PROGRESS 1: MARIO METHOD

BIRD SONG IDENTIFICATION

Thanyaporn Phinthuphan

1 Oct 2018

Outline

- Recap about the paper
- Preprocessing
- Segmentation
- Feature generation
- Classification & Result
- Problem & Next step

Bird Song Classification in Field Recordings

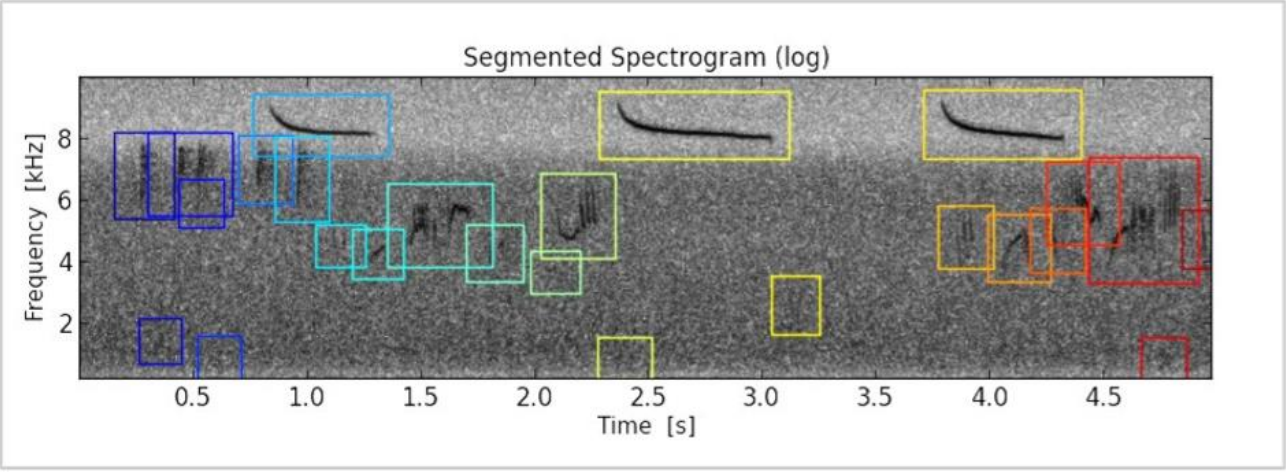
Introduction

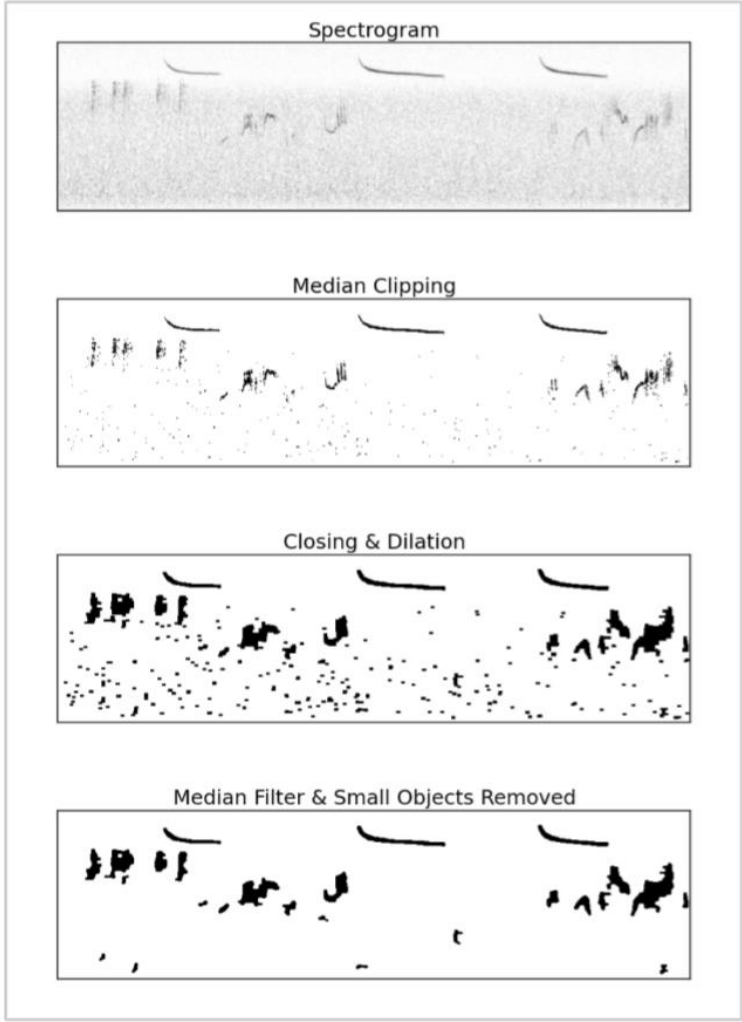
- author: Mario Lasseck
- the winning Solution for NIPS4B 2013 Competition
- 87 sound classes of birds (call/song)
- 687 audio file (WAV format) in the training set (length 1-5 sec)
- 1,000 test file

