ISSN 2278-3091
Volume 9, No.1, January – February 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse59912020.pdf
https://doi.org/10.30534/ijatcse/2020/59912020

A Brief on Snoring Data and Classification Methods

Tosin A. Adesuyi[1], Byeong Man Kim[2], Yoon Sik Shin[3]
[1]Department of Software Engineering, Kumoh National Institute of Technology, S. Korea, atadesuyi@kumoh.ac.kr
[2]Department of Software Engineering, Kumoh National Institute of Technology, S. Korea, bmkim@kumoh.ac.kr
[3]Department of Software Engineering, Kumoh National Institute of Technology, S. Korea, ysshin@kumoh.ac.kr

ABSTRACT

Recently, research on snoring sound had gained interest especially in the area of classification in Obstructive Sleep Apnea (OSA) and distinction from non-snoring sounds. Classifiers such as Support Vector Machine (SVM), Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) had been used to meet this interest. These approaches relies on several underlying techniques such as Mel Frequency Cepstral Coefficients (MFCC), Short-Time Fourier transform (STFT) and several others to extract features from snore/apnea Spectrogram images before classification process. However, achieving desirable classification accuracy depends on choice of classifier, feature extraction techniques and, available dataset. In the paper, we presented a brief survey on existing methods and snore data acquisition processes to quickly expose and ease new researches in this domain to make appropriate choice from available methods.

Key words: Snoring, Classification methods, Signal processing, Sleep, Survey.

1. INTRODUCTION

Sleep is a natural phenomenon in human that the body, mind and, nervous system are in a relaxed or inactive state. Occasionally, human does experience sleep disorder which may be classified as primary snoring (PS) that is has no medical co-morbidity [1], and snoring phenomenon associated with Obstructive Sleep Apnea (OSA) that implies health risk [2]. According to authors in [3-6], 44% of adult women above the age of 40 and more than 60% of adult men do snore. Snoring sound has been recorded to have approximately 90 to 100dB, and this has the potential to induce hearing loss to persons sleeping beside the snorer [7,8]. Aside this, the snorer can suffer from OSA: a severe health status that affect more than 10% of men and 6 % of women living in US [9]. Janott et. al. (2018) described OSA as a repeated episodes of reduced or completely halted airflow despite an ongoing effort to breathe. OSA symptoms include: morning headache, feels of sleepiness at daytime, and severe fatigue [10], all of which can lead to hypertension and myocardinal infarcture (also known as heart attack) [11]. Aside polysomnography which has high financial cost implication [12], other several approaches [3, 7], [13-22] has been proposed in literatures to detect, diagnose and remedy early snoring stage before getting to an unhealthy phase. Lately, methods in the field of signal processing and AI seems to be the promising area to detect snoring and apnea event. These approaches majorly include: support vector machine (SVM) [10, 16], recurrent neural network (RNN) [3, 22], convolutional neural network (CNN) [17], electromechanical film transducer (Emfit) signal [18], pulse transition time (PTT) [19], acoustic signal processing with KNN [20], deep neural network [13], and hybrid methods [12, 20, 22]. Several of these methods made use of publicly available snoring data or personally aggregated snoring dataset. Moreover, Mel Frequency Cepstral Coefficients (MFCC) techniques and several others are used to extract features from snore/apnea spectrogram images along with the aforementioned methods for classification process.

In the paper, we presented a brief survey on existing methods and snore data acquisition processes to quickly expose and ease new researches in this domain to make appropriate choice from available methods. The rest of the paper is organized as follows: Section 2 surveys popular techniques used in snoring classification phases and related works, while section 3 highlight classification result obtained from several researches, thereafter we conclude the paper in section 4.

2. POPULAR TECHNIQUES USED IN SNORING CLASSIFICATION PHASES

In this section an overview of snoring data acquisition, related works, frequently used methods and their underlying techniques are presented.

2.1 Snoring Data Acquisition

Based on our findings, we categorize snoring data sources into four namely: online available snore sound corpus, snore data provided by medical organizations, snore data created through subjects, and crowdsourced snore data.

