

### **Cross-Correlation Based Template Matching**

The goal with our pipeline was to create a product that could reliably and quickly create labeled audio data that could be used for training machine learning models. Our solution was the cross-correlation pipeline, which we use to label data automatically. At a basic level, the cross-correlation pipeline takes in a template audio file for a given bird's call and runs it across a longer audio file, identifying where this template sound or similar occurred. These similarities can allow us to identify the presence of a species in an unlabeled audio clip and label it for use in machine learning.

If we can improve the workflow and speed up the process of gathering training data without sacrificing accuracy, we will be able to learn more about the biodiversity of different species. Monitoring wildlife gives us a better understanding of rare, threatened, or endangered species to help us reallocate funds and provide assistance and protection to those species that need it.