Preprocessing and Segmentation

Bird Song Classification in Field Recordings

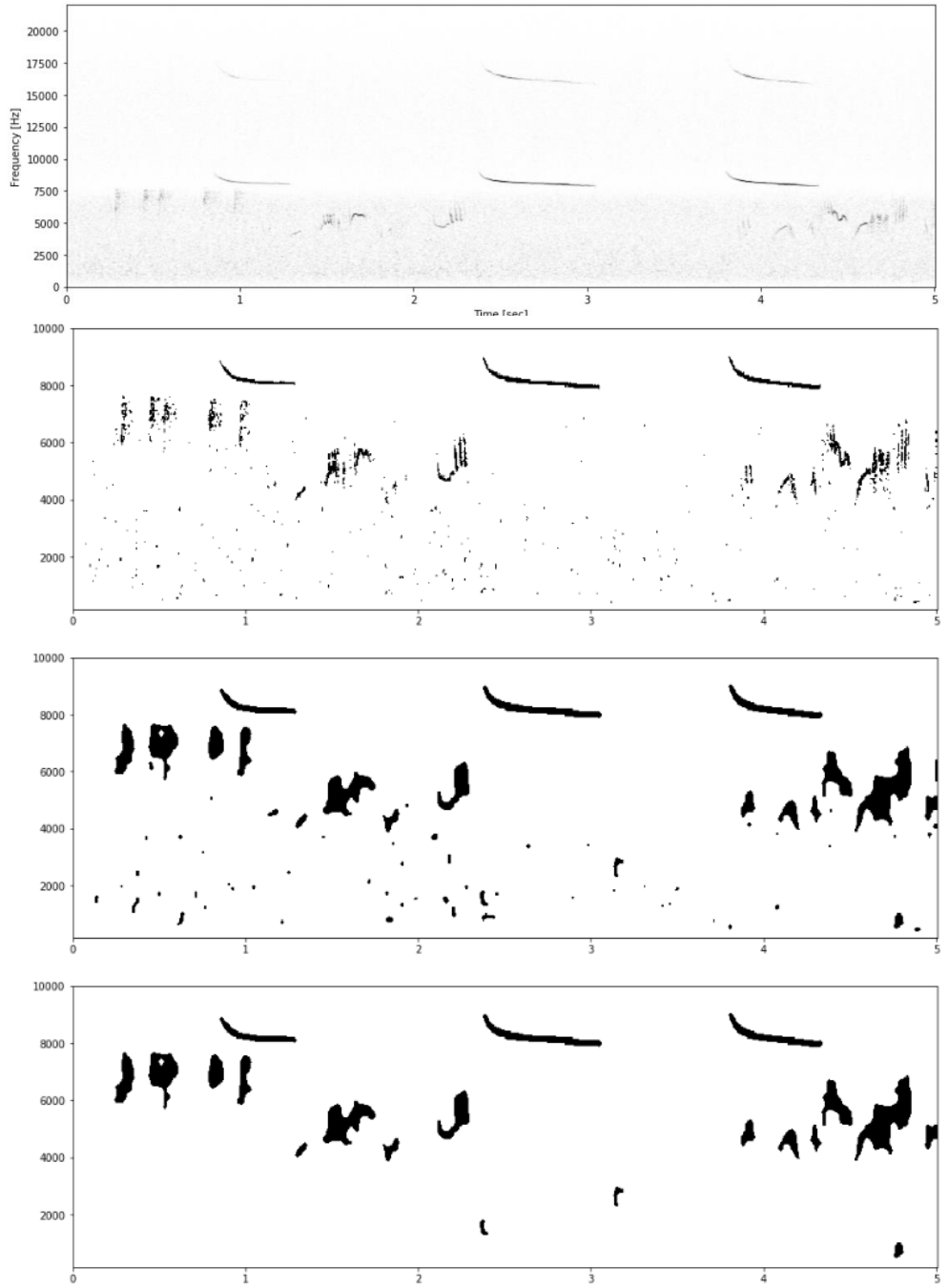
- STFT using hanning window → normalized
- reducing background noise with median clipping
- closing & dilation → segmentation (size/position)