2.1.1 Publicly Available Snore Sound Corpus

A typical example of this dataset is the Munich-Passau Snore Corpus which consist of 828 snore audio samples grouped into four classes. Each classes is based on the source of obstruction that resulted into snoring and they include: Velum (V), Oropharyngeal (O), Tongue (T), and Epiglottis (E) [17]. The dataset was made available through the INTERSPEECH 2017 ComParE Snoring Challenge [23]. Authors in [17, 21] stressed the challenge of unbalance class samples as the velum class dominated the distribution. They were able to bypass this by simply replicating other class sample to balance the classes. However, authors of [16] were able split the snore audio samples into 282 for training; 283 for development; and 263 for testing.

2.1.2 Snore data provided by medical organization

This includes data recorded from patients with apnea in a hospital or medical research centers. In [20], an overnight audio recording of Obstructive Sleep Apneas/Hypopnea Syndrome patients were provided by the department of Otolaryngology, Beijing Hospital. The audio were captured through a single channel low-noise microphone placed at 30cm above the patient's head. Moreover in [19], patient data with obstructive sleep apnea syndrome were collected from Firat University Research Hospital Sleep Room polysomnography recording. The Total sample size was 100 (including 50 female and 50 male). Janott et al. (2018), gathered a total of 2174 sample data (snore and non-snore) from three hospitals in Germany. The samples are from patients already diagnosed for OSA through polysomnography. They later down-sample the size to 219 with four class (as in Munich-Passau Snore Corpus) after further diagnosis using DISE (drug induced sleep endoscopy) examination. The dataset also suffer from unbalance class samples but were solved by adding more samples which subsequently increased the total size to 223 [10].

2.1.3 Snore data created through subjects

These are snoring audios captured from people who consented to be subjected to sleep or other form of sleep induced substance to record snoring activity. According to authors in [3], 8 subjects were recorded overnight (approximately 8hrs) for 3 days by placing their smart phones 1 meter away from their head. The audio samples recorded were classified into snore and non-snore sounds. [22] used a field recorder and non-contact microphone placed 70cm away from the bed top to record 20 subjects. The samples also were annotated into snore & non-snore events, and were divided into training (11), validation (3), and test (6) sets. Authors in [12] used Olympus Noise canceling Microphone (ME52) hung at a height within 20-30cm above 24 volunteer's head to record throughout the night. The samples consist of three events: snore, apnea, and silence. Perez-Macias et al. (2018) however used 30 subjects who has undergone polysomnography for one full night to captured 30 samples of snoring events [18].

2.1.4 Crowdsourced Snore Data

This is a process of collecting/aggregate snoring and related-snore data from different online sources. Khan (2018) did this successfully as he was able to source 1000 sound samples from different online sources. The dataset consist of 500 snoring and non-snoring sound each [13].

2.2 Frequently Used Methods and their Underlying Techniques

From existing literatures, we survey frequently used method and their underlying techniques as follows:

2.2.1 Spectrogram

A Spectrogram is used to transform a snore audio sample into image that enables feature extraction (see figure 1). A Spectrogram is formed by converting sounds into image via a time-frequency imaging techniques based on short-time fourier transform (STFT) [16]. Snoring spectrogram images are represented using color maps such as viridis (blue, green, and yellow), jet (blue, green, and red), and gray (black, gray, white) [17]. The color maps exist in the matplotlib: a library in python package.

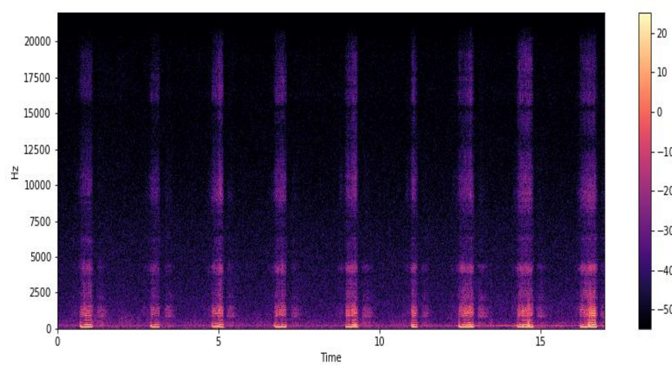

Figure 1: A sample spectrogram image of a snore signal

2.2.2 Mel Frequency Cepstral Coefficients (MFCC)

The MFCC is used as a feature extraction technique and had been found to be efficient in automatic speech recognition [24]. MFCC analyzes frame of particular length in an input signal based on Mel-frequency scale [3]. According to authors in [3] and [25-27], steps for estimating MFCC includes:

(a) Pre-emphasis: Generally, high frequency components especially in speech signals has low energy, and this hinders the extraction of useful features. Therefore, pre-emphasis is used to boost the energy from low to high. A signal that is pre-emphasized $\tilde{x}(n)$ is denoted as:

$$\tilde{x}(n) = x(n) - \alpha x(n-1) \quad (1)$$

$x(n)$ is the input signal, *n* is the sample number, and α has a value ranges from 0.9 to 1.0.

