as Support Vector Machine (SVM) and Adaptive Neuro Fuzzy Inference System (ANFIS) which showed the best classification results by subject independent validation. HECV are learning algorithms that construct a set of classifiers and then classify new data points by taking a weighted vote of their predictions. The first method for constructing ensembles manipulates the training examples to generate multiple hypotheses The Learning algorithm is run several times each time with a different subset of the training examples. The proposed methods specifically works well for the learning algorithms which are unstable, whose output classifier meets significant modifications with respect to small changes in the training snore sound features. These results encourage the use of snoring sounds examination to evaluate the upper airway analysis during OSA. Though the acoustics of snoring, as a diagnostic system is from a growing stage, there is an urgent need of new system to an accurate, huge data set. Detection is performed to unbalanced features of snores is also left as scope of future work to diagnose OSA.

## REFERENCES

1. F. R. de Almeida, N. T. Ayas, R. Otsuka, H. Ueda, P. Hamilton, F. C. Ryan, and A. A. Lowe, "Nasal Pressure Recordings to Detect Obstructive Sleep Apnea," *Sleep Breath*, vol. 10, pp. 62-69, 2003.
2. G. C. Chang and Y. P. Cheng, "Investigation of Noise Effect on Lung Sound Recognition," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, Kunming, China, vol. 3, pp. 1298-1301, 2008.
3. T. Young, L. Finn, et al., "Estimation of the Clinically Diagnosed Proportion of Sleep Apnea Syndrome in Middle Aged Men and Women," *Sleep*, vol. 20, pp. 705-706, 1997.
4. Dempsey, J. A., Veasey, S. C., Morgan, B. J. & O'Donnell, C. P. Pathophysiology of Sleep Apnea. *Physiol. Rev.* 90, 47-112 (2010).
5. Maimon, N. & Hanly, P. J. Does snoring intensity correlate with the severity of obstructive sleep apnea? *J. Clin. Sleep Med.* 6, 475-478 (2010).
6. Nakano, H., Furukawa, T. & Nishima, S. Relationship Between Snoring Sound Intensity and Sleepiness in Patients with Obstructive Sleep Apnea. *J. Clin. Sleep Med.* 4, 551-556 (2008).
7. Fiz, J. A. et al. Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea. *Eur. Respir. J.* 9, 2365-2370 (1996).
8. Xu, H., Huang, W., Yu, L. & Chen, L. Sound spectral analysis of snoring sound and site of obstruction in obstructive sleep apnea syndrome. *Acta Otolaryngol.* 130, 1175-1179 (2010).

## ENSEMBLE HETEROGENEOUS FEATURE SELECTION (EHFS) AND HETEROGENEOUS ENSEMBLE CLASSIFIER WITH VOTE (HECV) FOR AUTOMATIC DETECTION OF SNORE SOUNDS

