

# BIRD AUDIO DETECTION FOR DCASE 2018 CHALLENGE TECHNICAL REPORT

## Technical Report

Lianjie Tao, Xinxing Chen

**E-mail:** taolianjie007@cqu.edu.cn

The 2018 BAD challenge [1] requires to determine bird audio in a 10 seconds sound clips, the organizer gave us three development datasets for training our NN, and three evaluation datasets to evaluate our NN. The goal of the challenge is to maximize the recognition of audio in the birds.

**BAD Methodology :( results in 87.75% on <http://lsis-argo.lsis.org:8005/scores>)**

### Features log Mel-Spectrogram

Sample rate: 1.6khz

Frame size: 1024 samples

Hop size: 360 samples

Mel-scaled filter bank: 64

**Inputs :( 240, 64, 3)**

**Data-Augmentation** (we do a lot of work about this, but the result perform not well, so the last result do not use data-augmentation)

**Model** – Convolutional Neural Network (CNN)

We used the classical Neural Network- Xception [2], used keras with Tensorflow Backend, trained on a NVidia GTX X (Pascal) with 11 GB of RAM, training about 20 epochs with a pre-train model can achieve an 87.75% preview score.

### Xception Structure

The structure is completely referenced to xception, only the last fully connected layer is modified.

Some time later, I will submit my experiment code to my respository for reference.

<https://github.com/nicktao9?tab=repositories>

### References

[1] D. Stowell, Y. Stylianou, M. Wood, H. Pamuła, and H. Glotin. *Automatic acoustic detection of birds through deep learning: the first bird audio detection challenge*. Methods in Ecology and Evolution, 2018.

[2] Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017:1800-1807.