PhD Program
in
TECHNOLOGY FOR HEALTH

Automatic segmentation, classification and analysis for a priori known syllable repertoires of mice vocalizations by supervised learning methods.

PhD Candidate: **Daniele Baggi**

Email: d.baggi@unibs.it

XXXVI Cycle

Tutor: Prof. Emilio Sardini, Prof. Pierangelo Migliorati

Background

Since the early 1970s, the functional role of USVs has been a controversial issue. Those sounds are considered as simple effect of a physiological response to a thermal challenge. Other researchers argue that USVs have a communicative purpose and can represent an emotional state. In particular, USVs of pups arouse maternal care and they are interpreted as an early communicative signal of mother-pup interaction. Even adult mouse USVs are signals of internal emotional states and facilitate social communication during non-aggressive encounters and in particular during mating behavior. USVs have been largely studied in a rodent behavioral perspective in the last decades, but only recently they are becoming an important tool for the behavioral characterization of mice model of neurodevelopmental disorders that are similar to the ones in the human being model, starting from communicative deficits such as autism spectrum disorders (ASD). Indeed, altered ultrasonic communication is reported in several mouse models of ASD. Many studies have shown that environmental enrichment can revert and/or prevent diseases linked to genetic alteration and environmental insults. Since the main environment component of a pup is its mother, maternal care turns out to be the earliest source of environmental intensification. Therefore, it can be useful to manipulate mother-pup interactions, housing pups with an additional non-lactating female from birth until weaning. This early maternal enrichment has been demonstrated to have long-term positive effects on brain and behavior in mouse models of neurodevelopmental disorders.

In this context, the team in the department of Molecular and Translational Medicine is studying ultrasonic communication in a mouse model of neurodevelopmental disorders, the p50-knock out (p50-KO) mice that have a deletion of a gene coding, compared with the wild-type (WT) model. They perform the USVs analysis by using an ultrasound sensitive microphone for the sound recording and a specific software to analyze the audio tracks. Then USVs are classified in ten different classes with a manual method based on spectral features such as frequency, duration, amplitude and by looking at the USVs' spectrogram shape.

