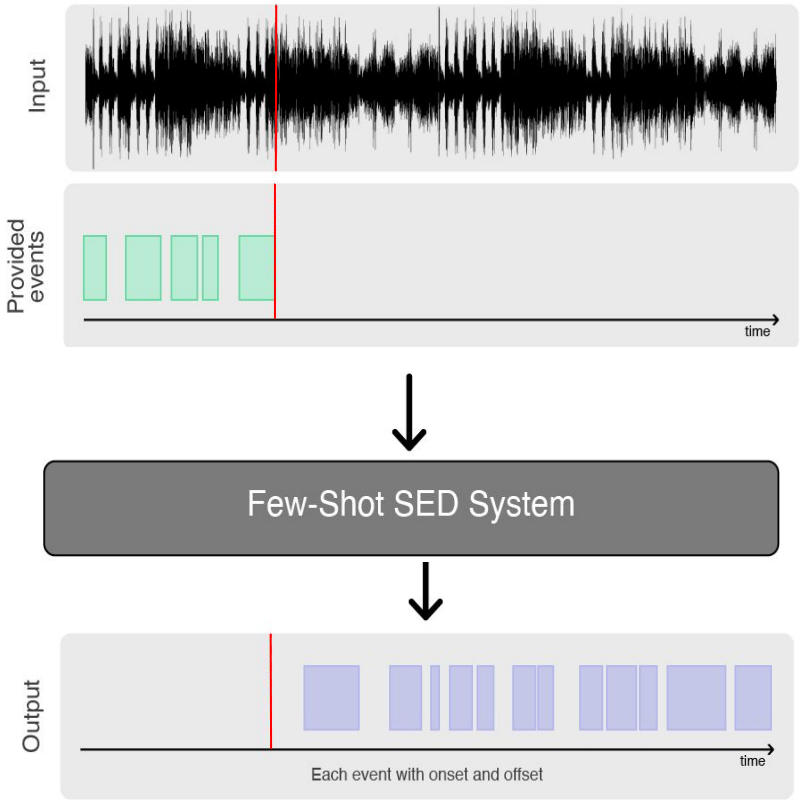
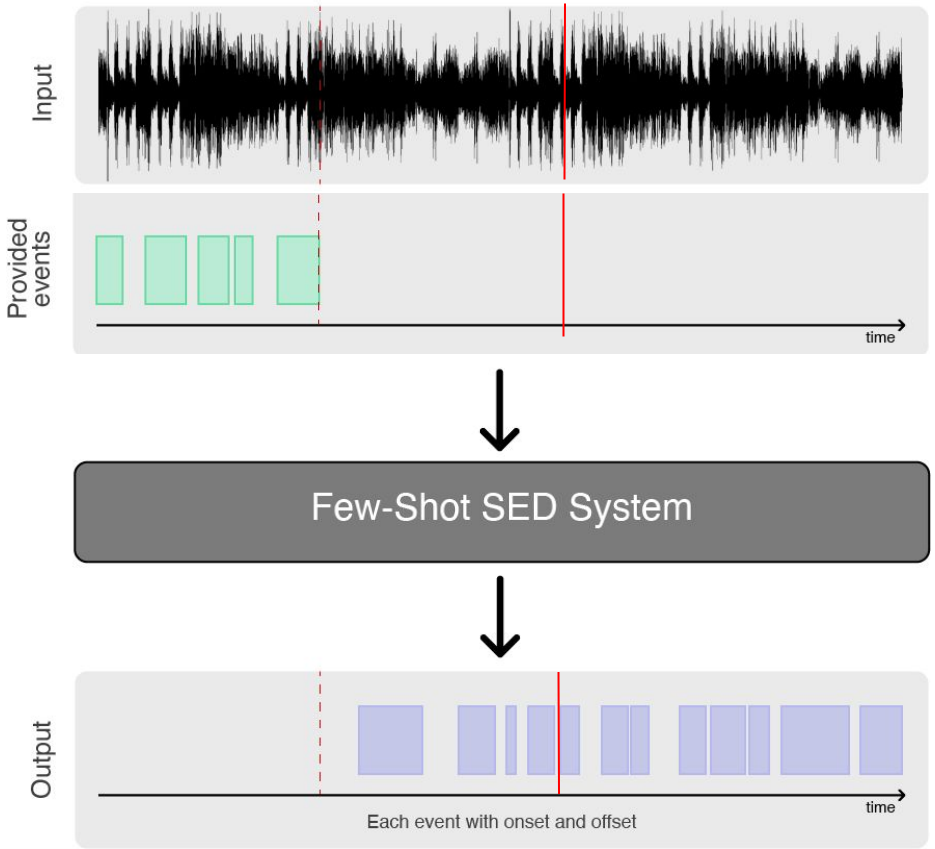
FEW-SHOT BIOACOUSTIC EVENT DETECTION