10. Ng, A. K., San Koh, T., Abeyratne, U. R. & Puvanendran, K. Investigation of obstructive sleep apnea using nonlinear mode interactions in nonstationary snore signals. *Ann. Biomed. Eng.* 37, 1796–1806 (2009).
11. Ng, A. K. et al. Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea? *Sleep Med.* 9, 894–898 (2008).
12. Alencar, A. M. et al. Dynamics of snoring sounds and its connection with obstructive sleep apnea. *Phys. A: Statistical Mechanics and its Applications* 392, 271–277 (2013).
13. T. Verse, A. Dreher, C. Heiser, M. Herzog, J. Maurer, W. Pirsig, K. Rohde, N. Rothmeier, A. Sauter, A. Steffen, S. Wenzel, B. Stuck, Ent-specific therapy of obstructive sleep apnoea in adults: a revised version of the previously published German s2e guideline, *Sleep Breath.* 20 (4) (2016) 1301–1311.
14. W. Huang, B. Guo, Y. Shen, X. Tang, A novel method to precisely detect apnea and hypopnea events by airflow and oximetry signals, *Comput. Biol. Med.* 88 (2017) 32–40.
15. B. A. Stuck and J. T. Maurer, “Airway evaluation in obstructive sleep apnea,” *Sleep Medicine Reviews*, vol. 12, no. 6, pp. 411–436, 2008.
16. A.H. Khandoker, C.K. Karmakar, M. Palaniswami, Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings, *Comput. Biol. Med.* 39 (1) (2009) 88–96.
17. Liu, Z., Luo, X., Lee, H. & Lu, C. Snoring source identification and snoring noise prediction. *J. Biomech.* 40, 861–870 (2007).
18. Ng, A. K., Koh, T. S., Baey, E. & Puvanendran, K. Role of upper airway dimensions in snore production: acoustical and perceptual findings. *Ann. Biomed. Eng.* 37, 1807–1817 (2009).
19. R. J. Beeton, I. Wells, P. Ebdon, H. Whittet, and J. Clarke, “Snore site discrimination using statistical moments of free field snoring sounds recorded during sleep nasendoscopy,” *Physiological Measurement*, vol. 28, no. 10, pp. 1225–1236, 2007.
20. Kim, T., Kim, J.W. and Lee, K., 2018. Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *Biomedical engineering online*, 17(1), p.16.
21. Jin, H., Lee, L.A., Song, L., Li, Y., Peng, J., Zhong, N., Li, H.Y. and Zhang, X., 2015. Acoustic analysis of snoring in the diagnosis of obstructive sleep apnea syndrome: a call for more rigorous studies. *Journal of Clinical Sleep Medicine*, 11(07), pp.765-771.
22. Alshaer, H., Hummel, R. and Bradley, T.D., 2017. Distinguishing Patients with Central from Obstructive Sleep Apnea Using Overnight Breath Sound Recordings, *European Respiratory Journal*,50: OA3204.
23. Benavides, A.M., Pozo, R.F., Toledano, D.T., Murillo, J.L.B., Gonzalo, E.L. and Gómez, L.H., 2014. Analysis of voice features related to obstructive sleep apnoea and their application in diagnosis support. *Computer Speech & Language*, 28(2), pp.434-452.
24. Mesquita J, Solà-Soler J, Fiz JA, Morera J, Jané R. All night analysis of time interval between snores in subjects with sleep apnea hypopnea syndrome. *Med Biol Eng Comput* 2012;50:373–81.
25. Fiz JA, Jané R, Solà-Soler J, Abad J, García MA, Morera J. Continuous analysis and monitoring of snores and their relationship to the apneahypopnea index. *Laryngoscope* 2010;120:854–62.
26. J. Kezirian, W. Hohenhorst, and N. de Vries, “Drug-induced sleep endoscopy: the vote classification,” *European Archives of Oto-Rhino- Laryngology*, vol. 268, no. 8, pp. 1233–1236, 2011.
27. Saeys, Y., Abeel, T. and Van de Peer, Y., 2008, Robust feature selection using ensemble feature selection techniques. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 313-325.
28. Rokach, L., 2010. Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1-2), pp.1-39.
29. Kodovský, J., Fridrich, J.J. and Holub, V., 2012. Ensemble classifiers for steganalysis of digital media. *IEEE Trans. Information Forensics and Security*, 7(2), pp.432-444.
30. Polat, K. and Güneş, S., 2007. An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digital Signal Processing*, 17(4), pp.702-710.
31. Güler, I. and Übeyli, E.D., 2005. Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. *Journal of neuroscience methods*, 148(2), pp.113-121.
32. Çaydaş, U., Hasçalık, A. and Ekici, S., 2009. An adaptive neuro-fuzzy inference system (ANFIS) model for wire-EDM. *Expert Systems with Applications*, 36(3), pp.6135-6139.
33. Mohandes, M., Rehman, S. and Rahman, S.M., 2011. Estimation of wind speed profile using adaptive neuro-fuzzy inference system (ANFIS). *Applied Energy*, 88(11), pp.4024-4032.
34. Pal, M. and Mather, P.M., 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5), pp.1007-1011.

