

PROPOSAL

BIRD SONG IDENTIFICATION

Thanyaporn Phinthuphan
19 Nov 2018

Outline

- Introduction
- Literature review
- Methodology
- Result
- Conclusion & Next step

Introduction

Background & Objective

เพื่อพัฒนาโปรแกรมสำหรับจำแนกชนิดนกจากเสียงร้อง โดยใช้วิธีการทางคอมพิวเตอร์ และ machine learning ในการประมวลผล เพื่อนำไปใช้ประโยชน์ในการศึกษาเกี่ยวกับนกในประเทศไทย และการสำรวจประชากรนกในแต่ละท้องถิ่น

Introduction

Scope

- จำแนกเสียงนกในประเทศไทย โดยมีความตั้งใจจะพัฒนาให้รองรับเสียงนกมากกว่า 40 ชนิด
- เป็องต้นทดลองบนเสียงนกจำนวน 11 ชนิด ซึ่งเป็นนกที่พบได้บ่อยในกรุงเทพฯ



Introduction

Planning

- รวบรวมข้อมูลเสียงนกที่พบได้บ่อยในกรุงเทพฯ เพื่อใช้ในการสร้างและทดสอบโมเดล
- ศึกษางานวิจัยที่เกี่ยวข้อง เพื่อหาแนวทางและตัวอย่างขอบเขตของโครงการ
- สร้างโมเดลโดยอ้างอิงจากวิธีการของงานวิจัยตัวอย่าง
- ทดสอบโมเดลบนข้อมูลเดียวกันกับงานวิจัยตัวอย่าง เพื่อเปรียบเทียบประสิทธิภาพของโมเดล
- นำโมเดลมาประยุกต์ใช้บนข้อมูลเสียงนกที่รวบรวมไว้ และทดสอบประสิทธิภาพเทียบกับข้อมูลจากงานวิจัย
- พัฒนาโมเดลเพิ่มเติมเพื่อให้มีประสิทธิภาพดีขึ้น ถูกต้องแม่นยำมากขึ้น และรวดเร็วขึ้น
- พัฒนาโปรแกรมในส่วนของฟังก์ชันการใช้งาน ให้สามารถใช้งานได้ง่าย และสวยงาม

Introduction

Benefits

- สามารถนำโปรแกรมที่ได้ไปใช้จำแนกชนิดนกจากเสียงร้องตามธรรมชาติได้จริง
- ผลลัพธ์ที่ได้มีความแม่นยำมากกว่าร้อยละ 90 ขึ้นไป

Literature review

- Spectrogram
- Short-time Fourier transform
- Random forest
- AUC
- K-Fold Cross Validation

Literature review

- Bird Song Classification in Field Recordings
- Winning Solution for NIPS4B 2013 Competition by Mario Lasseck
- จำแนกเสียงนกจำนวน 87 ชนิดที่อัดจากสถานที่ต่าง ๆ ในฝรั่งเศส
- ผลของงานวิจัยนี้ ได้คะแนน AUC score เท่ากับ 91.75%

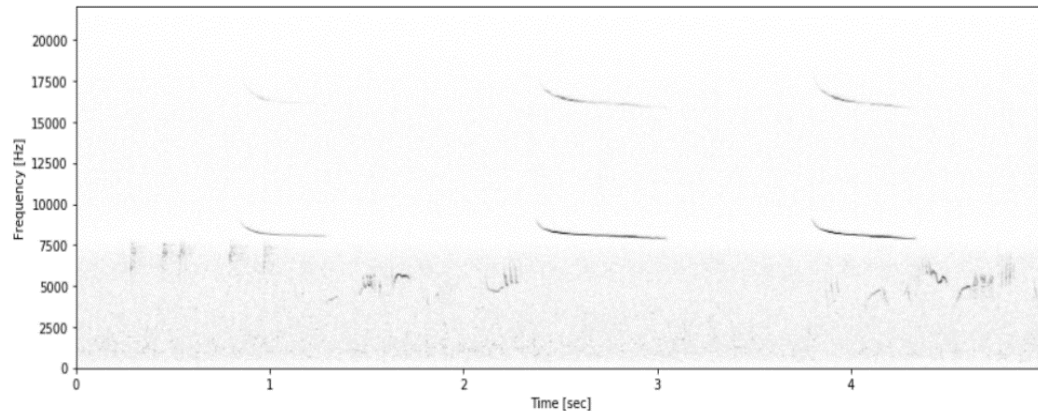
Methodology

- ขั้นตอนที่ 1: Preprocessing
- ขั้นตอนที่ 2: Segmentation
- ขั้นตอนที่ 3: Feature generation
- ขั้นตอนที่ 4: Classification

