Characterization of Audio Snores (CA-Snores)

Ali Salloum IST (197827) and Pierre Mougenot (IST 197706) and Barbara Leikam (IST 197702)

***Abstract*— Snoring is common in the general population and the irregularity could lead to the presence of Obstructive sleep apnea. Diagnosis of OSA could therefore be made by snoring sound analysis. However, there is still a shortage of robust methods to automatically detect snoring sounds without the need to calibrate for every individual. In this project we would like to explore the richer information present in audio level sampling of the snore signals. One goal is to find common and severity specific of snoring patterns. Therefore, we were looking at the time-frequency characteristics and snore envelope morphology. This project was based on several steps in order to produce a robust full-night feature extraction and we tried different classifier to characterise the signal. For identification, the main focus was set on the random forest classifier, which performed with an accuracy of 78%.For comparison, Decision Tree, Gaussian Naıve Bayesian and Quadratic Discriminant Analysis classifiers have been tried. They all showed good results as well but did not achieve the RFC's high accuracy. Running the classifiers with different input settings also showed a large difference in significance towards the classifier's performance between snoring-features.**

Snoring sound | Whole-night | Classification | Machine learning | Prognosis

I. PROBLEM AND MOTIVATION

Snoring is increasingly being recognized as of public health concern [1]. Epidemiological studies have shown that nearly 40% of males and about 20% of females are snorers. Snoring sounds are favored by various physiologic and social factors, whereas these sounds yield direct information on the dynamics and ventilation of the upper airways, for example the diagnosis of apneas. Snoring could be related to sleep deprivation and thus to other severe pathologies [2].

In fact, Obstructive sleep apnea (OSA) is an independent risk factor for cardiovascular diseases, stroke, hypertension, and myocardial infarction and in high-risk cases a targeted medical intervention is necessary. Numerous surgical methods have been proposed to cure snoring, but many of them do not solve completely the issue if the origin of the vibration is not precisely located. Although the exact mechanism of snoring sound generation is highly influenced by the individual anatomy, the typical. Prolonged partial obstruction is usually associated with sustained crescendo snoring. Using a snoring signal to assess prolonged partial obstruction is problematic since there is no means to differentiate between being snoring and snoring associated with marked partial obstruction. There for more research needs to be done [3] [4].

No European standards are available for the measurement of snoring, but three methods (acoustic sensor, nasal pressure transducer, piezoelectric vibration sensor.) are considered equivalent in the scoring manual of the AASM [5]. The standard method for detecting snoring signals is Polysomnography. It is usually done at a sleep disorders unit within a hospital or at a sleep center. This is very time intensive and is time intensive to the cost of the patients. Snoring sounds are having a high diversity and are highly variable.

Snoring sounds can change from one breath to the other in their frequency domains and in time. In consequence numerous possible classifications of snoring sounds are possible. This variability in sound patterns may arise due to :

- Altering (geometric, physical) characteristics of resonating cavities in the upper airways such as the pharynx or mouth cavity. The geometry and apertures of cavities significantly change when airways temporarily occlude or fully dilate.
- Movement of the site of collapse upstream or downstream the airway also contributes to the sound variability.

Mostly the following classifications are based on [6] :

1) The location of the sound source.
2) The diagnostically relevant type of snoring.
3) The waveform of snoring sounds in the time domain.

If an accurate classification of audio signals is possible, maybe a long term stay in a hospital is not necessary anymore and patients would be more flexible.

II. BACKGROUND AND RELATED WORK

Several works have been presented in the recent years on multi-feature acoustic analysis methods with the aim to classify and segment snore/non-snore sleep sounds.

In [7] the method is based on the characterisation of spectral energy distribution of snoring signals and a two-dimensional projection via principal component analysis (PCA). Within this paper the data from 30 subjects (18 simple snorers and 12 OSA patients) with different apnoea/hypopnea indices were classified using the proposed algorithm. The system was tested by using the manual annotations of an ENT specialist as a reference. The accuracy for simple snorers was found to be 97.3% when the system was trained using only simple snorers' data.

