Canapy

Animal vocalization analysis and annotation tool

Keywords

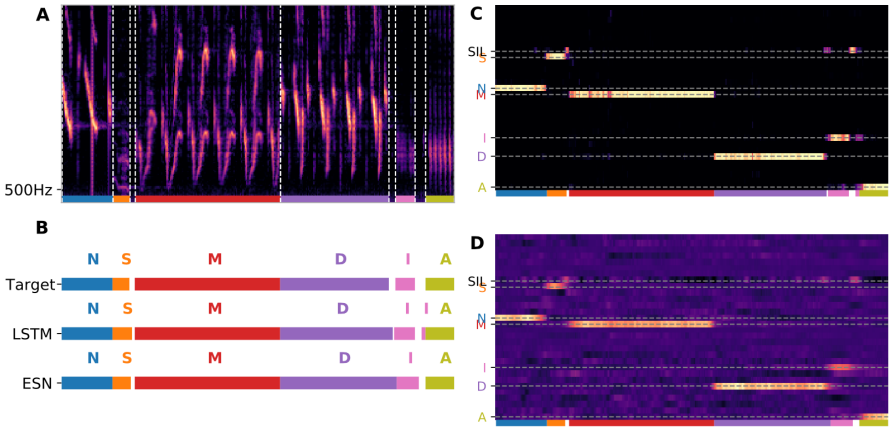
Machine Learning, Recurrent Neural Network (RNN), Reservoir Computing, Web development, User interface, Data analysis, MLOps

Supervisors

Xavier Hinaut, Researcher at Inria (Talence, France)
xavier.hinaut@inria.fr
Nathan Trouvain, PhD Student at Inria **(main contact)** (Talence, France)
nathan.trouvain@inria.fr
Arthur Leblois, Researcher at IMN (Bordeaux, France)
arthur.leblois@u-bordeaux.fr

Research team & internship location

Mnémosyne Team: Inria Bordeaux Sud-Ouest, LABRI & Institut des Maladies Neurodégénératives (Centre Broca Aquitaine, Carreire campus)
https://team.inria.fr/mnemosyne



Spectrogram of a canary song and recognition by two decoders (LSTM and ESN).

Introduction and scientific context

The study of songbirds is relevant for understanding how (human) children learn language, or more generally for understanding how our brains learn and process sound sequences (e.g. music, ...).

Songs structure analysis will help shed new light on experimental data, to better understand singing behavior and the underlying neural processes. In particular, it will be the main ingredient in the analysis and decoding of neural activity in birds' brain regions related to learning and song production.

However, retrieving this type of data requires to go through a slow annotation process, during which the sound recordings are sorted and labeled by hand. Recent developments [1, 2] provide Machine Learning (ML) tools to automate much of this process. But these tools are

difficult to master and to maintain. A first prototype of software based on a web interface developed in Python (*panel*, *bokeh*) offer a simple tool to set up a semi-automated annotation pipeline using a graphical interface. This tool intended audience is researchers or amateurs from any scientific horizon (biologists as well as computer scientists). This prototype could be greatly improved, in collaboration between the Mnemosyne team (developping the annotation tool [1]), the team at IMN studying the neural mechanisms of song learning in birds, and members of the international research community having also developed an annotation tool [2].

Objectives

The main objectives of the internship will be:

1. to develop a graphical interface to train vocalization annotation models, to visualize their performance and to re-annotate parts of the dataset accordingly (in a similar fashion as semi-supervised learning);

2. to develop the corresponding software backend: data management (audio and annotations), *serving* and local persistence of the models (MLOps);

3. to collaborate with the project members to define the needs, establish the specifications or integrate pre-existing tools. This objective also implies collaborating with international researchers, and making an open source tool available to the public.

The development will be incremental: a first prototype will allow to train models and to present their evaluation on the interface. A second prototype will offer advanced editing possibilities of the dataset (re-annotation of parts of the audio according to the results of the model), and the final version will integrate advanced analysis tools (dataset errors detection, spectrograms dimensionality reduction for visualization and/or clustering, syntactic analysis of song sequences, ...)

Methods

The student will have to develop an interface, preferably web, in *javascript*/*typescript* (*React*...) or directly in *Python* (*bokeh/panel/holoviz*...).
The software backend will serve Machine Learnnig models defined in Python (type *scikit-learn*/*reservoirpy* at first, eventually type *tensorflow/pytorch*).
The tool could be inspired by or integrated with the *VocalPy* initiative [3]. The student will be encouraged to collaborate with the project collaborators. For example, the data could follow the convention defined by the *VocalPy crowsetta* package.

Once complete, the tool will be made public, on Github, along with its documentation. The goal is to impact a large international community, like *ReservoirPy* [4], a library already developed in the Mnemosyne team for the ML community.

Required skills

- Expertise in web/frontend programming

- Expertise in Python

- Solid knowledge in software project management

- General knowledge in machine learning (training procedure, evaluation, metrics, visualizations, Python tools...) is a plus

- General knowledge about artificial neural networks

Applications

Deadline: as soon as possible; as soon as a good application is received it will be considered. Please contact Nathan Trouvain for any questions, and send your application (CV + cover letter) to the three supervisors of the project: Nathan Trouvain, Xavier Hinaut and Arthur Leblois. (Emails at the top of the internship proposal).

References

[1] N. Trouvain et X. Hinaut, « Canary Song Decoder: Transduction and Implicit Segmentation with ESNs and LTSMs », in ICANN 2021 - 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, sept. 2021, vol. 12895, p. 71-82. doi: 10/gq43sk.

[2] Y. Cohen, D. A. Nicholson, A. Sanchioni, E. K. Mallaber, V. Skidanova, et T. J. Gardner, « Automated annotation of birdsong with a neural network that segments spectrograms », eLife, vol. 11, p. e63853, janv. 2022, doi: 10/gq43sd.

[3] « VocalPy ». https://github.com/vocalpy

[3] « ReservoirPy ». https://github.com/reservoirpy/reservoirpy