Methodology

Preprocessing

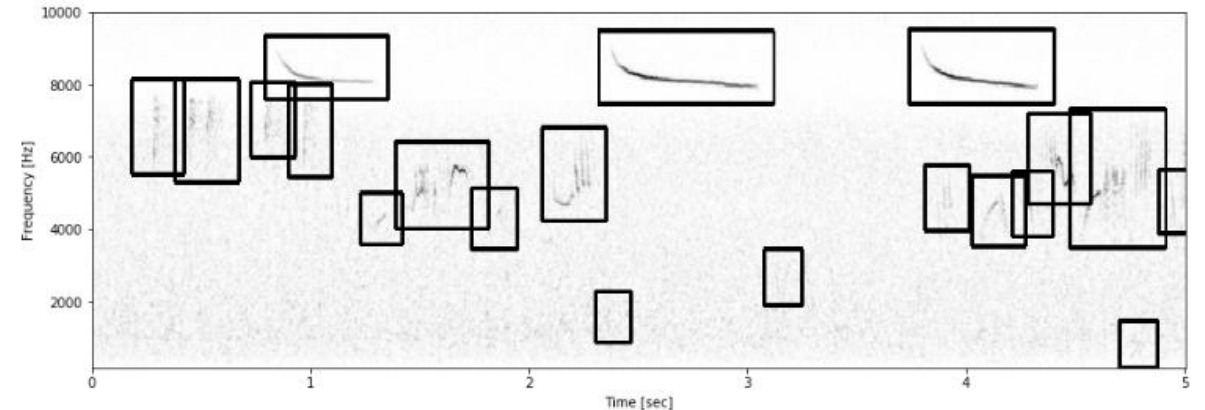
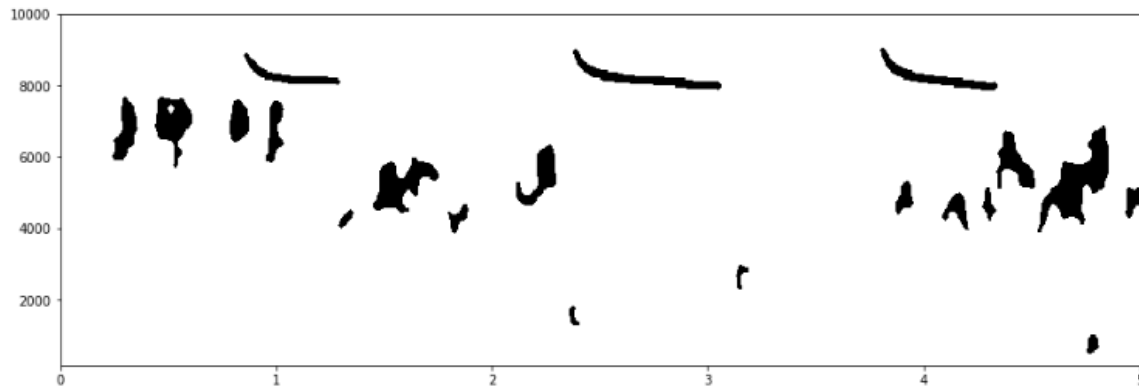
- แปลงเป็น spectrogram ด้วยเทคนิค STFT โดยใช้ฟังก์ชันหน้าต่างแบบ Hann window
- normalize ค่าใน spectrogram ให้อยู่ในช่วง 0 ถึง 1
- คัดเลือกช่วงความถี่ที่เหลือเพียงช่วง 170 – 10,000 Hz
- ผลลัพธ์ในขั้นตอนนี้คือ spectrogram ในช่วงความถี่ที่เหมาะสม ซึ่งเสมือนเป็นรูปภาพขาวดำ



Methodology

Segmentation

- ลบสัญญาณพื้นหลัง (background noise) ในภาพออกด้วยเทคนิค median clipping
- closing, dilation, median blur โดยใช้ OpenCV
- แยกส่วนประกอบโดยการหาเส้นขอบ (contour)
- ผลลัพธ์ในขั้นตอนนี้คือ รายการของภาพชิ้นส่วนสำคัญของภาพ พร้อมกับช่วงความถี่ของชิ้นส่วนนั้น



Methodology

Feature generation

- นำเสียงของนกมาทำขั้นตอนที่ 1 และ 2 เพื่อหาชิ้นส่วนสำคัญของเสียงของนกแยกตามชนิด
- หา feature โดยหาเปอร์เซ็นต์ความคล้ายสูงสุดของไฟล์เทียบกับแต่ละชิ้นส่วนของนกชนิดนั้น ๆ
- ใช้เทคนิค normalized cross-correlation
- เทียบเฉพาะช่วงความถี่ที่ใกล้เคียงกันเท่านั้น