(b) Framing: Framing is used to divide $\tilde{x}(n)$ into N time of frame with adjacent frames separated by P frame shift. It is assumed that there exist a constant signal properties within each frame, however, abrupt division of signal (at both ends) by framing do lead to information or feature loss [27]. Based on time measurement N ranges from 10 to 30ms and $P \le 0.5$. Each frame $\tilde{f}_j$ is estimates as:

$$\eta = \frac{P+[\tau-N]}{p} \tag{2}$$

$$\tilde{f}_j(n) = \tilde{x}(Pj+n), \tag{3}$$

$0 \le n \le N-1$, $0 \le j \le \eta$. η is the numbers of frame in the signal and τ is the total samples in the signal.

(c) Hamming window: Windowing is used to avoid the process of information loss that may occur during framing. Furthermore, it is used to prevent truncation in frame continuity at both ends of the signal (snore sound). To perform windowing on a signal, frames f are multiplied by the hamming window ω(n) as [3]:

$$f_j = \omega(n) \times \tilde{f}_j(n), \ 0 \le n \le N-1 \tag{4}$$

$$\omega(n) = \left[-\beta \cos\left(\frac{2\pi n}{N-1}\right) - (\beta - 1)\right], \tag{5}$$

$$0 \le n \le N-1$$

Parameter β is set to 0.46 [27].

(d) **Fast Fourier Transform:** FFT uses continuous and periodic signal in a frame and convert each signal in time domain to frequency domain.

(e) Mel filter bank and DCT: The mel filter bank quantifies the level of energy at low frequency while the Discrete Cosine Transform (DCT) logs the energy from the filter bank and convert it from frequency scale to time scale [3]. A Mel frequency Mel (*f*) can be defined as [27]:

$$Mel(f) = 2595 log_{10}\left(\frac{f}{700} + 1\right) \tag{6}$$

f is the linear scale frequency.

$$\text{DCT} = \sum_{i=1}^{N} \cos\left[\frac{(i-0.5)\times 3.142c}{N}\right] \times E_i \tag{7}$$

$c = 1,2,3,\dots,m$, where E_i is the log value of i^{th} mel filter coefficient m and N are numbers of Mel-scale cepstral coefficients, and number of Mel filter respectively. A sample representation of MFCC is using snoring signal is given in figure 2.

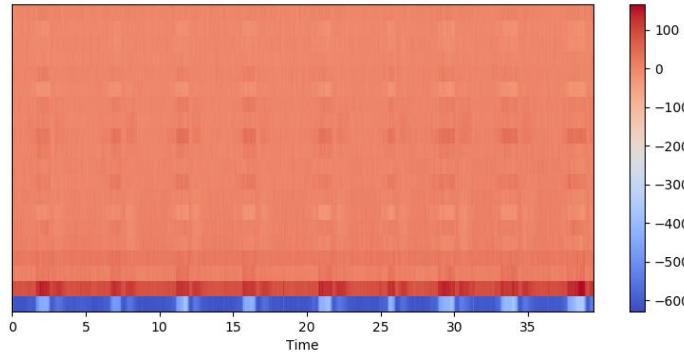

Figure 2: A sample representation of MFCC from a snore signal.

2.2.3 Short-Time Fourier transform (STFT)

The STFT is an improvement on the Fourier transform technique to create time-frequency representation of sounds/signals by windowing according to time [3]. It has the capacity to extract features in frequency and time through unvaried time length L and time function for windowing ω [3]. STFT for a discrete-time signal $X(n)$ is represented as:

$$X_\omega(nL, \mu) = \sum_{c=-\infty}^{\infty} x(c) \times \omega(nL - c)e^{-j\mu c} \tag{8}$$

2.2.4 Histogram of Oriented Gradient (HOG)

HOG is a good texture descriptor [16] that divides an image into a dense grid and compute the gradient and also the histogram of each region of interest based on the voted weight of each pixel in the image [28]. According to authors in [16] each regions is grouped into blocks and their histograms are normalized. They applied this technique to extract features from snore spectrogram image by varying the feature dimension and color channels.

