PROGRESS 5: MODEL PRE-RESULT

BIRD SONG IDENTIFICATION

Thanyaporn Phinthuphan

11 Apr 2019

Outline

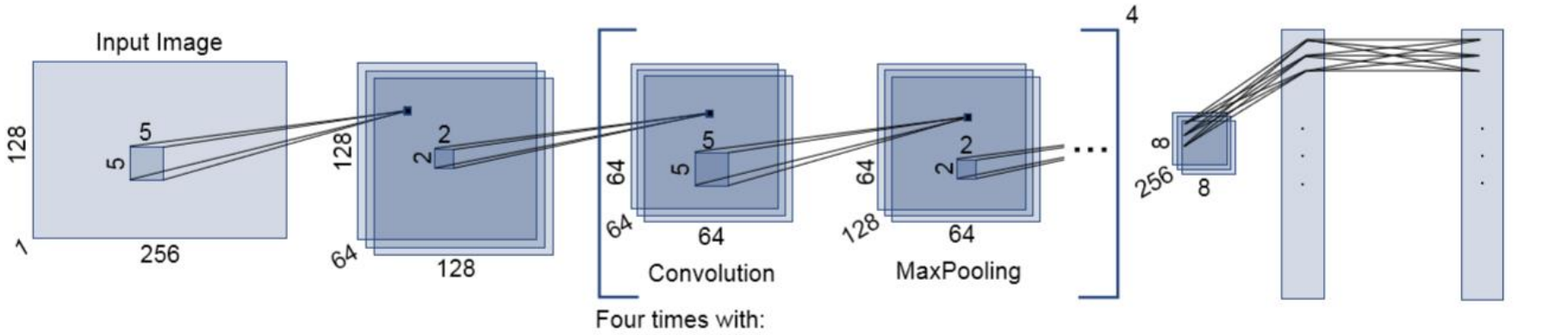
- Recap
- Methodology
- Preliminary Result
- In progress
- Next step

Recap

- data with 11 bird sound class 1,847 files (100-200 per class)
- using CNN model based on winning solution of BirdCLEF 2016

Recap

- separate signal / noise part with median clipping method
- split spectrogram of signal part into chunks 512 pixel
- not apply data augmentation



Methodology

Training data

- train on A/B quality files 1,614 files (87%)
- class imbalance → selected ~70 files/class = 766 file (41%)
- train : test = 90 : 10

Methodology

Data chunks

- 4,349 training chunks / 516 testing chunks
- all file → 12,432 chunks
- train CNN with 0.001 learning rate ~1,000 epoch

Methodology

Model evaluation

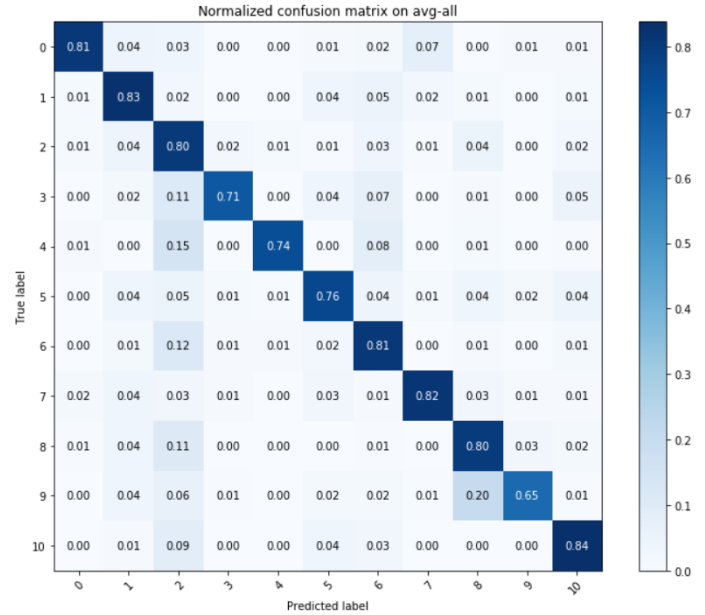
- accuracy on chunks
- accuracy on files
  - simple mode
  - average mode with confidential values
  - only one true
- confusion matrix
- mean average precision (MAP)

