

Research Topic: *Deep Learning, Audio Processing*  
Title: *"Advances in Deep Learning for Audio Processing"*

### **A. Background and Significance**

Machine learning methods have been utilized for the design and development of computer-aided tone analysis for over a decade now. These automated diagnostic systems are meant to provide an interpretation and baseline with quick and accurate diagnosis/detections with an intention to aid in interpretation of language and advances in communication. [1-9]. Over the past few years deep learning strategies have found their use in the baby cry processing domain for segmentation and classification tasks. However, unlike natural adult language, baby cry suffer from significantly fewer choice of words to be descriptive and lower contextual variations along with added challenges imposed by similar sound cues, variabilities across by phonic uniqueness.

In this project, we will survey several methods where domain knowledge guided features and machine learning methods match or outperform recent deep learning methods that were not strategized based on domain expertise. Our goal is to build methods that enhance diagnostic efficiency, reproducibility and method explainability with baby cries, where "large data sets" do not exist.

### **B. Method and Deliverables**

We will implement a variety of deep learning methods to answer some of the following research questions:

1. Does transfer learning apply between different types of baby cries, especially for varying degrees of vowel selection and vowel stretching?
2. It is well established that segmentation is the best form of diagnosis. Is there a good way to generate augmented data for the baby cries segmentation tasks?
3. Most baby cries have variations. Can converting single baby cry to grouping of meaning aid data augmentation?
4. How can we learn features of interest (explainability) using deep learning models?

### References

[1] Petit Journey "Understand the different cries of your baby"[Online]:  
<https://www.petitjourney.com.au/understand-the-different-cries-of-your-baby/>

[2] Audeering. "audinterface provides user interfaces to apply any machine learning model or digital signal processing algorithm to audio files or databases" [Online]:  
<https://audeering.github.io/audinterface/usage.html/>