2.2.5 Zero-Crossing Rate (ZCR)

The zero-crossing rate is used to analyze voiced and unvoiced signals. It is regarded as the rate at which positive signal changes to negative signal and vice versa. It does this by measuring the number of times a waveform crosses the zero amplitude [29]. ZCR divides audio signal into K frames such that $\{f_j(n): 1 \le j \le K\}$ and each frame zcr_i is computed as [30]:

$$\text{zcr}_j = \sum_{n=1}^{r-1} sgn\left[f_j(n) \times f_j(n-1)\right] \tag{9}$$

r is the number of samples in each frame and

$$sgn(f_j(n)) = \begin{cases} 1, & if\ f_j(n) > 0 \\ 0, & otherwise \end{cases} \tag{10}$$

2.2.6 Classification Methods

Commonly used classification methods include: SVM, KNN, CNN, RNN and hybrid of these methods.

(a) SVM: The SVM was originally developed to classify data with two classes. According to authors in [31], it is regarded as a technique used for classification of linear and non-linear dataset. It is also referred to as a classifier that maps binary input unto a high-dimensionality within a feature space and repeatedly finds an hyperplane that maximize the distance between the origin and training inputs [32]. For an example, let x represent data in the input space S such that it is mapped (Φ is the mapping function) into a high dimension space D as: $x \in \mathbb{R}^S \mapsto \Phi(x) \in \mathbb{R}^D$ along with a kernel function $\phi(x)$ to locate the separating hyperplane [27]. SVM testing phase is represent in Eq. 11 such that from class x, label class y can be determined by:

$$y = \begin{cases} n, & if\ \delta_n(x) + h > 0 \\ 0, & if\ \ \delta_n(x) + h \le 0 \end{cases} \tag{11}$$

$$\delta_n(x) = \max\{d_j(x)\}_{j=1}^{N} \tag{12}$$

where h is the classification threshold and $d_j(x)$ denotes the distance between x and the hyperplane corresponding to class i [32].

(b) KNN (K Nearest Neighbor): KNN searches for an n-dimensional pattern space using the training data and thereafter looks for the k training samples in close proximity to the sample to be sorted by a pre-defined distance measures, and finally classified them to k number based on nearest neighbors [33]. Sequel to basic KNN principle of finding the shortest distance between training data, Euclidean distance equation can be substituted for this process as [34]:

$$\delta(x, y) = \|x - y\|^2 = \sqrt{\sum_{j=1}^{m}(x_j - y_j)^2} \tag{13}$$

where *m* is the data dimension, δ is the distance function, *x* and *y* are sample and test data respectively.

(c) CNN: It is a type of deep neural networks that works best with image recognition. CNN networks has been used in video and image applications such as objects/image detection [35]. It is based on the convolution of images and extraction of salient features based on filters that are learned by the network during training phase [36]. Aside the input layer, the stacked layers of Convolutional neural network include: convolution layer, activation layer, pooling layer, and fully-connected layer. Mathematically we can represent computation at layer l with filter size $s \times s$ and convolution depth d such that neuron $x_{i,j,k}^{l}$ is computed with activation function $f(.)$ as:

$$a_{i,j,k}^{l} = \sum_{p=1}^{d}\sum_{q=0}^{s}\sum_{r=0}^{s} b_k^l + w_{i,j,k}^{l,k}\, x_{i+q,j+r,p}^{l-1} \tag{14}$$

$$x_{i,j,k}^{l} = \max(0, a_{i,j,k}^{l}) = f(a_{i,j,k}^{l}) \tag{15}$$

where b is the bias applied at each filter. We can also denote the max pooling function task as $mpool(.)$, hence we have:

$$z_{i,j,k}^{l} = mpool(x_{m,n,k}^{l}) \forall (m, n) \in \mathcal{R}_{ij}$$

$\mathcal{R}_{ij}$ is a local neighborhood around location (i, j) [37]. During backpropagation difference between estimated value $\hat{y}$ and target value y (label) is computed as error via the cross entropy loss function L_θ as:

$$\hat{y} = softmax(z^L)$$

$$L_\theta = -\sum_{i=1}^{|D|} y_i \log \hat{y_i} + (1 - y_i)\log(1 - \hat{y_i}) \tag{16}$$