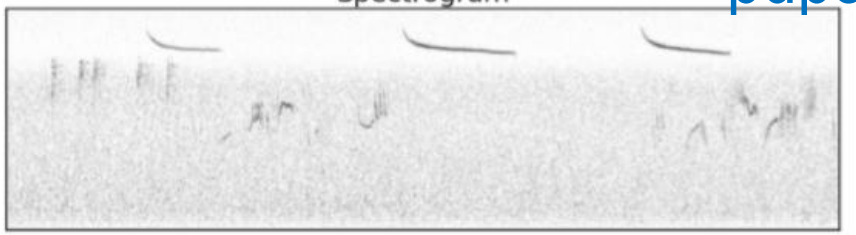


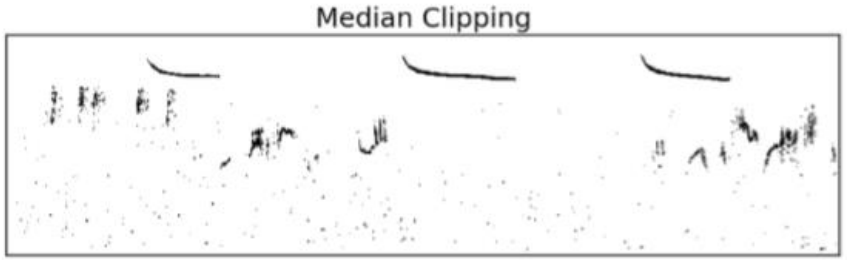
Preprocessing

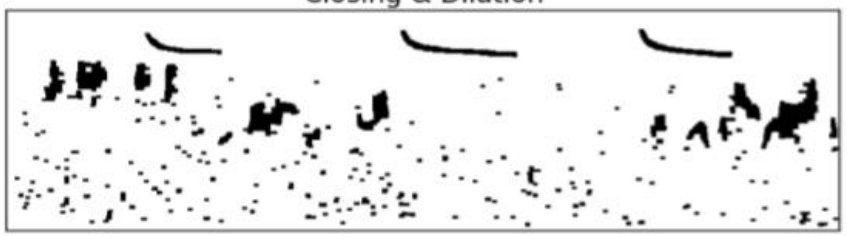
- using OpenCV
- only tune parameter for closing, dilation, blur

