

MD4 AI 技術を用いた生物音響モニタリング: 大規模鳥類識別モデルの日本の複数の生物分類群への応用

Bioacoustics Monitoring Using AI Technology:

Application of a Large-Scale Acoustic Bird Classifier to Multiple Taxonomic Groups in Japan

地球循環共生工学領域 28H22029 佐藤匠 (Takumi SATO)

Abstract: Deep learning models have been actively developed to analyze long-term recording data to assess biodiversity. In birds, BirdNET, which can identify more than 6,000 bird species, has been developed. This study evaluated BirdNET 1) for its generalization performance on Japanese bird call data, and 2) for its applicability to bioacoustic monitoring of multiple taxonomic groups. In case 1), BirdNET was evaluated for four tasks: a) identifying individual calls, b) identifying species composition, c) detecting Red List species, d) identifying seasonal variation. BirdNET showed performance comparable to human surveys, but the predicted seasonal variation showed low correlation coefficient with human survey. Further Japanese bird data collection is needed to improve performance. In case 2), transfer learning successfully achieved high F1 scores to identify birds, amphibians, insects, and bats. This result suggests the applicability of BirdNET to taxa other than birds, even with a small number of training data.

Keywords: BirdNET, Transfer learning, Deep learning, Automatic classification, Passive acoustic monitoring

1. 背景と目的

生物多様性を長期録音データで評価するための深層学習モデルが盛んに開発されている。指標生物の一つである鳥類の分野では、六千種を超える種の判別が可能な BirdNET が開発された¹⁾。しかし、日本の鳥類に対して BirdNET がもつ識別性能が評価されていないことや、鳥類以外の生物分類群では十分なデータ数を確保するのが難しくモデルの構築が困難であることが課題である。そこで本研究では、大規模鳥類識別モデルである BirdNET について、1) 日本の鳥類の鳴き声に対して汎化性能を持つかを評価するとともに、2) 転移学習で鳥類、両生類、昆虫類、コウモリ類などの複数の生物分類群の識別へ応用可能であるかを検証した。

2. 研究手法

2. 1 日本の鳥類に対する BirdNET の識別性能の評価

まず(1)日本全国²⁾、(2)千葉県全域³⁾、(3)奈良県春日山⁴⁾で録音された3種類の正解ラベル付きの鳥類の鳴き声の音声データセットを構築した(表 1)。この音声データセットを用いて、日本の鳥類に対して BirdNET v2.4 が持つ汎化性能を a) 鳥類の鳴き声の音響イベントをどれだけ正確に識別できるか、b) 鳥類の種組成を推論できるか、c) レッドリストに含まれる種を検出できるか、d) 長期録音データから種別の活動量の変化を推論できるかの4つのタスクで評価した。タスク a) では Overall Accuracy, Precision, Recall, F1 score, b) c) では Precision, Recall, F1 score, d) では長期録音された時系列データの正解ラベルと識別結果の相関を評価した。

2. 2 転移学習による複数の生物分類群の識別

鳥類以外の生物分類群について(4)16種の両生類・昆虫類³⁾、(5)29種のコウモリ類⁵⁾、(6)111種の両生類・昆虫類・コウモリ類・鳥類³⁾⁵⁾の鳴き声を収録した正解ラベル付きの3つの音声データセットを構築した(表 2)。音声データセット(1)と(4)–(6)に対して、音声ファイルごとにグループ化した層化5分割交差検証を行い、BirdNET の転移学習の評価を行った。転移学習では、鳴き声の特徴の埋め込み層の重みを固定して全結合層を訓練した。モデルの交差検証誤差は Precision, Recall, F1 score で評価した。

表 1 日本の鳥類に対する識別性能の評価に使用したデータセットの諸元

No.	録音地点	音声ファイル数	平均録音時間 (s)	種数	種別の音声ファイル数		
					平均	最小	最大
1	日本全国	1111	18 ± 8	419	3 ± 2	1	17
2	千葉県	8838	2 ± 1	48	180 ± 152	50	500
3	奈良県春日山	768	2084 ± 324	63	457 ± 835	1	3427

表 2 BirdNET の転移学習に用いた鳥類以外の生物分類群のデータセットの諸元

No.	含まれる生物分類群	音声ファイル数	平均録音時間 (s)	種数	種別の音声ファイル数		
					平均	最小	最大
4	両生類・昆虫	764	4 ± 2	16	45 ± 42	10	137
5	コウモリ	1414	44 ± 20	29	47 ± 61	3	292
6	両生類・昆虫・コウモリ・鳥類	11897	7 ± 15	111	106 ± 154	3	1080

表 3 日本の鳥類に対する識別性能と
転移学習による複数の生物分類群の識別結果

タスクの種類	Precision	Recall	F1 score
BirdNETの汎化性能の評価			
個々の鳴き声の判別	0.847	0.498	0.580
BirdNETの転移学習			
鳥類	0.913 ± 0.012	0.880 ± 0.021	0.889 ± 0.018
両生類・昆虫	0.977 ± 0.029	0.970 ± 0.028	0.972 ± 0.029
コウモリ	0.775 ± 0.005	0.665 ± 0.027	0.693 ± 0.020
両生類・昆虫・コウモリ・鳥類	0.884 ± 0.008	0.833 ± 0.005	0.848 ± 0.006
地域に生息する鳥類の種の検出			
鳥類の種組成の把握			
データセット(2)	0.331	0.938	0.489
データセット(3)	0.471	0.905	0.620
レッドリスト種の検出			
データセット(2)	0.311	1.00	0.474
データセット(3)	0.500	0.897	0.642