The paper [8] tracheal respiratory signals are recorded and snore segments are detected by extracting 10 temporal and spectral features and an Artificial Neural Network (ANN).In this case the snore detection algorithm was applied to the tracheal sounds of nine individuals with different OSA severities. On the testing dataset, the classifier achieved a sensitivity and specificity of 95.9% and 97.6% respectively.

In another paper [9] the automatic sound segmentation into snoring/breathing/noise episodes is performed using an adaptive effective-value threshold method for noise reduction, feature extraction of both linear and nonlinear descriptors and a Support Vector Machine (SVM) classifier. Therefor the proposed automatic detection method achieved over 94.0% accuracy when identifying snoring and non-snoring sounds despite using small training sets.

Specifically regarding the determination of the vibration or occlusion mechanisms, the use of different acoustic feature sets has been proposed in [2], while in [10] a k-nearest neighbour (k-NN) classifier is fed with different acoustic features. A performance comparison of different feature sets in combination with frequently used classifier model is shown in [11].

In [12], [13] the authors exploit deep convolutional neural networks pre-trained on image datasets for feature extraction from Short Time Fourier Transform (STFT) representation of the snore sounds. Then, the outputs of the bottom layers are used to feed a SVM model which provides the snore sound classification.

In another paper they used an algorithm for snoring sounds classification based on Deep Scattering Spectrum (SCAT), Gaussian Mixture Models (GMM) Supervectors and Deep Neural Networks (DNN) with an performance of the algorithm has been assessed on the Munich-Passau Snore Sound Corpus (MPSSC), composed of recordings of Drug-Induced Sleep Endoscopy (DISE) examinations. The results are expressed in terms of Unweighted Average Recall (UAR) and a remarkable improvement with respect to the state-of-the-art performance has been registered, achieving a score up to 67.14% and 74.19% respectively on the devel and test datasets.
Also, the project work of last year had on high impact on this paperwork [14].

III. Approach and Uniqueness

A. Material

Ten large stereo whole-night recordings (.wav format) of male and female patients. The set are from different degree of severity of the snoring signals. These recordings were made with fixed non-contact microphones on the left and on the right side of the patient's bed.

B. Methods

Because the data are too large to be properly opened and worked on, the algorithm works on samples (windows of 5min) one at a time until the end of the recording. This allows the code to require less RAM and processing power. These samples are converted to mono and filtered using a wiener filter [15]. This TSNR technique removes the reverberation effect while maintaining the benefits of the decision-directed approach. However, classic short-time noise reduction techniques, including TSNR, introduce harmonic distortion in the enhanced speech. To overcome this problem, a method called HRNR is implemented in order to refine the a priori SNR used to compute a spectral gain able to preserve the speech harmonics as proposed in [15].

A snore detection algorithm is then applied on each sample. This algorithm will window the sample (50ms of duration) and search for every spikes (or events) in this window by comparing it with the heuristic. The number of events needs to be superior to 1500 for the algorithm to continue its work. If it is the case, a more deep analysis is applied and only windows that contain events will be kept. These events need to be verified as snores to be outputted. To do so, the algorithm relies on k-harmonic to find MFCCs. If these can be found the event is classified as a snore.