(d) RNN (Recurrent Neural Network)*:* Basically, RNN is fashioned to learn from sequential information. The terminology "recurrent" depict that current computation is correlated with previous task [38]. Usually, at any time step of sequence, RNN compute its prior memory and the current input. The computed memory is used to predict the current time step and it is forwarded to the next step as input [34]. The memory or hidden stated vector h_t is computed as:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}\, x_t) \tag{17}$$

$$z = W_{hz}\, h_t + b_z \tag{18}$$

Where x_t is the input, W_{xh} is the weight matrix connecting the input to the memory state, W_{hz} is the weight matrix connecting the memory states to the output, b_z is the bias, and f represents activation function [36]. For multi-class classification, RNN make use of the SoftMax activation function on the output z hence, we have:

$$z_t = SoftMax(z) \tag{19}$$

2.3 A Brief Survey of Existing Related Works

Figure 3 is a generic architecture for snoring sound classification process. In existing researches, snoring/apnea sounds are converted into spectrogram images and thereafter, feature extraction techniques are applied. The extracted features are finally passed into a classifier for classification. Few works had been found to bypass the feature extraction technique and applied a classifier directly while some extract the features directly from the snore audio without using spectrogram image and subsequently forward the extracted features to a classifier. Example of such works are given in the following review (also see table 1). Authors of [19] transformed polysomnography (PSG) data of snoring and sleep apnea to pulse transition time (PTT) signal. The signal was later converted to a spectrogram image and feed as input into AlexNet and VGG16 convolutional networks. These processes were used as pre-training phase to extract features from the PPT signal. The feature vectors obtained were fed into a combination of SVM and KNN classifier. The classification accuracy was higher than 90%. [20] also used

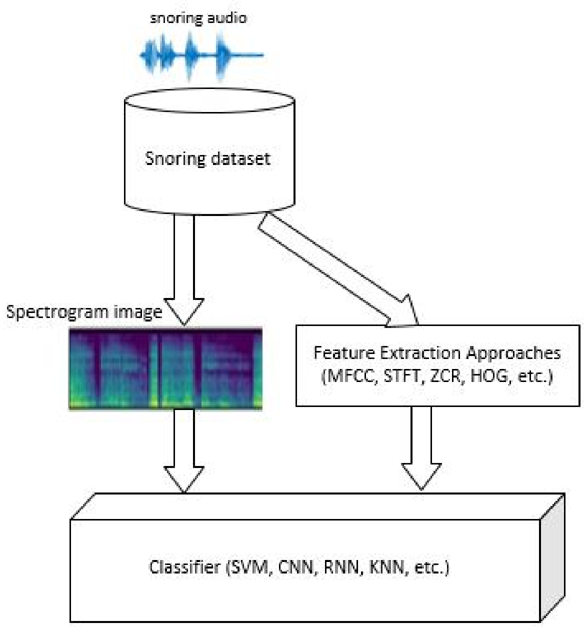

Figure 3:. A generic architecture for snore/apnea sound classification.

KNN classifier combined with MFCC to detect, segment, and classify snore related sound. The approach was very effective and yielded an excellent result. Approach in [12] applied a linear prediction coding (LPC), MFCC and sub-band for feature extraction along with hybrids of deep neural networks (DNN + LSTM; CNN + LSTM) to classify snoring, apnea and, silence event. However, the classification accuracy for both hybrids were less than 90%.
Moreover, [16] made a comparative analysis between Local Binary Pattern (LBP) and HOG for features extraction from snoring sub-challenge dataset. The extracted features were classified using SVM and the combination of both (LBP+HOG) seem to achieve a better result than each feature extraction techniques. Authors in [22] focused on RNN and MFCC technique to detect snoring and non-snoring sound and were able to achieved accuracy higher than 90%. Furthermore, [3] also used RNN classifier to classify snoring and non-snoring sound but leverage on multi-feature extraction techniques that include; Zero-crossing rate (ZCR), STFT, and MFCC. Though their dataset was developed from relative very small number of subjects (8) however, the result was close to 99%. A hybrid of dual convolutional neural network (CNN) and gated recurrent unit (GRU) was proposed in [21]. Features from their spectrogram image were fed into two parallel convolutional layers. Their average were merged together as a single channel slide before going into the GRU. The result was not as good as the aforementioned approaches. Khan (2019) applied MFCC technique to extract features from 1000 sound samples in order to detect snoring and notify vibration using a smart wearable gadget. . In the work, a CNN model was developed that produced a classification accuracy of 96% [13]. In contrast, the work in [17] leverage on existing learning models of AlexNet and VGG19 to pre-trained and extract features from a deep spectrum snore image. The output features were sent into SVM classifier. However, the pre-trained features from AlexNet had a better development and test result when used with viridis color maps. Authors of [10] likewise used SVM classifier to classify snore dataset into four classes. They applied multi-level feature extraction techniques that include: MFCC, harmonic-to-noise ratio (NHR), pitch, spectral harmonicity, voicing, and microprosodic features. An unweighted average recall of 55.8% was recorded.

