

日本国内における夜間フライトコール (NFC) の自動検出

加藤 義清 (バードリサーチ会員)

大坂英樹 (トリルラボ)

鳥類学大会2024

鳥の鳴き声の録音調査の課題

大量の音声データを確認し鳴き声を同定する作業に手間と時間を要する

機械学習を応用した鳴き声の自動検出

➡音声データの確認作業に要する手間を大幅に削減可能

➡広域大規模な調査が現実的なものに

- 事前学習モデルを利用することにより、タスク固有データを大量に用意しなくても比較的高性能な分類器の構築を目指す



発表の目的

日本国内で2024年春に実施された録音調査のデータに夜間フライトコール (NFC) の自動検出手法を適用し、その性能を検証した。その結果に基づき、今後の調査方法の実現性を検討する。

渡り鳥の夜間フライトコール

ホオジロ属 (*Emberiza*) やツグミ属 (*Turdus*) など、渡りの途中、夜間にフライトコール (Nocturnal Flight Call (NFC)) と呼ばれる声を発しながら飛ぶことが知られている

渡りの調査方法として、欧米を中心にNFCの録音調査が他の手法を補完する形で採用されている

日本におけるNFC録音調査の試み

市民科学プロジェクト*として、春と秋の渡りのシーズンに日本各地の協力者から録音データの提供を受け、録音調査から渡りの状況を明らかにすることを目指している

2024年度は基礎検討フェーズで自動検出技術の実データを用いた性能検証を行う

* 2023年度バードリサーチ調査研究支援プロジェクト「皆で地鳴きから夜の小鳥の渡りを調べる」

2024年の夜間フライトコール調査

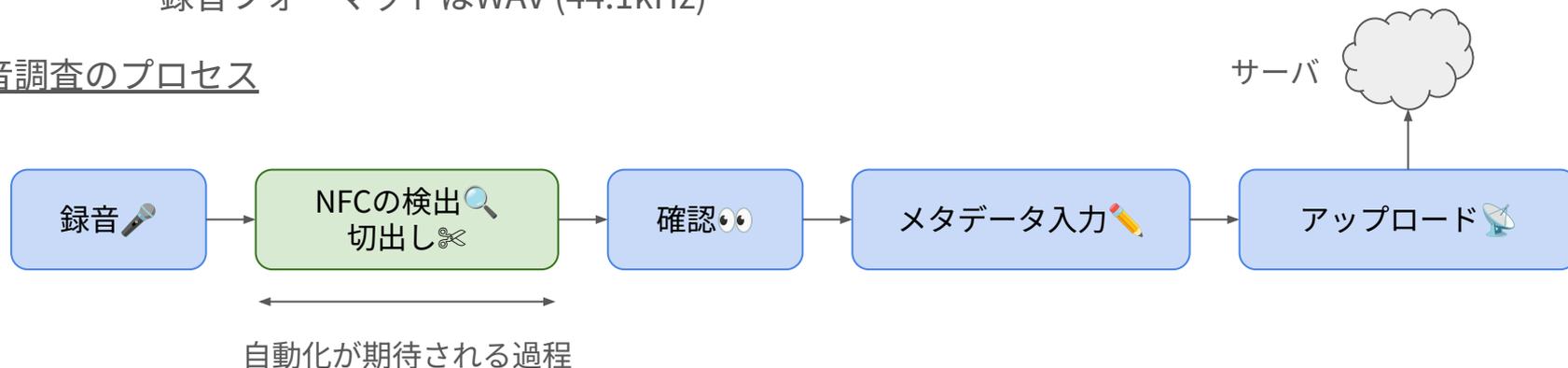
調査期間 春季2024年4月1日～6月30日、秋季2024年9月1日～11月30日