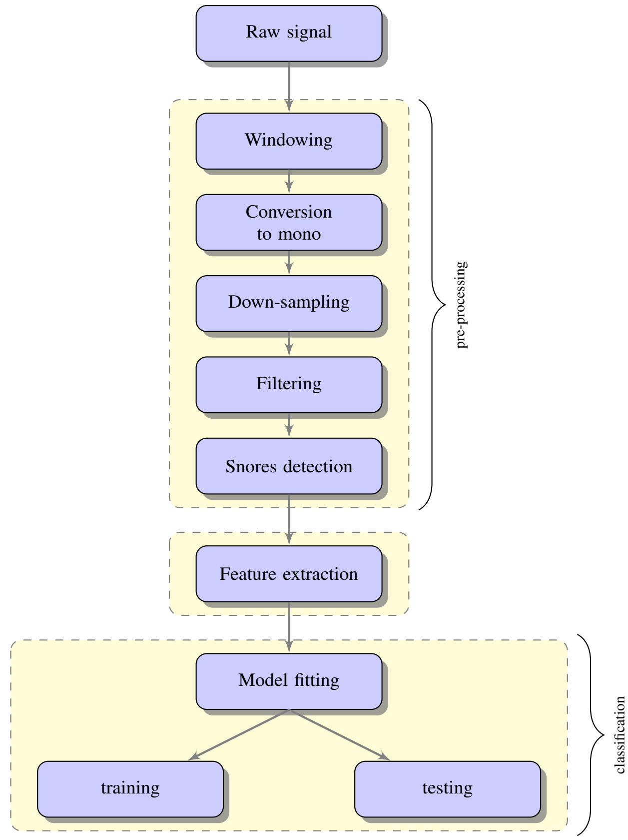

Fig. 1. Flow chart of the methods

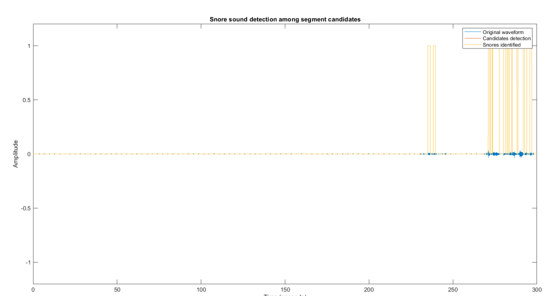

Fig. 2. Snores detected in a sample

The features are then extracted from these snores. For the time domain, only the snore duration is computed. For the frequency domain, however, more features had been chosen.

- Maximum frequency of given band: 'mf'
- Maximum average frequency of given band: 'mavgf'
- Widths of the bin at frequency mavgf: 'wl' and 'wu'
- Time difference between maximum average frequencies

maximum activation

mf, mavgf, wl and wu are computed for low and high frequencies.

In addition to studying and exploring these features and how separable the points are, we also tried classifying the snores. Although we classified a single snores to classes which were patients, in the future works one could classify the all snores into distinct sleeping conditions.

The main classifier used in this work was the **Random Forest Classifier** (RFC). It is a machine learning algorithm that consists of several uncorrelated decision trees. Their majority vote or average prediction is the final classification for a data sample. This method reduces the tendency of individual classification tree for overfitting resulting in more flexible and accurate classifier.

A decision tree with a maximum of 100 nodes is grown using the whole training data. All variables of a subset at each node are tested and the best one is selected, which is called greedy splitting. This is repeated multiple times to finally obtain a random forest consisting of multiple individual decision trees. These individual trees make predictions for a new unseen data sample, and their majority vote determines the predicted class of the sample. To know which input parameters lead to the best accuracy in the end a randomised search on hyperparameters was used. As a result, a number of trees in random forest of 400, a maximum number of levels in tree of 100, a minimum number of samples required to split a node of 5 and a minimum number of samples required at each leaf node of 1 is optimal for the greatest possible accuracy.

In addition to RFC, we explored three other classification algorithms: A single decision tree, Gaussian naive bayes based classifier and Quadratic discriminant analysis. None of these outperformed our main algorithm RFC. Table below (fig. 3) shows the basic statistics of each algorithm.

IV. RESULTS AND CONTRIBUTIONS

Algorithm	Precision	Recall	Accuracy	Time(s)
RFC	**0.78**	**0.78**	**0.78**	31
GNB	0.68	0.68	0.49	**0.01**
SDT	0.68	0.68	0.68	0.19
QDA	0.52	0.51	0.51	0.05

Fig. 3. Results of the classification where "RFC" = Random Forest Classifier, "GNB" = Gaussian Naive Bayes, "SDT" = Single Decision Tree and "QDA" = Quadratic Discriminant Analysis.