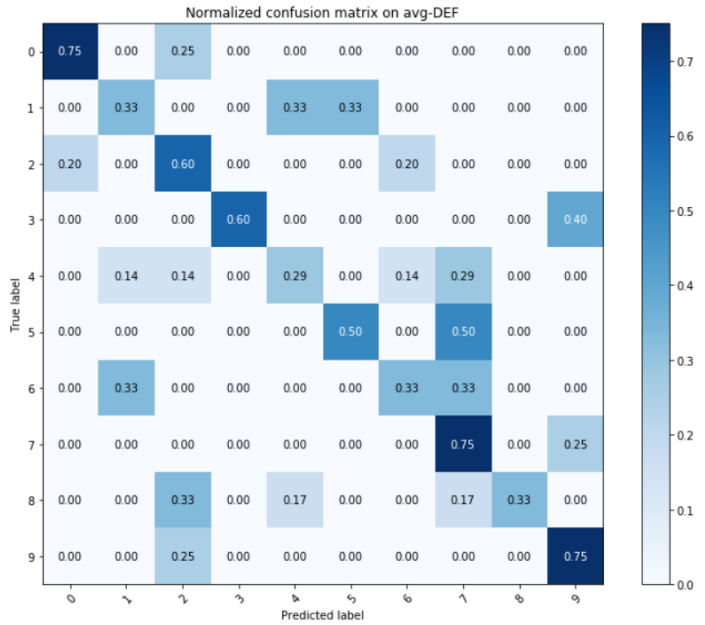
Preliminary result

Accuracy

	chunk	mode	avg	some
train	86.76	91.15	92.31	95.21
test	85.85	81.82	84.42	89.61
all	75.01	74.93	76.77	83.6

all	76.77
A	86.31
AB	80.17
ABC	77.38
DEF	51.16

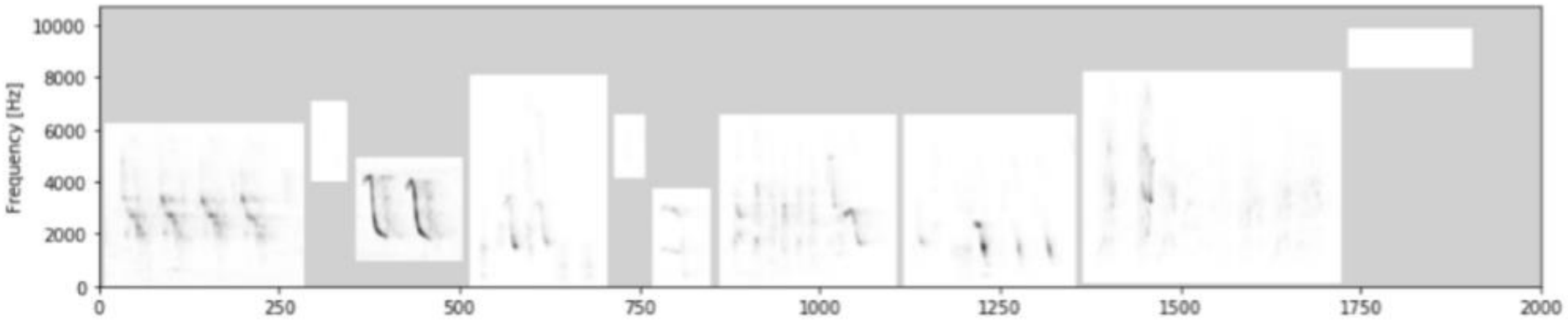




Preliminary result

Discussion

- the worst class is นกเอี้ยงสาริกา, same as former model
- test by recording real bird sound at park → too much error



ภาพที่ 4.4 ภาพแสดงชิ้นส่วนสำคัญในการจำแนกเสียงนกเอี้ยงสาลิกา (AUC score 0.6227)

In progress

New scope data

- select bird that very common / have to identify with sound in Thailand

  → query by science name on xeno-canto.org
- some species are restricted due to conservation concerns → cannot download

  (นกกางเขนบ้าน นกกางเขนดง นกขมิ้นท้ายทอยดำ นกขุนทอง นกเอี้ยงด่าง)
- and select only class that have more than 10 files → 80 class / 8,746 files
- goal: try to use most class → if some class not work, remove them!

