Low Level Texture Features for Snore Sound Discrimination

Fatih Demir[1], Abdulkadir Sengur[1], Nicholas Cummins[2], Shahin Amiriparian[2,3], Björn Schuller[2,4]

[1]Firat University, Technology Faculty, Electrical and Electronics Engineering Dept., Elazig, Turkey

[2]Chair of Embedded Intelligence for Health Care and Wellbeing, Augsburg University, Germany

[3]Machine Intelligence & Signal Processing Group, Technische Universität München, Germany

[4]GLAM – Group on Language, Audio & Music, Imperial College London, U.K.

INTRODUCTION

- Association between snoring and serious health risks, such as obstructive sleep apnea and heart disease.
- Adaption of **spectral** and **energy features** in particular for the task of snore sound detection.
- Motivation: recently proposed **Deep Spectrum** approach which utilised pre-trained image classification CNNs as feature extractor [1].
- Investigation of the use of **low level image texture features** extracted from spectrograms for snore sound classification.

DOWNLOAD

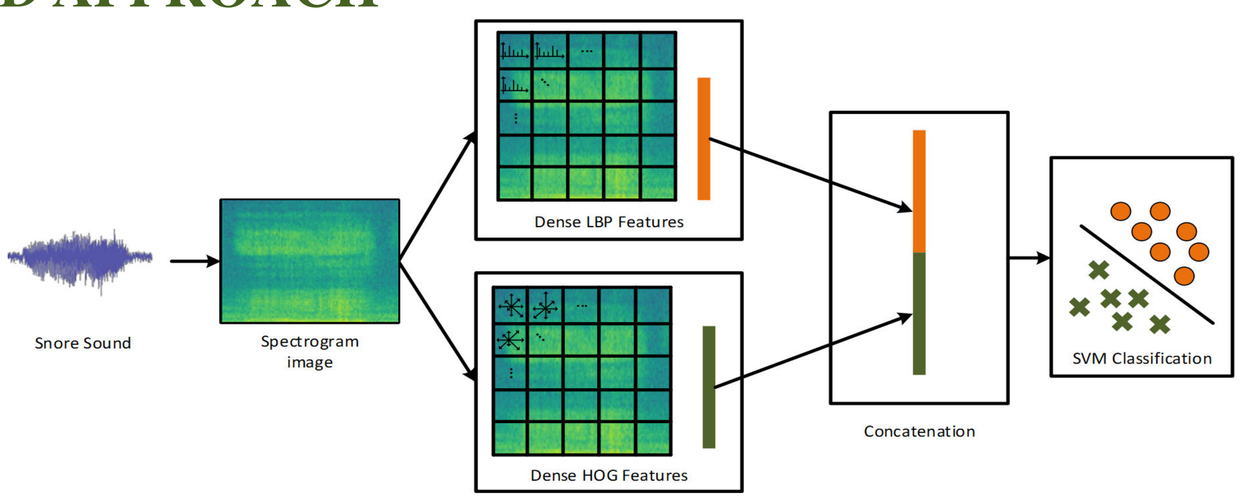
- The paper and poster are available to download via:

DATABASE

- The **Munich-Passau Snore Sound Corpus** (MPSSC), the database of the ComParE-2017 challenge [2].
- The dataset contains **828** snore audio files grouped into four categories: (i) **V**elum, (ii) **O**ropharyngeal lateral walls, (iii) **T**ongue base and (iv) **E**piglottis.

PROPOSED APPROACH

- First, the recorded snore sounds are converted to **spectrograms** images.
- We then extract **local binary pattern (LBP)** and **histogram of oriented gradients (HOG)** image descriptors from the spectrogram.
- Finally, classification is perform using a **Support Vector Machine (SVM)**.



RESULTS

- All results given in terms of *Unweighted Average Recall*

Feature	# Features	Channel	C (SVM)	Devel.	Test
LBP features	2891	Grey	0.1	33.0	-
	2891	R channel	10	34.2	-
	2891	G channel	0.01	34.5	-
	2891	B channel	1	34.2	-
	5782	R, G channels	0.1	33.1	-
	5782	**R, B channels**	**1**	**35.4**	**62.7**
	5782	G, B channel	10	33.9	-
	8673	**R, G and B**	**0.01**	**34.6**	**64.4**
HOG features	1296	Gray	1	37.7	60.2
	1296	R channel	0.1	36.1	-
	1296	G channel	0.01	37.1	-
	1296	B channel	0.1	34.4	-
	2592	R, G channels	0.1	37.3	-
	2592	R, B channels	100	35.5	-
	2592	G, B channel	0.01	35.0	-
	3888	**R, G and B**	**0.1**	**35.0**	**59.5**
LBP + HOG	**8374**	**R, B channels**	**0.1**	**36.8**	**69.0**
	8374	**G, B channel**	**0.1**	**37.8**	**72.6**
	12561	R, G and B	0.01	35.6	-

COMPARITIVE ANALYSIS

- All results given in terms of *Unweighted Average Recall*

Method	Devel.	Test
Baseline End-to-End	40.3	40.3
Baseline functionals	40.6	58.5
Deep Spectrum [1]	44.8	67.0
Deep Spectrum & CSO [3]	57.6	66.5
Proposed method	37.8	**72.6**

CONCLUSIONS

- LBP features appears better suited than the HOG features for snore sound classification.
- A combination of the color channels yields stronger results over using a single channel.
- Concatenation of the LBP and HOG descriptors achieved the highest score.
- Compared to the ComParE 2017 baseline, our approach achieved a **relative percentage increase** in unweighted average recall of **23.1%**.

REFERENCES

[1] S. Amiriparian, et al., "Snore Sound Classification Using Image-based Deep Spectrum Features," in Proceedings INTERSPEECH 2017, Stockholm, Sweden, pp. 3512–3516, ISCA, August 2017.
[2] B. Schuller, et al., "The INTERSPEECH 2017 Computational Paralinguistics Challenge: Addressee, Cold & Snoring," in Proceedings INTERSPEECH 2017, Stockholm, Sweden, pp. 3442–3446, ISCA, August 2017.
[3] M. Freitag, et al., "An 'End-to-Evolution' Hybrid Approach for Snore Sound Classification," in Proceedings INTERSPEECH 2017, Stockholm, Sweden, pp. 3507–3511, ISCA, August 2017.

Acknowledgements. This work was supported by the European Union's Seventh Framework and Horizon 2020 Programmes under grant agreements No. 338164 (ERC StG iHEARu)