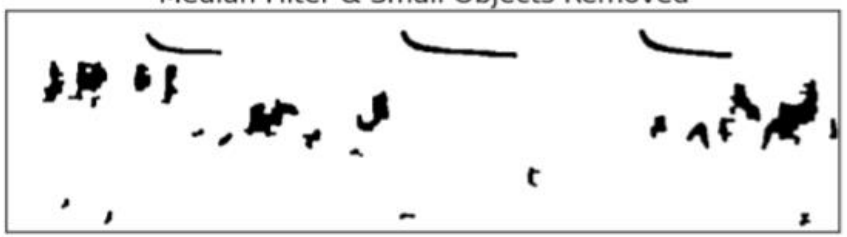


paper

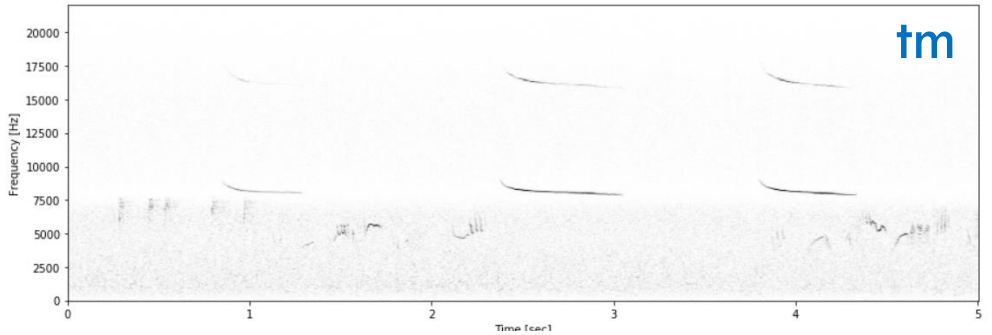

Spectrogram

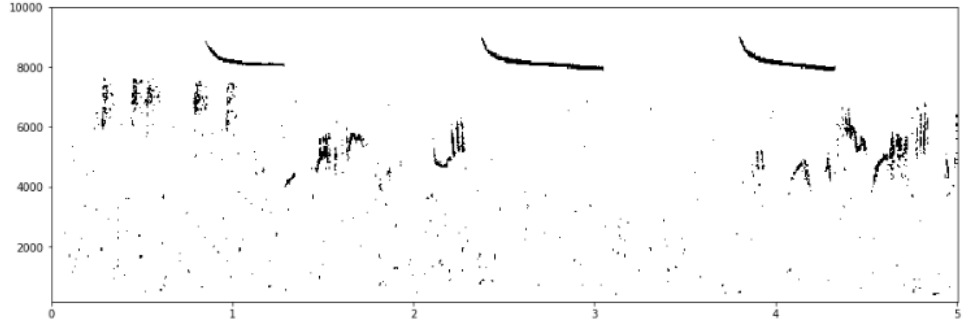

Median Clipping

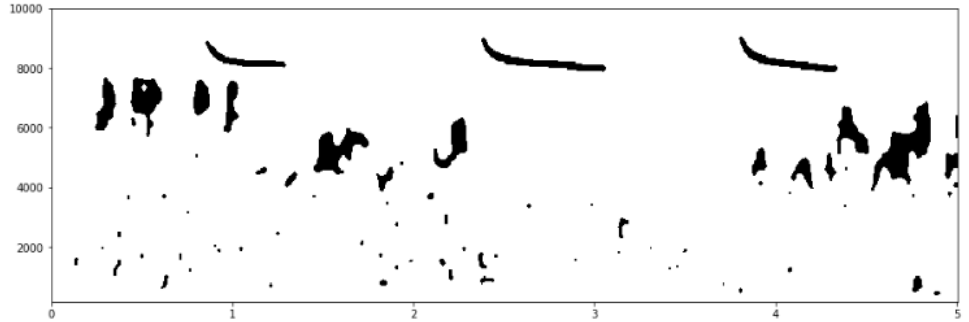

Closing & Dilation

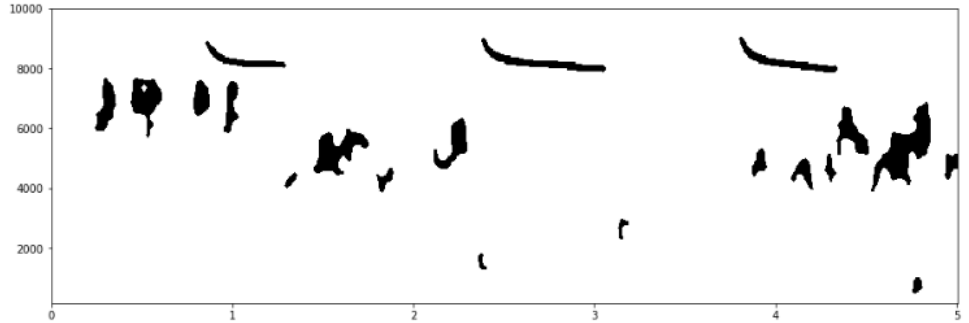

Median Filter & Small Objects Removed

tm

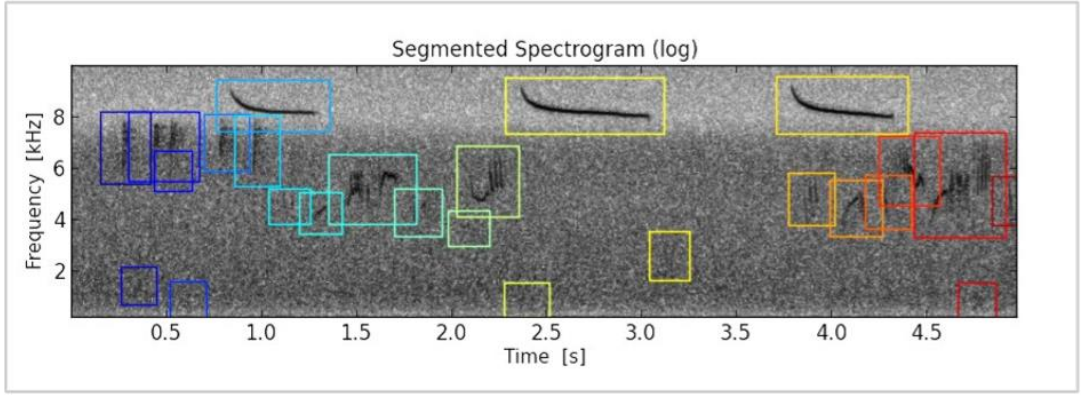

Frequency [Hz]
Time [sec]