実施方法 全国の調査協力者(18名)が夜間にICレコーダで録音し、音声データを提供

サイト数 北海道から沖縄まで15ヶ所

- 録音方法
- ・庭やベランダなど身近な場所にICレコーダを設置し、タイマー録音
 - ・録音頻度は週1回から毎日まで、1晩に3回 (21:00-22:00, 0:00-1:00, 3:00-4:00)
 - ・録音フォーマットはWAV (44.1kHz)

録音調査のプロセス



NFCの自動検出

音声埋め込みに基づく分類器 [1] を利用

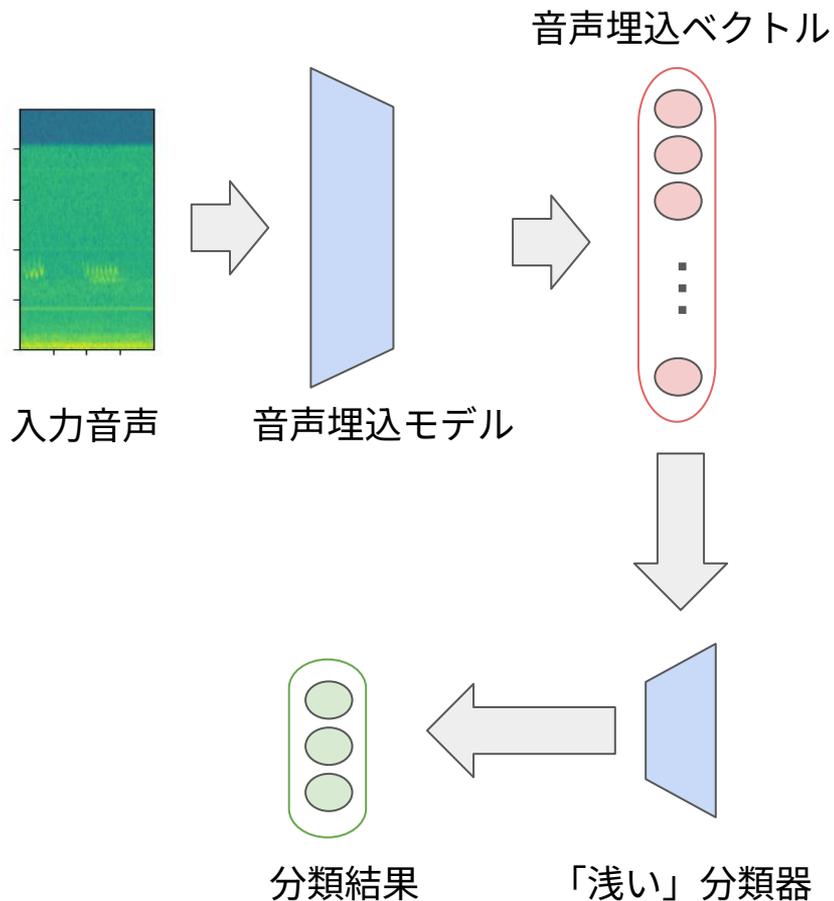
特徴

- 入力音声を事前学習モデルで音声埋め込みに変換
- 音声埋め込みを入力とした「浅い」分類器を構築

利点

- 大量のデータで訓練された事前学習モデルは入力音声の重要な特徴を埋め込みに表現可能
- 少量のデータでも高い性能の分類器が構築可能

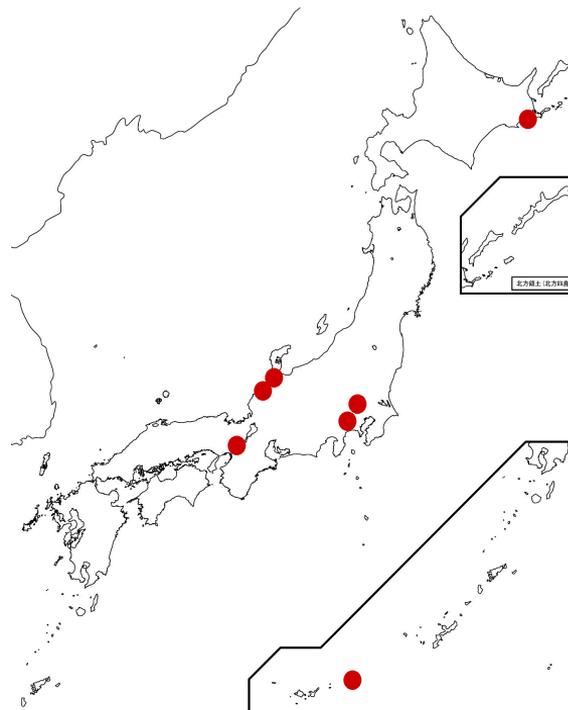
[1] 大坂, 加藤: [渡り鳥の夜のフライトコール録音調査に向けた自動判別技術の有効性](#). [人工知能学会第64回AIチャレンジ研究会](#), ポスター発表 (2024)



2024年春のNFC録音調査データへの適用

概要

調査期間	2024年4月～6月
サイト数	7
のべ録音日数	141日
総録音時間	522時間



NFC自動検出の適用

分類器構成 音声埋込モデル: Perch[2]、分類手法: SVM