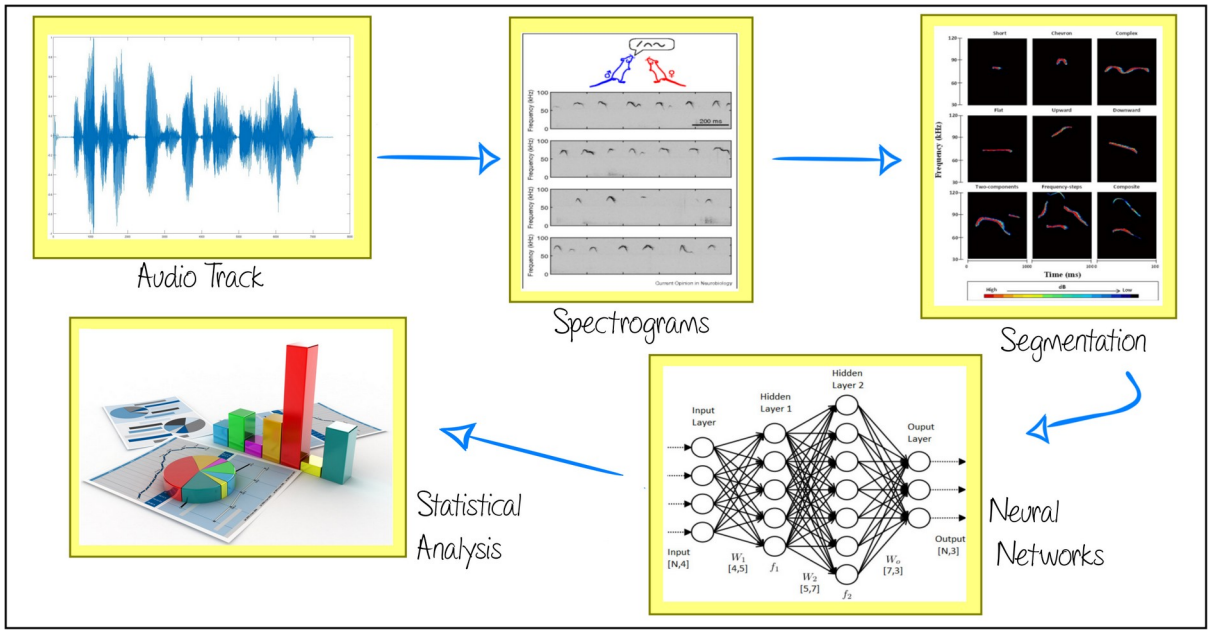
Objectives

Up to now the USV analysis is carried out manually with the support of a proprietary software that extract the audio track spectrograms with the aim to be manually segmented by the operator and finally the software compute and return some useful values relating to each segments

This project aims to automatize and speed up the mice USVs analysis process by automatically segment the audio track, classify the obtained segments into one of the vocalization classes and finally return a statistical analysis on the vocalization sequences.

Methodologies

The project, is composed by 3 main parts: the audio track *segmentation*, the USVs *classification* and finally the third part dedicated to a *statistical analysis* of the recognized vocalizations so that some pattern in the USVs sequences could be identified and linked to particular physiological conditions.



The audio track segmentation consists into moving from the whole audio track to a finite set of audio segments, each containing only a single vocalization included in the track. Unfortunately, the audio tracks are noisy and so a de-noising process or some noise attenuation methods should be considered in order to have good results. Once the audio track is completely segmented it could be useful to find a way to discard all the segments containing noise so that the classification task could be facilitated. It could be interesting to do it automatically by creating, for example, an additional noise class that could be added to the ten ones published by Scattoni et al so that a completely automatic system could be built.

The second part of the project is the classification one. For this task a large dataset is needed. A dataset is already created by a PhD student from the

department of Molecular and Translational Medicine. It was built by manually splitting the audio tracks with the AviSoft software and manually classifying the collected segments into ten classes. In order to achieve a complete automation of the system, it would be better if the learning algorithm that performs the classification is trained over a dataset of segments that are extracted by the designed automatic segmentation method and then manually labeled by a single human operator. Therefore, a new dataset should be created with the segments obtained by the designed segmentation method before moving to the next task.

Once the dataset is built the classification part can be implemented. Many learning algorithms could be tested during this phase. In the case of machine learning algorithms such as Support Vector Machines or Random Forest , the main problem is to find a set of features able to properly characterize each class of vocalizations making them distinguishable from each other. The features could be extracted directly from the audio signal or extracted from its spectrogram that can be considered as an image. Another possibility is to apply Deep Learning Algorithms such as Deep Neural Networks. Deep learning algorithms does not need an explicit phase of feature extraction because it is an intrinsic capability of the network, but the architecture must be optimized to obtain the best result.

Finally, when the segmentation and the classification phases provide good results, a statistical analysis could be implemented on the base of the useful information needed by the team that works on the mouse USVs. Indeed, the communication differences between different mouse genetic strains at the same physiological conditions are not linked to the vocalization type only, but are also linked to the number of vocalizations per class and their succession in a USV sequence.

To conclude the project a graphical user interface could be also implemented in order to make the automatic system more user friendly.

Some expansions of the project could also be considered. For example, it could be interesting to make the vocalizations analysis on both pup and adult mice. Again, it could be useful to analyze the association of the mice USV sequences with their displacements in the cage and their social interactions by analyzing the recorded video signals of their behavior. Other medical context can also be explored in order to find some different application for this type of analysis.

Expected Results and Impact

In the end of this project I want to find effective methods for segmentation and classification in order to go over the 85% of accuracy that is the higher result in literature. Once the classification is reliable enough some statistical analysis will be implemented in order to help the job of the research team with which I cooperate.

In general an artificial intelligence approach will automatize and speed up the process of vocalization selection and classification that is carried out manually up to now and so I think that this project will have a very good impact to the research in this field by shrinking the analysis times.