3. RESULTS FROM EXISTING APPROACHES

In this section, we presented a cross section of result obtained in the classification of snore/apnea dataset from existing works based on the classifier used and the features extraction techniques. From table 1, it is evident that a multi-feature extraction techniques (ZCR+STFT+MFCC) along with RNN classifier gave the best result (98.8%) having used one of the largest data size. On the contrary, absence of feature extraction technique may seriously affect the result of a classifier and thus decline classification accuracy as cited in Ref. [21]. The result form table 1 also indicate that RNN classifier when used with MFCC as a feature extractor does yield a consistent accuracy above 90%. CNN in like manner, can achieve such result but only with a large enough snoring dataset. However, the combination of SVM and KNN may be suitable to produce high classification accuracy when the snoring/apnea data size is small. Overall, good choice of feature extractor is very germane but combination of two deep neural networks classifier may not give desirable result.

Table 1: Classification results in existing researches

Ref.	FET.	CL.	S/TDS	T. Acc %
[16]	LBP+HOG	SVM	-/828	72.0
[3]	ZCR+STFT+ MFCC	RNN	8/5600	98.8
[12]	MFCC	CNN+ LSTM	24/24	88.28
[12]	LPC	CNN+ LSTM	24/24	88.12
[22]	MFCC	RNN	20/5670	95
[13]	MFCC	CNN	-/1000	96
[21]	-	Dual CNN+ GRU	-/828	63.8
[19]	PPT signal + AlexNet+ VGG16	SVM+KNN	100/100	92.78
[17]	Deep Spectrum	AlexNet	-/828	67.0
[28]	SCAT+GMM +MAP	MLP	224/282	67.71

Ref.: authors reference, FET.: feature extraction techniques, CL.: classifier, S/TDS: subject/training data size, T. Acc: test accuracy, SCAT: Deep Scattering Spectrum; GMM: Gaussian Mixture Model; MAP: Maximum a Posteriori; Multi-Layer Perceptron.

4. CONCLUSION

This paper presented a brief on snoring data and classification methods. It further expatiate on the snoring/apnea data aggregation processing by category. Also, it reviews underlying techniques for feature extraction that foster good classification accuracy with classifiers. In summary, the paper gave a brief exposition on research activities undertaken in the quest for snoring/apnea data classification such that it enables quick and easy guide for interested researcher in the research area.

ACKNOWLEDGEMENT

This research was supported by Kumoh National Institute of Technology (2019-104-074).