As noted before, the Random Forest classifier performed quite well with an accuracy of 0.78 and F1-score 0.77. The result was relatively surprising as when we tried to reduce the dimension of the feature vectors and plot the “snore clusters” separately, it was quite challenging. We explored three different dimension reduction algorithms (PCA, t-SNE and Autoencoder), and t-SNE performed the best visually. Yet the overall performance of the dimension reduction process into 2D was poor (fig. 4).

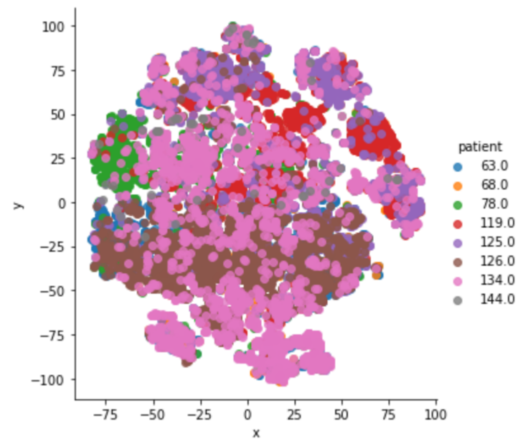

Fig. 4. t-SNE

V. DISCUSSIONS AND CONCLUSIONS

In this project, we focused on three themes. First we delved into characteristics of snoring sounds, secondly we came up with features that were extracted from the time series representing snoring. These features were explored and tried to plot in 2D with the help of three distinct dimension reduction algorithm. Lastly, we explored the potential of a classifier that could identify and detect certain snoring related sleep disorders.

During the project we faced a few challenges and limitations that we would like to comment.

The size of data

We had to work with data sets having very large sizes without an access to proper computer facilities. This heavily restricted our pipeline. Moreover, we encourage to consider applying data stream modelling with this type of data.

Lack of labelled data

Ideally we would have labelled data describing each snoring patient. As then we would focus on designing a classifier that could detect certain sleep disorders based on the snoring data instead of patients. This is critical if we wanted to create an online solution where anyone could upload a file of their snoring. The web application could then analyse the audio file with the help of our classifier.

Unsupervised learning methods requires more data samples and better feature extraction to be a promising approach. In medical context, one often prefer labelled data.

Exploring more features

Reducing the dimensions to plot the features reveals that most of snores share similar characteristics. For instance in figure 5 where the bottleneck layer of the autoencoder is reduced two two dimensional space with t-SNE, the projected points overlap with each other.

Becoming more familiar with the snores, segmenting it or using a template approach, could help us captivate the rich information in it better. This would also help us to discriminate the snore types more accurately.

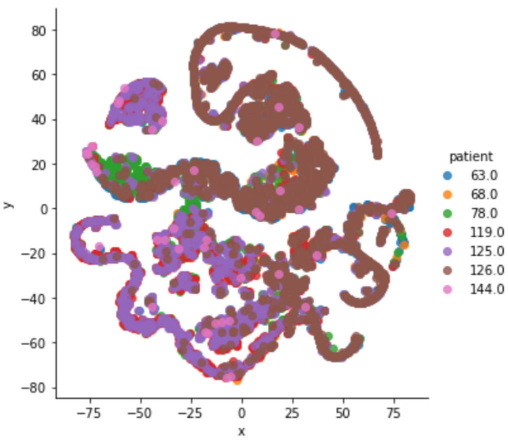

Fig. 5. Autoencoder with bottleneck reduction

Algorithmic engineering

Even if you had an access to the best algorithms, the nature of the data will affect your results. But assuming having well structured data, we recommend considering boosting, for instance Adaboost. Random Forest model is still strongly recommended as it is not a black box model. Hence, it could help us identify the critical questions regarding the snoring diagnostics. Unlike neural networks or other deep learning methods, RFC could answer questions like *which maximum frequency of a given band possibly correspond to sleep apnea* or *does the interval between snores tell something about the severity of the condition.*

REFERENCES

[1] Olanisun Olufemi Adewole et al. "Prevalence and correlates of snoring among adults in Nigeria". eng. In: *African health sciences* 8.2 (June 2008). PMC2584328[pmcid], pp. 108–113. ISSN: 1729-0503. URL: https://pubmed.ncbi.nlm.nih.gov/19357760.