Methodology

Classification

- แยกออกเป็นหลายโมเดลตามจำนวนชนิดของนก
- แต่ละโมเดลจะแทนนกชนิดหนึ่งและจะทำนายว่ามีเสียงนกชนิดนั้นอยู่ในไฟล์เสียงที่นำมาทดสอบหรือไม่
- ใช้โมเดล random forest
- ถ้าค่าคะแนนมากกว่าค่า threshold หนึ่ง ๆ เราจะตอบว่ามีเสียงนกชนิดนั้นอยู่ในไฟล์เสียงที่นำมาทดสอบ

Result

- ผลจากการทดสอบบนข้อมูลจากงานวิจัยที่เกี่ยวข้องบน Kaggle
- ผลจากการทดสอบบนข้อมูลที่ใช้ในโครงการ

Result

On Kaggle

- AUC score 90.17%
- คิดเป็นอันดับที่ 4 ในการแข่งขัน



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](#)

GoldSilverBronze

#	Δpub	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

Result

On Thai birds data

- ไฟล์เสียงนกที่ใช้รวบรวมจากเว็บไซต์ www.xeno-canto.org
- จำนวน 120 ไฟล์ ความยาวประมาณ 4 – 40 วินาที
- ประกอบด้วยเสียงนก 11 ชนิด ชนิดละ 10 ไฟล์ และเสียงนกอื่น ๆ อีก 10 ไฟล์
- แต่ละไฟล์อาจมีเสียงนกหลายชนิดปนกันได้

Result

On Thai birds data

- ประเมินผลด้วยค่า AUC score
- ใช้วิธี 5-Fold Cross Validation
- AUC score เฉลี่ยเท่ากับ 92.6%

Class no	Class Name	#features	AUC score
1	นกเขาใหญ่ (<i>Spilopelia chinensis</i>)	351	0.9864
2	นกเขาชวา (<i>Geopelia striata</i>)	423	0.9477
3	นกเอี้ยงสาธิต (<i>Acridotheres tristis</i>)	306	0.6227
4	นกกระแตแต้แว๊ด (<i>Vanellus indicus</i>)	574	0.9909
5	นกกระจอกบ้าน (<i>Passer montanus</i>)	334	0.9955
6	นกกาเหว่า (<i>Eudynamys scolopaceus</i>)	399	0.8833
7	นกกิ้งป๋อเล็กเหลือง (<i>Cinnyris jugularis</i>)	368	0.9773
8	นกตีทอง (<i>Psilopogon haemacephalus</i>)	458	0.9591
9	นกยางเปี่ย (<i>Egretta garzetta</i>)	277	0.9045
10	นกอีกา (<i>Corvus macrorhynchos</i>)	246	0.9363
11	นกอีวาบตักแตน (<i>Cacomantis merulinus</i>)	407	0.9818

Result

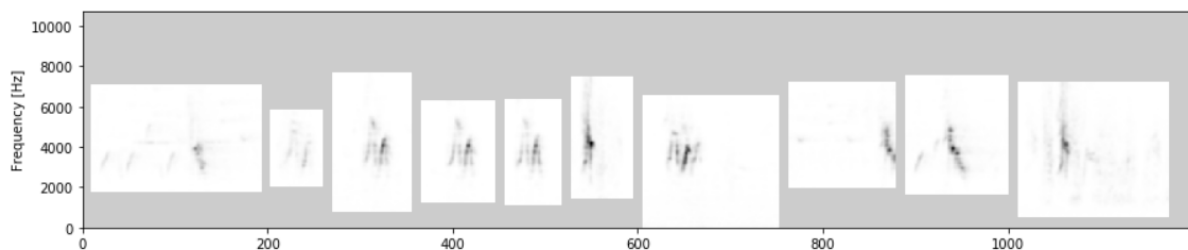
On Thai birds data

- โมเดลที่ได้จึงควรมีค่า recall ที่มาก
- เลือกค่า threshold เท่ากับ 0.11
- recall เท่ากับ 74.55%
- precision เท่ากับ 54.09%