REFERENCES

1. M. M. M. De Meyer, W. Jacquet, O. M. Vanderveken, L. A. M. Marks, **Systematic review of the different aspects of primary snoring**, *Sleep Medicine Reviews, Elsevier*, 45 pp. 88-94, 2019. https://doi.org/10.1016/j.smrv.2019.03.001
2. O. Parra, A. Arboix, J. Montserrat, L. Quinto, S. Bechich, and L. Garcia-Eroles, **Sleep-related breathing disorders: impact on mortality of crebrovascular disease**, *European Respiratory Journal*, vol. 24(2) pp. 267-272, 2004.
3. S. J. Lim, S. J. Jang, J. Y. Lim, J. H. Ko, **Classification of Snoring Sound Based on a Recurrent Neural Network**, *Expert Systems with Applications, Elsevier*, 123 pp. 237-245, 2019.
4. F. Dalmasso, R. Prota, **Snoring: Analysis, Measurement, Clinical Implications and Applications**, *European Respiratory Journal*, 9 pp. 841-853, 1996. https://doi.org/10.1183/09031936.96.09010146
5. W. D. Duckitt, S. K. Tuomi, T. R. Niesler, **Automatic Detection, Segmentation ans assessment of snoring from ambient acoustic data**, *Physiological Measurement*, 27 pp. 1047-1056, 2006.
6. E. Lugaresi, F. Cirignotta, G. Coccagna, C. Piana, **Some Epidemiological Data on Snoring and Cardiovasculatory disturbarnces**, *Sleep*, 3 pp. 221-224, 1980. https://doi.org/10.1093/sleep/3.3-4.221
7. X. Zhang, X. Qiu, **Performance of a Snoring Noise Control System Based on an Active Partition**, *Applied Acoustics, Elsevier*, 116 pp. 283-290, 2017.
8. M. G. Sardesai, A. K. W. Tan, M. Fitzpatrick, **Noise-Induced Hearing Loss in Snorers and their Bed Partners**, *J Otolaryngol*, 32 pp. 141-145, 2003.
9. P. E. Peppard, T. Young, J. H. Barnet, M. Palta, E. W. Hla**, Increased Prevalence of Sleep-Disorder Breathing in Adults**, *Am. J. Epidemiol*, 177(9) pp. 1006-1014, 2003. https://doi.org/10.1093/aje/kws342
10. C. Janott, M. Schmitt, Y. Zhang, K. Qian, V. Pandit, Z. Zhang, C. Heiser, W. Hohenhorst, M. Herzog, W. Hemmert, et. al., **Snoring Classified: The Munich-Passau Snore Sound Corpus**, *Computers in Biology and Medicine*, 94 pp. 106-118, 2018.
11. T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, S. Badr, **The Occurrence of Sleep: Telephone interview survey of a United Kingdom Population Sample**, *BMJ*, 314 pp. 1230-1235, 1993.
12. B. Kang, X. Dang, R. Wei, **Snoring and Apnea Detection Based on Hybrid Neural Networks**, *In Proc. International Conference on Orange Technologies, Singapore*, pp. 57-60, 2017.
13. T. Khan, **A deep Learning Model for Snoring Detection and Vibration Notification Using a Smart Wearable Gadget**, *Electronics*, 8, 987, 2019.
14. M. Lechner, C. E. Breeze, M. M. Ohayon, B. Kotecha, **Snoring and Breathing Pauses During Sleep: Interview survey of a United Kingdom Population Sample Reveals a Significant Increase in the Rates of Sleep Apnoea and Obesity Over the Last 20 years-Data from the UK Sleep Survey**, *Sleep Medicine, Elsevier*, 54 pp. 250-256, 2019.
15. B. T. Woodson, K. S. Tadokoro, S. G. Mackay, **Radiofrequency Ablation of the Lateral Palata Space for Snoring**, *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 3 pp. 106-109, 2017.
16. F. Demir, A. Sengur, N. Cummins, S. Amiriparian, B. Schuller, **Low Level Texture Features for Snore Sound Discrimination**, *In Proc. 40th Annual International Confernce of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 413-416, 2018.
17. S. Amiriparian, M. Gerczuk, S. Ottl, N. Cummins, M. Freitag, S. Pugachevskiy, A. Baird, B. Schuller, **Snore Sound Classification Using Image-Based Deep Spectrum Features**, *In Proc. Interspeech*, pp. 3512-3516, 2017. https://doi.org/10.21437/Interspeech.2017-434
18. J. M. Perez-Macias, M. Tenhunen, A. Varri, S. –L. Himanen, J. Viik, **Detection of Snores Using Source Separation on an Emfit Signal**, *IEEE Journal of Biomedical and Health Informatics*, Vol. 22(4), pp. 1157-1167, 2018.
19. S. A. Tuncer, B. Akilotu, S. Toraman, **A deep Learning-Based Decision Support System for Diagnosis of OSAS Using PTT Signals**, *Medical Hyptheses, Elsevier*, 127 pp. 15-22, 2019.
20. K. Qian, Z. Xu, H. Xu, Y. Wu, Z. Zhao, **Automatic Detection, Segmentation and Classification of Snore Related Signal from Overnight Audio Recording**, *IET Signal Processing*, Vol. 9(1), pp. 21-29, 2015.
21. J. Wang, H. Stromfelt, B. W. Schuller, **A CNN-GRU Approach to Capture Time-Frequency Pattern Independence for Snore Sound Classification**, *26th European Signal Processing Conference (EUSIPCO)*, pp. 997-1001, 2018.
22. B. Arsenali, J. V. Dijk, O. Ouweltjes, B. D. Brinker, D. Pevernagie, R. Krijn, M. V. Gilst, S. Overeem**,**