機材 CPU: Core i5-6500 (3.2GHz), GPU: NVIDIA GeForce GTX 1070

処理時間 24時間分の音声データにつき1時間
=>全データの処理に約22時間

性能評価用データの作成方法と結果

サイト毎に、分類器スコア*に基づく層化抽出**を実施

スコアの分布（右下図）に基づき、層化区間を[0, 0.2], (0.2, 0.4], (0.4, 0.6], (0.6, 0.8], (0.8, 1]として、各サイト、各区間から最大10標本を抽出し、273標本を得た

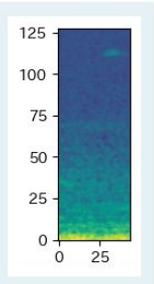
* 分類対象（音声）が特定のクラス（NFC）である確率や確信度を表す値

** 統計学における標本調査の手法。部分母集団（層）に偏りがある時に、層ごとに標本抽出をすることで推定誤差を小さくすることができる

抽出標本に対し、スペクトログラムの目視および音声を確認してラベルを付与

アノテーションインターフェース

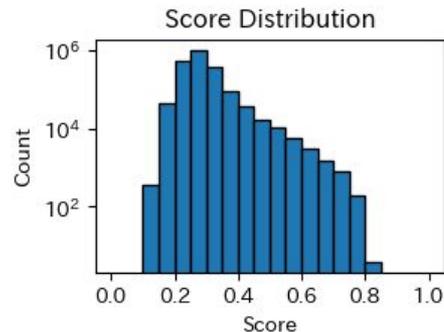
Filename	/data/nfc/survey/2024年録音生データ/ (石川県加賀市) /240415_000002_032250_石川県加賀市_240414_0015.wav
Position	9275.0
Score	0.610822



0:01 / 0:01

Item 5:

- none
- background
- maybe
- nfc**



性能指標 (参考)

再現率 (Recall)