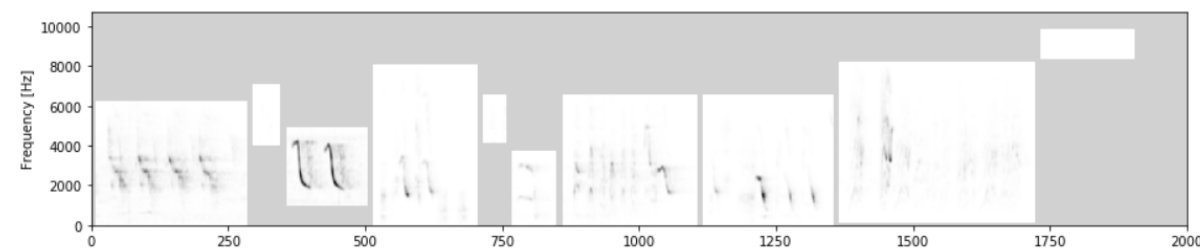
Class no	Class Name	recall	precision
1	นกเขาใหญ่ (<i>Spilopelia chinensis</i>)	1.0000	0.5933
2	นกเขาชวา (<i>Geopelia striata</i>)	0.7000	0.6833
3	นกเอี้ยงสาลิภา (<i>Acridotheres tristis</i>)	0.0000	0.0000
4	นกกระแตแต้แว๊ด (<i>Vanellus indicus</i>)	0.9333	0.7000
5	นกกระจอกบ้าน (<i>Passer montanus</i>)	0.9000	0.7333
6	นกกาเหว่า (<i>Eudynamys scolopaceus</i>)	0.5333	0.5167
7	นกกิ้งปล็อกเหลือง (<i>Cinnyris jugularis</i>)	1.0000	0.6467
8	นกตีทอง (<i>Psilopogon haemacephalus</i>)	0.8000	0.3567
9	นกยางเปี่ย (<i>Egretta garzetta</i>)	0.5000	0.6000
10	นกอีกา (<i>Corvus macrorhynchos</i>)	0.9000	0.4533
11	นกอีวาบตักแตน (<i>Cacomantis merulinus</i>)	0.9333	0.6667

Result

On Thai birds data



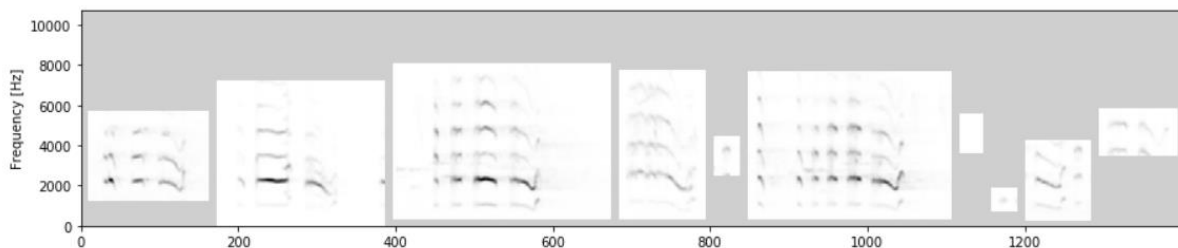
นกกระจอกบ้าน (AUC score 0.9955)



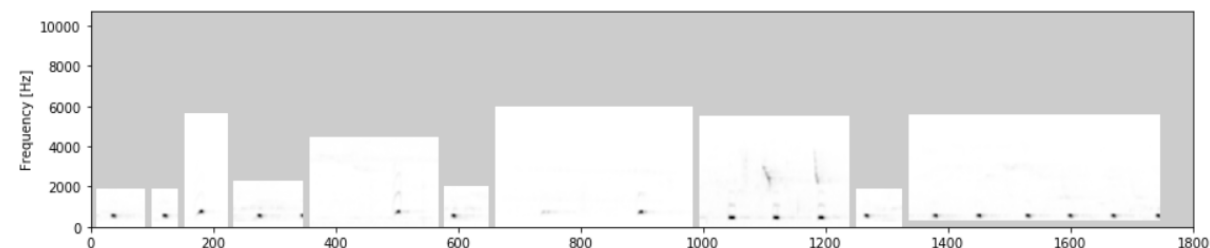
นกเอี้ยงสาลิกา (AUC score 0.6227)

Result

On Thai birds data



นกกระแตแต้แว้ด (AUC score 0.9909)



นกตีทอง (AUC score 0.9591)

Conclusion & Next step

- เวลาที่ใช้ในการจำแนกเสียงนกจะแปรผันตามความยาวของเสียงที่นำมาทดสอบ
- ปรับโมเดลในการประมวลผล โดยอาจใช้เทคนิค deep learning เข้ามาช่วย
- ขยายขอบเขตของโครงการให้รองรับเสียงนกมากกว่า 40 ชนิดขึ้นไป
- รับข้อมูลเสียงและประมวลผลแบบ real-time
- มีการแสดงผลลัพธ์เป็นชื่อนก พร้อมรูปภาพ
- มีระบบเก็บข้อมูลชนิดนกที่ได้ยินในแต่ละเวลา

Q&A