Recurrent Neural Network for Classification of Snoring and Non-Snoring Sound Events, *In Proc. 40th Annual International Confernce of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 328-331, 2018.
https://doi.org/10.1109/EMBC.2018.8512251

23. B. Schuller, S. Steidl, A. Bathlinear, E. Bergelson, J. Krajewski, C. Janott, A. Amatuni, M. Casillas, A. Seidle, M. Soderstrom, et al., **The INTERSPEECH 2017 Computational Paralinguistics Challenge: Addressee, Cold & Snoring**, *In Proc. INTERSPEECH 2017, 18TH Annual Conference of the International Speech Communication Association, ISCA. Stockholm, Sweeden: ISCA*, 2017.
24. R. L. Rabiner, W. R. Schafer, **Theory and Applications of Digital Speech Processing**, Pearson Education, Inc., 2010.
25. P. G. N. Priyadarshani, N. G. J. Dias, A. Punchihewa, **Dynamic Time Warping based speech recognition for Isolated Sinhala words**, *IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 892-895, 2012.
26. C. W. Weng, C. Y. Lin, J. S. R. Jang, **Music Instrument Identification Using MFCC: Erhu as an Example**, *In 9th International Conference of the Asia Pacific Society for Ethnomusicology*, pp. 42-43, 2004.
27. S. Jothilakshmi, V. N. Gudivada, **Large Scale Data Enabled Evolution of Spoken Language Research and Applications**, *in Handbook of Statistics Elsevier*, Vol. 35, pp. 301-340, 2016.
https://doi.org/10.1016/bs.host.2016.07.005
28. N. Dalal, B. Triggs, **Histograms of Oriented Gradients for Human Detection**, *In Proc. IEEE International Conference Computer Vision Pattern Recognition*, pp. 886-893, 2005.
29. N. Chauhan, T. Isshiki, D. Li, **Speaker Recognition Using LPC, MFCC, ZCR Features with ANN and SVM Classifier for Large Input Database**, *In Proc. 4th International Conference on Computer and Communication Systems*, pp. 130-133, 2018.
30. A. Ghosal, R. Chakraborty, R. Chakraborty, S. Haty, B. C. Dhara, S. K. Saha, **Speech/Music Classification Using Occurrence Pattern of ZCR and STE**, *In Proc. 3rd IEEE International Symposium on Intelligent Information Technology Application*, 435-438, 2009.
31. A. A. Munya, B. Sangita, **Survey of Machine Learning Techniques in Medical Imaging**, *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8(5), pp. 2107-2116, 2019.
https://doi.org/10.30534/ijatcse/2019/39852019
32. F. Vesperini, A. Galli, L. Gabrielli, E. Principi, S. Squartini, **Snore Sounds Excitation Localization by Using Scattering Transform and Deep Neural Networks**, *International Joint Conference on Neural Networks (IJCNN)*, 2018.
33. J. Huang, Y. Wei, J. Yi, M. Liu, **An Improved KNN Based on Class Contribution and Feature Weighting**, *10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 313-316, 2018.
34. M. Athoillah, M. I. Irawan, E. M. Imah, **Study Comparison of SVM-, K-NN-, and Backpropagation-Based Classifier for Image Retrieval**, *Journal of Computer Science and Information*, Vol. 8(1), pp. 11-19, 2015.
35. A. S Mohamed, N. Marbukhari, H. Habibah, **A Deep Learning Approach in Robot-Assisted Behavioral Therapy for Autistic Children**, *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8(1.6), pp. 437-443, 2019.
https://doi.org/10.30534/ijatcse/2019/6381.62019
36. S. Pattanayak, **Pro Deep Learning with Tensorflow: A Mathematical Aproach to Advanced Artificial Intelligence in Python**, *Apress Media*, pp.178, 2017.
37. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, L. Wang, G. Wang, J. Cai, T. Chen, **Recent Advances in Convolutional Neural Networks**, arXiv: 1512.07108v6 [cs.CV], 2017.
38. S. P. Sheetal, B. K. Nilesh, **Review on Text Sequence Processing with use of different Deep Neural Network Model**, *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8(5), pp. 2224-2230, 2019.
https://doi.org/10.30534/ijatcse/2019/56852019