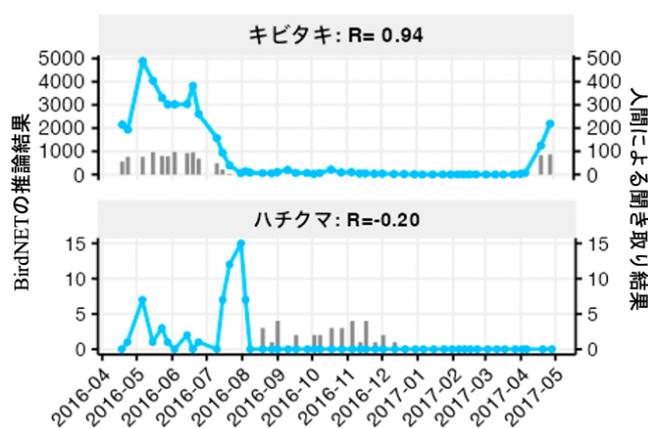


図 1 相関係数が高かったキビタキ(上)と低かったハチクマ(下)の BirdNET の予測(青, 折れ線)と人間による聞き取り結果(黒, 棒グラフ)の時系列変化

3. 結果と考察

3. 1 日本の鳥類に対する BirdNET の識別性能の評価

a) 個々の鳴き声の識別, b) 鳥類の種組成の推論, c) レッドリストに含まれる種の検出のタスクに対する評価結果を表 3 に示す. a) 個々の鳴き声の識別は Overall accuracy が 54.1% であり, 北米や欧州の先行研究と同程度の識別性能を示した. b) 鳥類の種組成の推論, c) レッドリストに含まれる種の検出では, 人間の聞き取り調査による精度と同等であったが, BirdNET のパラメータの設定によっては, 誤検出する種や見逃す種の数が増加した. よって調査目的に応じてハイパーパラメータを校正することが重要であると示唆された. d) 鳥類の種別の活動量の変化の推論では, BirdNET の推論と人間による聞き取り調査の結果の相関係数が 0.95 の高い相関を示す種もあったが, 全体の半数近くの種では相関が低く変動の推論には課題を抱えていた (図 1).

3. 2 BirdNET の転移学習による複数の生物分類群の識別結果

BirdNET を複数の生物分類群に転移学習した結果を表 3 に示す. BirdNET を日本国内の鳥類の鳴き声のデータセット (1) で転移学習した結果, F1 score は 0.580 から 0.889 に改善された. また, 鳥類以外の生物音響データベース (4), (5), (6) に対して転移学習した結果, F1 score のマクロ平均はそれぞれ 0.981, 0.692, 0.848 であった. (4) 両生類・昆虫類では各種の学習データ数が 12 - 241 個と少数にも関わらず, 16 種中の 12 種で F1 score が 0.980 以上と高い精度を達成した. (5) コウモリ類では, 29 種中の 13 種で F1 score が 0.800 以上を達成したが, 8 種では 0.600 を下回った. さらに, データセット (6) で両生類・昆虫類・コウモリ類・鳥類を同時に転移学習したところ, 111 種中の 62 種で F1 score が 0.900 以上を達成し, 0.600 を下回ったのは 11 種のみであった. これらの結果から, BirdNET の転移学習はデータ数が限られている鳥類以外の生物分類群の生物音響モニタリングにも有効であり, 生物多様性研究での効率的かつ多様な評価に貢献できる可能性が示された.

4 今後の課題

日本国内に生息する生物音響モニタリングに向けては, 大規模識別モデルの汎化性能が低い種の精査や学習していない種の音声データを蓄積することが重要である. また音響識別技術を応用した生態系モニタリング技術の開発や地理空間情報や視覚情報を用いたマルチモーダルな識別モデルへの拡張が重要である.

参考文献

- 1) Kahl, S (online) BirdNET-Analyzer [Python]. <https://github.com/kahst/BirdNET-Analyzer> 参照 2023 年 11 月 13 日
- 2) Bird research (online) バードリサーチ 鳴き声図鑑 https://www.bird-research.jp/1_shiryo/index.html, 参照 2023-11-13
- 3) 千葉県立中央博物館 (online) どうぐ-tools: 耳をたよりにプロジェクト. <https://www.chiba-muse.or.jp/NATURAL/special/hearing/html/tools.html> 参照 2023-11-13.
- 4) 大矢樹, 岡本真帆, 武田紗季, 伊東明, 名波哲 (2017) 春日山原始林において観察記録が途絶えた鳥類の音声による再確認. 地域自然史と保全 = Bulletin of Kansai Organization for Nature Conservation / 関西自然保護機構 編 39: 61-70.
- 5) Kobayashi, K. et al. (2021) Development of a species identification system of Japanese bats from echolocation calls using convolutional neural networks. Ecological Informatics 62, 101253. <https://doi.org/10.1016/j.ecoinf.2021.101253>