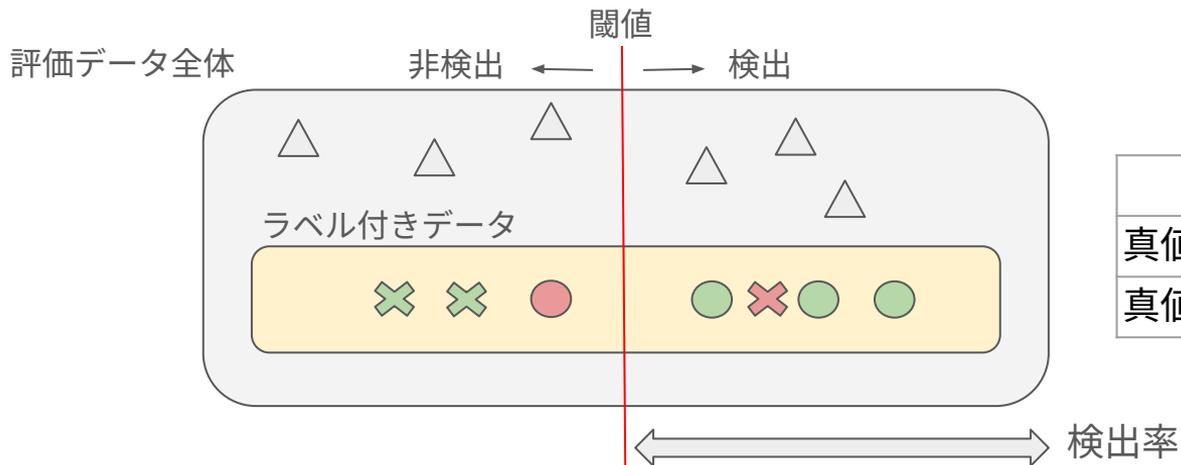
評価データ中の正例のうち、正しく検出された割合
(高いほど正例の見逃しが少ない)

適合率 (Precision)

検出された事例のうち、正しく正例である割合
(高いほど誤検出が少ない)

検出率 (Detection Rate)

データ全体のうち、「検出」とされる事例の割合
(低いほど確認すべき事例が少なくて済む)