Ines Nolasco, Shubhr Singh, Ester Vidana Vila, Vincent Lostanlen, Ariana Strandburg-Peshkin, Emily Grout, Lisa Gill, Hanna Pamula, Joe Morford, Michael Emmerson, Frants Jensen, Helen Whitehead, Ivan Kiskin, Veronica Morfi and Dan Stowell.



Task Description



Few Short Learning

Meta Learning

Learn on support set, predict on query set.

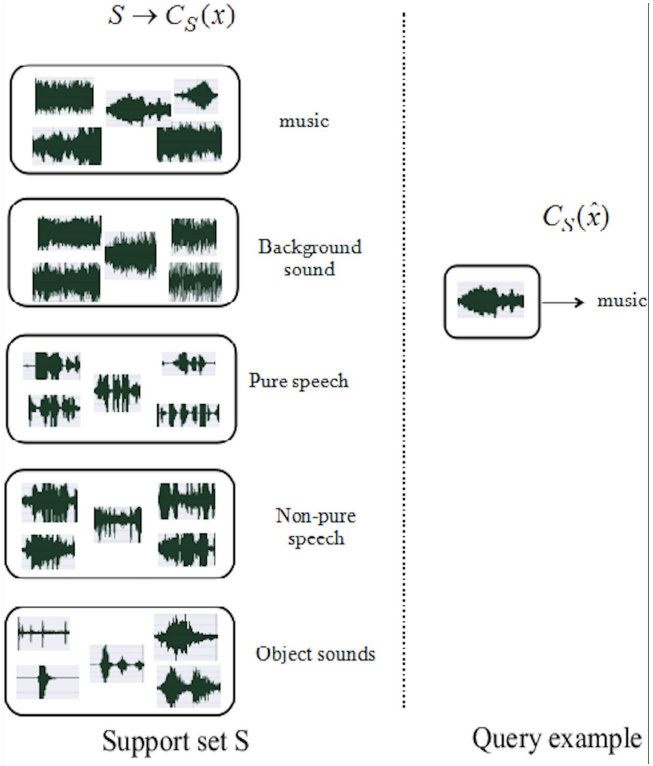
Learn embeddings using a distance/similarity loss function

Few Shot Learning for Audio

Transfer Learning

Use a Pretrained network

Fine tune model on data augmented support set



Zhang et. al. 2019

Challenges

- **Sparse Sound Events -** The sound events in the audio files are sparse with background noise, making it difficult to learn robust embeddings.
- **Sound Event Duration** - Length of sound events in training set is different to validation/evaluation set, making it difficult to predict the temporal boundaries of the sound events.
- **Data augmentation** - Augmenting on few data samples might lead to a skewed data distribution, leading the model to overfit.
- **Transfer Learning** - Fine tuning a pre-trained model on a few data samples hinders feature migration.
- **Distinct tasks** - Meta learning framework not only samples the data space but also the task space, hence when the training task and the target task are distinctly different, the effect of meta-learning is minimal (Song et. al. 2022)

Thank you to all the organizers and participants!