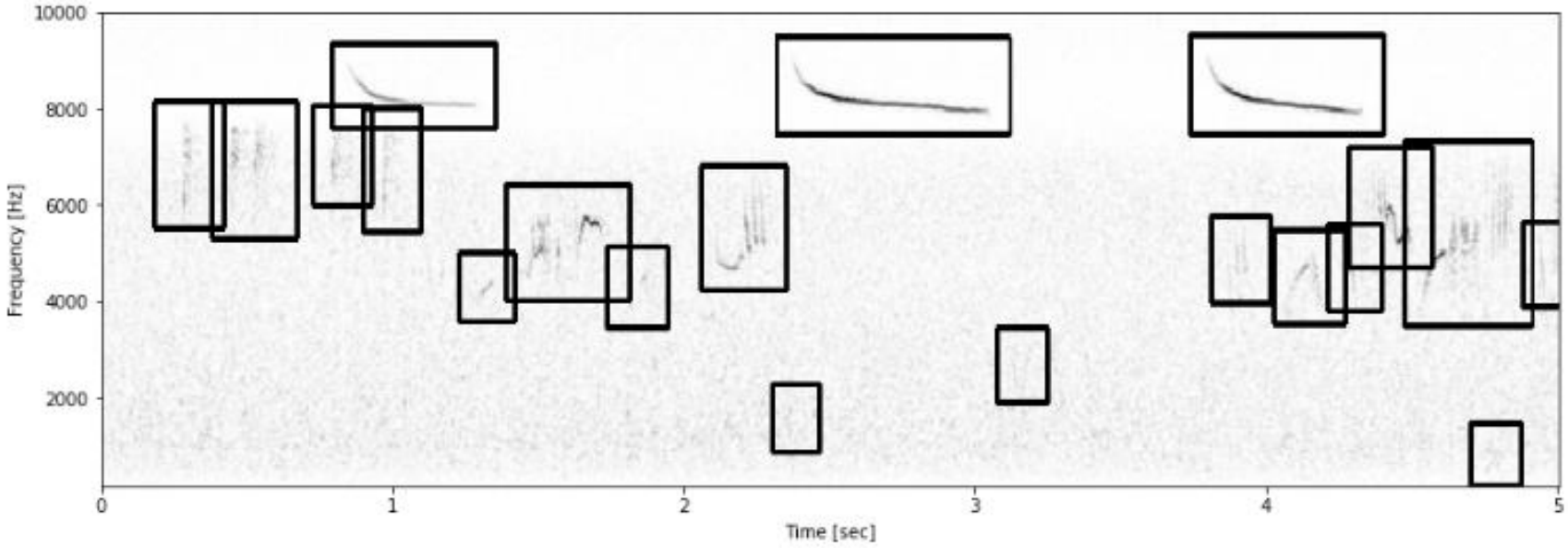




Segmentation



- 7,821 segments (paper has 9,198 segments)




```
Total segment: 20
```

Feature Generation

Bird Song Classification in Field Recordings

- File-statistics

  max, min, mean, std for all values of spectrogram + 16 divided spectrogram

- Segment-statistics

  count + max, min, mean, std for weight, height, frequency position

- Segment-probabilities

  highest matching all segments using normalized cross-correlation

- 68 + 13 + 9,198 (number of segments in training) features per file

Feature Generation

- File-statistics

  max, min, mean, std for all values of spectrogram + 16 divided spectrogram

- Segment-statistics

  count + max, min, mean, std for weight, height, frequency position

- Segment-probabilities

  highest matching all segments using normalized cross-correlation

- 7,821 (number of segments in training) features per file

Bird Song Classification in Field Recordings

Feature Extraction

- File-statistics

  max, min, mean, std for all values of spectrogram + 16 divided spectrogram

- Segment-statistics

  count + max, min, mean, std for weight, height, frequency position

- Segment-probabilities

  highest matching all segments using normalized cross-correlation

- 68 + 13 + 9,198 (number of segments in training) features per file