	予測: 検出	予測: 不検出
真値: 正例	● 真検出	● 誤不検出
真値: 負例	✗ 誤検出	✗ 真不検出

性能評価の結果（全体）

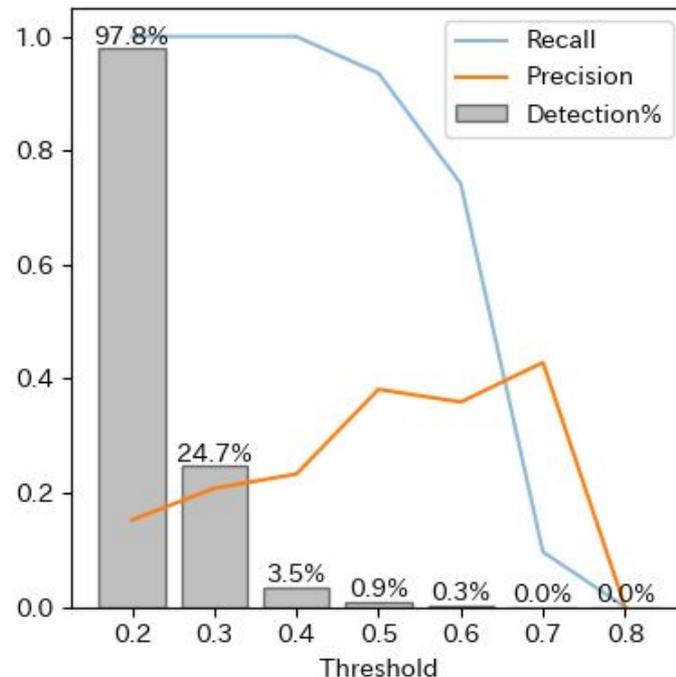
再現率 閾値が0.5を超えると急激に低下する

適合率 閾値が高いほど高くなる傾向にあるが、0.8の時に急激に悪化する
これは特定のサイトの影響によるもの（後述）

検出率 閾値とともに急激に減少する

閾値が0.5を採用すれば、再現率 0.935、適合率 0.382、
検出率 0.9%となる

→確認すべきデータを大幅に削減しつつ、検出漏れを最小限にできることが期待できる



性能評価の結果（サイト単位）

再現率 閾値0.5まで1近くを維持する傾向がみられた

検出率 閾値0.3以上で急激に低下する傾向がみられた

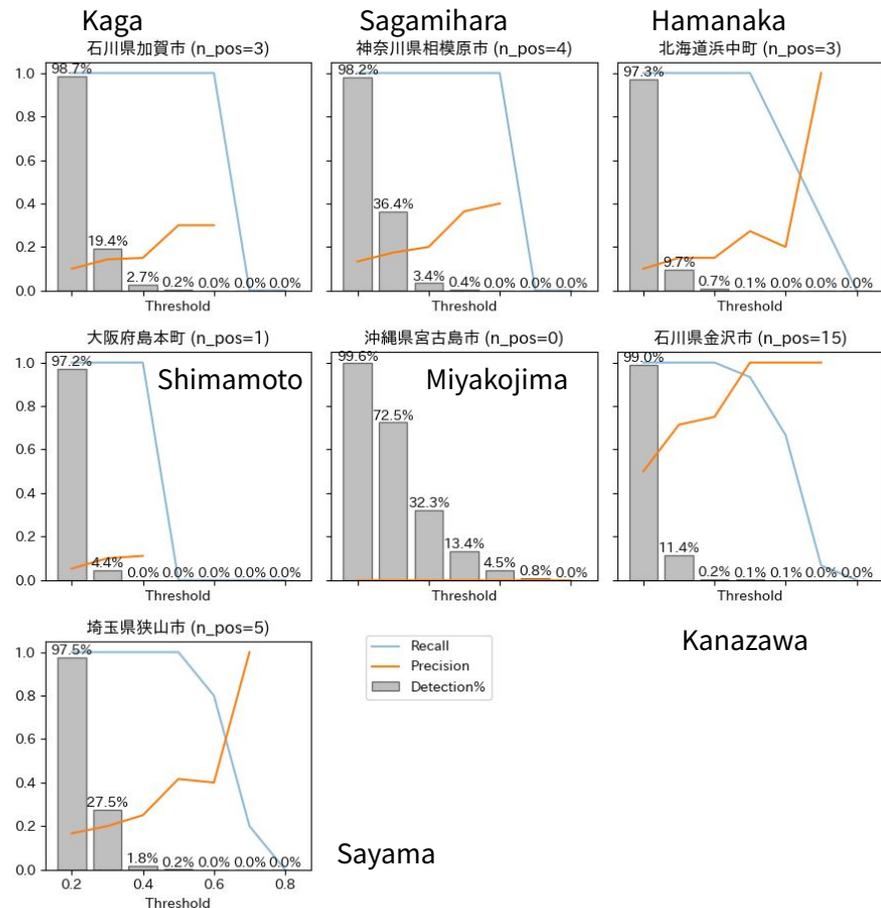
適合率 多くのサイトで全体評価と同様の傾向を示す
一方、例外的に性能が良いサイトもみられた
金沢@Th=0.5: P=1.0, R=0.93, DR=0.1%

1サイト（宮古島）は抽出事例に正例を見つけることができなかった

→全体評価における適合率低下に寄与

宮古島のスコア分布は他サイトと傾向が異なる

（正例がなかったこととの関係は不明だが、両生類などの鳴き声があり、生物相が異なることによる可能性がある）



誤検出エラーの評価

誤検出例を抽出(n=10)して、エラーの原因を検証

カテゴリの1、2、4については、今後学習データを拡張することにより改善することが期待される

3に関しては、アノテーション品質の問題

→対策として、アノテーターを増やしたり、アノテーターに対する訓練の実施などが考えられる

No.	カテゴリ	事例数
1	NFCではない鳥の鳴き声	3
2	昆虫の声	3
3	判断がつかず	3
4	自動車	1

NFC自動検出結果（1）

日別NFC検出数 (大阪、神奈川、埼玉)