[2] Hui Jin et al. "Acoustic Analysis of Snoring in the Diagnosis of Obstructive Sleep Apnea Syndrome: A Call for More Rigorous Studies". eng. In: *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 11.7 (July 2015). jc-00184-14[PII], pp. 765–771. ISSN: 1550-9397. DOI: 10.5664/jcsm.4856. URL: https://doi.org/10.5664/jcsm.4856.

[3] Ulla Anttalainen et al. "Prolonged partial upper airway obstruction during sleep - an underdiagnosed phenotype of sleep-disordered breathing". eng. In: *European clinical respiratory journal* 3 (Sept. 2016). 31806[PII], pp. 31806–31806. ISSN: 2001-8525. DOI: 10.3402/ecrj.v3.31806. URL: https://doi.org/10.3402/ecrj.v3.31806.

[4] Asaf Levartovsky et al. "Breathing and Snoring Sound Characteristics during Sleep in Adults". eng. In: *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 12.3 (Mar. 2016). jc-00201-15[PII], pp. 375–384. ISSN: 1550-9397. DOI: 10.5664/jcsm.5588. URL: https://doi.org/10.5664/jcsm.5588.

[5] Erna Arnardottir et al. "How to measure snoring? A comparison of the microphone, cannula and piezoelectric sensor". In: *Journal of sleep research* 25 (Nov. 2015). DOI: 10.1111/jsr.12356. URL: https://doi.org/10.1111/jsr.12356.

[6] J. R. Perez-Padilla et al. "Characteristics of the snoring noise in patients with and without occlusive sleep apnea". In: *Am. Rev. Respir. Dis.* 147.3 (Mar. 1993), pp. 635–644. DOI: 10.1164/ajrccm/147.3.635. URL: https://www.ncbi.nlm.nih.gov/pubmed/8442599.

[7] M. Cavusoglu et al. "An efficient method for snore/nonsnore classification of sleep sounds". In: *Physiological measurement* 28 (Sept. 2007), pp. 841–53. DOI: 10.1088/0967-3334/28/8/007. URL: https://doi.org/10.1088/0967-3334/28/8/007.

[8] M. Shokrollahi et al. "Snoring sound classification from respiratory signal". In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016, pp. 3215–3218.

[9] Can Wang et al. "Automatic snoring sounds detection from sleep sounds via multi-features analysis". In: *Australasian Physical & Engineering Sciences in Medicine* 40 (Dec. 2016). DOI: 10.1007/s13246-016-0507-1. URL: https://doi.org/10.1007/s13246-016-0507-1.

[10] Kun Qian et al. "Automatic Detection, Segmentation and Classification of Snore Related Signals from Overnight Audio Recording". In: *IET Signal Processing* 9 (Feb. 2015), pp. 21–29. DOI: 10.1049/iet-spr.2013.0266. URL: https://doi.org/10.1049/iet-spr.2013.0266.

[11] Kun Qian et al. "Classification of the Excitation Location of Snore Sounds in the Upper Airway by Acoustic Multi-Feature Analysis". In: *IEEE Transactions on Biomedical Engineering* 64 (Aug. 2016), pp. 1731–1741. DOI: 10.1109/TBME.2016.2619675. URL: https://doi.org/10.1109/TBME.2016.2619675.

[12] Shahin Amiriparian et al. *Snore Sound Classification Using Image-Based Deep Spectrum Features*. Aug. 2017.

[13] Michael Freitag et al. *An 'End-to-Evolution' Hybrid Approach for Snore Sound Classification*. Aug. 2017.

[14] João Ribeiro and Pedro Constantino. *ASNORE: Assessing snore health through classification of whole-night audio data*. 2019.

[15] Yi-Wen Chen. *MATLAB - Noise Reduction by wiener filter*. eng. 2019. URL: https://www.github.com/JarvusChen/MATLAB-Noise-Reduction-by-wiener-filter.