THESIS/INTERNSHIP PROPOSAL

Subject:	**Automatic emotion recognition from the audio of respiration signal**
Supervisors and Co-supervisors:	**Radoslaw Niewiadomski**
Field of research:	**Affective Computing**
Motivations and general objectives:	Significant effort has been made over the last two decades in the field of automatic emotion recognition from nonverbal signals. The main focus has been on visual cues (e.g., full-body movements, facial expressions) and audio cues (e.g., prosody) of affective behavior, while other modalities have been rarely studied. Recent studies show that the respiration signal can be a promising modality for collecting data about human internal states [1]. It is expected that various emotional states can be recognized from the way humans breathe. The aim of this work is to: collect a new dataset of spontaneous affective reactions that includes the respiration signal, and develop supervised machine learning models for automatic emotion recognition using breath audio. In the first part of the thesis, a new dataset needs to be developed that contains breath audio recorded with a standard headset/boom microphone. The dataset will include spontaneous affective reactions, such as those occurring while playing video games or watching emotion-eliciting content on a PC, smartphone, or tablet. An important question to be addressed when creating the dataset is the choice of stimuli (e.g., video games) that can elicit specific emotional reactions in players under controlled conditions, the creation of methods for emotional self-assessment, multimodal data recording, and data annotation. The output of this stage is a set of short audio segments containing the respiration sounds, each labeled with an emotion label. Next, the student will propose a baseline model for signal analysis and classification of affective states from audio of breath. This includes using the standard feature-based approaches (e.g., [1]) that are used to process other audio data (e.g., extraction of MFCC, etc.), as well as the application of standard machine learning classifiers. The expected outcome of the thesis is a new dataset for studying affective respiration, as well as a simple classifier to identify certain emotions from respiration audio. The student will have the opportunity to develop innovative solutions and contribute to scientific publications.

Required skills:	• notions of audio signal processing • basic notions of deep/machine learning
Work Plan:	The student is expected to carry out the following tasks: • make a short survey of the literature on emotion elicitation in the lab conditions and on emotion recognition from the respiration signal, • design a set of experimental protocols (e.g., the choice of appropriate video stimuli or games) to collect the physiological data of spontaneous emotional reactions of the players, • collect the data recordings and self-assessments and deliver the dataset, • develop machine learning model using audio signals, • write the thesis report.
References:	[1] Lussu, V., Niewiadomski, R., Volpe, G., Camurri, A., The role of respiration audio in multimodal analysis of movement qualities, Journal on Multimodal User Interfaces, Volume 14, February 2020. *doi: 10.1007/s12193-019-00302-1*
Contacts:	radoslaw.niewiadomski@unige.it