Classification & Result

- using 87 random forest
- score of 90.17% AUC

Multi-label Bird Species Classification - NIPS 2013

Identify which of 87 classes of birds and amphibians are present into 1000 continuous wild sound recordings

32 teams · 5 years ago

Public Leaderboard | Private Leaderboard

The private leaderboard is calculated with approximately 67% of the test data.
This competition has completed. This leaderboard reflects the final standings.

Refresh

Gold Silver Bronze

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲2	Mario			0.91751	29	5y
2	▼1	les bricoleurs			0.91577	34	5y
3	▼1	Rafael			0.91251	38	5y
4	▲2	DB2			0.89624	15	5y

Submission and Description	Private Score	Public Score
sub002.csv 8 days ago by tmmm first result	0.90170	0.90407

Problem

- Runtime ~12 hours !!!

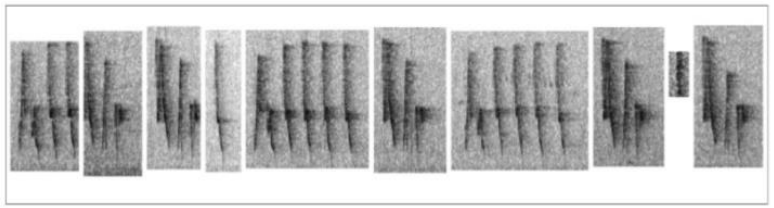
  (Including generate train/test feature (seg-prob)

Next Step

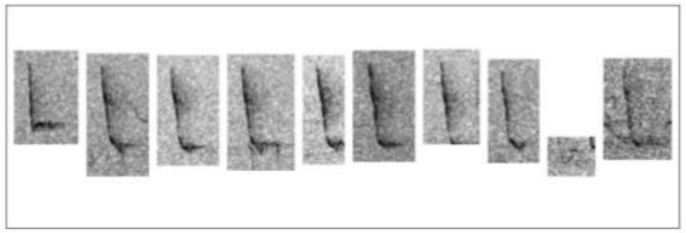
- Try to use these feature/method on my dataset
- Find the way to identify on continuous sound
- Extract only important features to reduce runtime

Bird Song Classification in Field Recordings

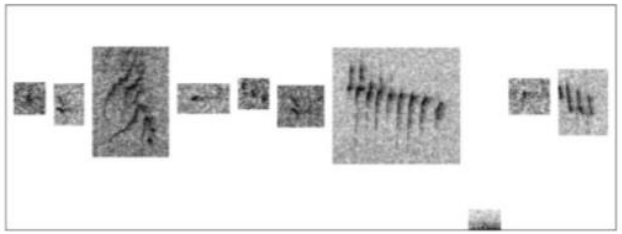
Classification



song of Cettia cetti



song of Phylloscopus collybita



call of Serinus serinus

Q&A