ピンク色背景は録音日を示す

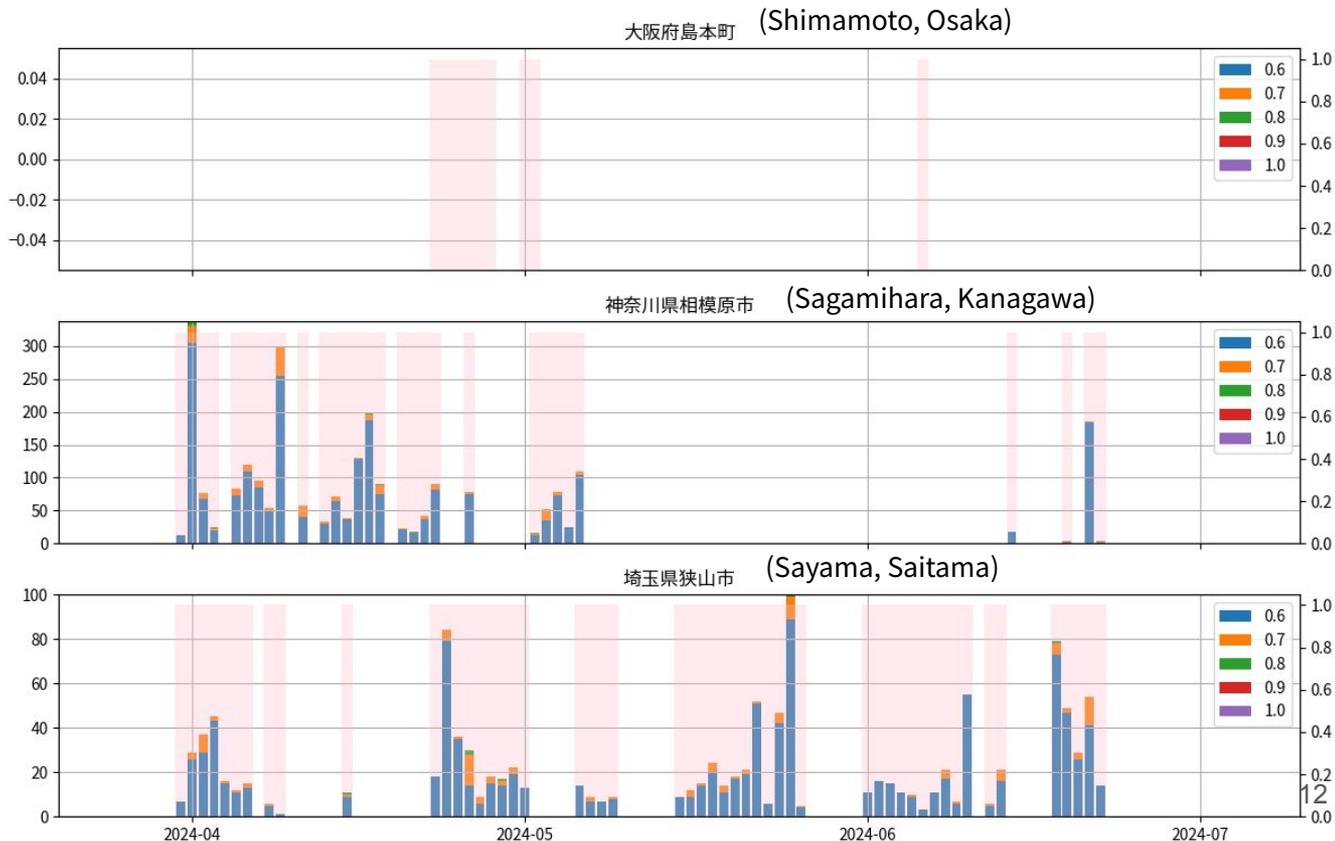
棒グラフの色は分類器スコアの区間に対応

例: “0.6” -> スコアが区間(0.5, 0.6]の事例に対応

縦軸は事例数

備考

対象とした7サイトのうち、
宮古島は抽出事例中に正例が
認められなかったため、結果
から除外した



NFC自動検出結果（2）

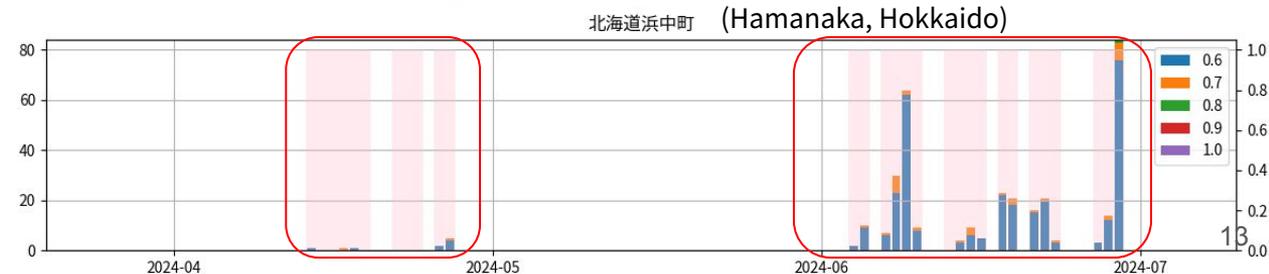
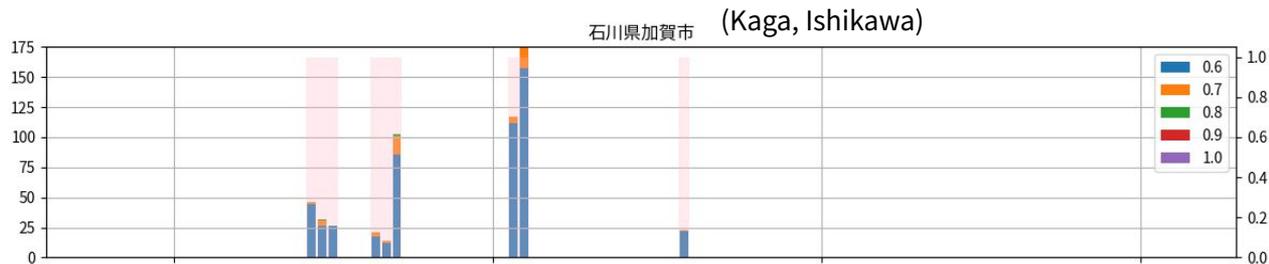
日別NFC検出数 (石川、北海道)

結果

関東地方や北陸地方では4月から6月にかけてNFCが検出されている

浜中町（道東）では4月にNFCが少数検出されているが、6月の録音でははるかに多く検出される
→渡りの結果、4月では少なかった個体数が6月には増加したか

大阪府島本町のデータは標本調査で正例は認められたものの、全体的にスコアが低い傾向にあり、自動検出数は0となった



まとめ

NFC自動検出手法の実調査データを用いた性能評価を実施

→適切な閾値 (0.5) を選べば、高い再現率を保ちつつ、比較的低い検出率を達成できることを示した

→録音調査の労力削減が期待される (1時間の録音につき30秒相当の事例を確認すれば良い)

実調査データに適用した結果、録音調査と自動検出を組み合わせることで、各地のNFC頻度の季節変動を捉えられることが示唆された

録音状態や環境がサイト毎に異なり、検出率に影響を与えるため、NFC検出数の絶対値はサイト間で直接比較できない

(録音方法の検討は大坂らが別途報告)

今後の課題

性能の向上 録音調査データから学習データを増やすことを検討

種の判別 検出されたNFCに更に種のラベルを付与する種判別器の構築を検討

録音サイト間の性能のばらつき

→ドメイン適用によりサイト毎に性能を改善できないか検証する

(例えば、南西諸島など現地生物相に合わせた学習など)

謝辞

「皆で地鳴きから夜の小鳥の渡りを調べる」プロジェクトのデータを利用させていただきました。録音データを提供してくださった調査協力者の皆様に感謝いたします。

本研究を進めるにあたってアドバイスを頂いたTom Denton氏に感謝いたします。