In progress

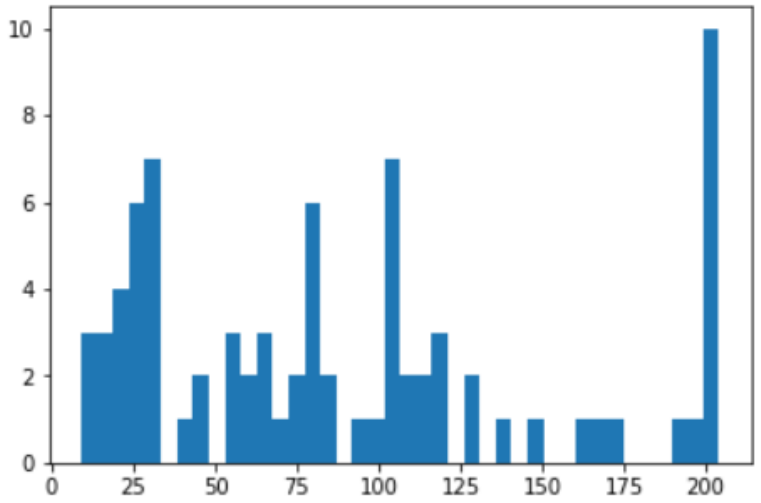
New scope data

- select training data with quality A/B/C with not more than 200 files/class

  → 7,170 files (82%)

  (add C because of number of files and to help handle user recording)

- 30,532 training chunks / 2,863 testing chunks



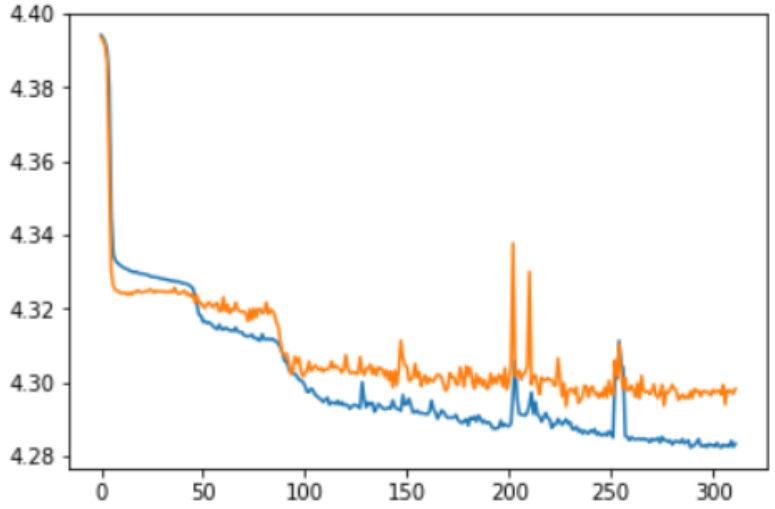
In progress

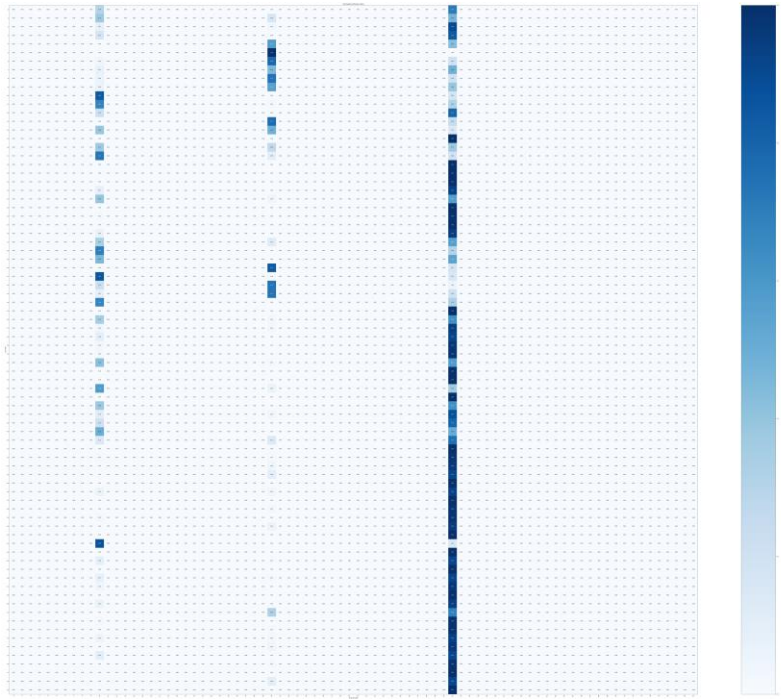
Problem

- CUDA error: device-side assert triggered
- OSError: [Errno 12] Cannot allocate memory

- try to train on 20% of files with 300 epoch

  → 13% of accuracy / predict only 3 class!





Next Step

- train model on 80 class
- find the way to keep accuracy on user record data

Q&A