

Deep learning-based detector of invasive alien frogs, *Polypedates leucomystax* and *Rhinella marina*, on an island at invasion front

Kaede Kimura^{1*}, Ibuki Fukuyama^{2,3}, Kinji Fukuyama⁴

¹ Graduate School of Science, Kyoto University, Kyoto, Japan. ORCID ID 0009-0001-4009-1424

² Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan. ORCID ID 0000-0001-6887-9582

³ Wakayama Experimental Forest, Field Science Center for Northern Biosphere, Hokkaido University, Wakayama, Japan.

⁴ Department of Biology, Keio University, Yokohama, Kanagawa, Japan

* Corresponding author: kaede.kmr@gmail.com

Abstract

Early detection is crucial for managing invasive alien species at invasion fronts. Automated bioacoustic monitoring using deep learning can support this by detecting species calls. However, its application for invasive alien species remains limited. Additionally, the lack of local training data in pre-establishment areas, where invasive species are absent or rare, can limit model performance due to domain shifts. This study aimed to develop and evaluate a deep learning-based detector for Southeast Asian treefrog (*Polypedates leucomystax*) and cane toad (*Rhinella marina*) on Iriomote Island, a Natural World Heritage Site located 30 km from established populations of these alien species on the nearby Ishigaki Island. We trained the BirdNET model with acoustic data collected on Ishigaki Island, where these species were common, as well as native frog calls from Iriomote Island. Model performance was evaluated using (1) wild sounds on Ishigaki Island, and (2) sounds obtained by playing back the alien species call on Iriomote Island. Model precision and recall for wild sounds were 0.972 and 0.906 for *P. leucomystax*, and 0.899 and 0.705 for *R. marina*, respectively. Playback survey dates were identifiable from the high number of detections when the detection thresholds were properly adjusted, except on a night with dense choruses of a native frog species, *Microhyla kuramotoi*. These results suggest that data from Ishigaki Island enabled training a model with adequate generalization across invasion front.

Keywords

Passive Acoustic Monitoring (PAM), preparedness, introduced species, call playback

Introduction

Biological invasion is a major threat to biodiversity worldwide, and eradication of introduced species is most feasible during the early phase of invasion (IPBES 2023). Early detection of invasive alien species has been facilitated by novel monitoring technologies like remote sensing, environmental DNA (eDNA), and machine learning (Fricke and Olden 2023). Deep learning, a form of machine learning, is gaining attention for automated biological monitoring because of its high accuracy and scalability (Norouzzadeh et al. 2018; Kahl et al. 2021). Recently, several studies attempted to detect invasive alien species in the wild with deep learning. Examples include detecting *Anolis* lizards (Aota et al. 2021), lionfish (Martínez-González et al. 2021), and Asian black hornets (O'Shea-Wheller et al. 2024) from images. Although its applications in acoustic monitoring are more limited, studies have explored the detection of barred owls (Kelly et al. 2023) and bullfrogs (Bota et al. 2024) from sounds. These studies suggest that a deep learning model can detect invasive alien species when local training samples (i.e., samples collected at sites of actual application) are available.

However, it is impossible to acquire local training data when the purpose of a project is preparedness at an invasion front, and the deep learning-based detectors trained on *ex situ* data can have lower performance due to domain shift (Beery et al. 2018). Domain shift refers to the situation where training and target data have different characteristics, often reducing model accuracy. For example, detection models trained using visual data from camera traps recording at one location can perform poorly if applied to data recorded at a different location due to differences in background and lighting conditions (Beery et al. 2018; Schneider et al. 2020; Norman et al. 2023). At the front of biological invasion, early detection systems are deployed in locations where training data of the target alien species are often unavailable (e.g., Taylor et al. 2017). Though this dilemma is likely common, the expected accuracy reduction of a model trained on *ex situ* data and the significance of such accuracy reduction in practical settings have rarely been examined.

Iriomote Island, a tropical island designated as a Natural World Heritage Site in Japan, has been threatened by the invasion of two alien frog species, cane toad (*Rhinella marina*) and Southeast Asian treefrog (*Polypedates leucomystax*) (National Institute for Environmental Studies 2024). *Rhinella marina* is not only a voracious predator but also a toxic prey, and especially the latter causes significant negative impacts on local biodiversity (Shine 2010). The ecological impacts of *P. leucomystax* have not been studied well, but possible negative effects on native anuran through resource competitions, such as *Zhangixalus* spp. which appear to utilize similar habitats during both larval and adult stages, are concerning (Ota et al. 2004; National Institute for Environmental Studies 2024). Additionally, *P. leucomystax* is suggested to be involved in the transfer of helminth parasites, potentially spreading them to native species (Hasegawa and Ota 2017). These two species have established dense populations on the nearby Ishigaki Island (about 30 km apart) and have been unintentionally transported to Iriomote Island dozens of times. Continuous detection and eradication efforts have prevented these species from establishing and spreading on Iriomote Island to date (Nakajima et al. 2005), but constraints on resources and manpower limit surveillance

frequency and coverage. Therefore, developing a system that can automatically detect the invasive alien frogs on Iriomote Island are needed. Given that male frogs produce species-specific calls during their breeding seasons, passive acoustic monitoring (PAM) has commonly been used to monitor their population status (Dorcas et al. 2009). Analyzing PAM data with deep learning is a promising approach to enable efficient and continuous monitoring of invasive frog species (Bota et al. 2024), potentially supporting early detection and rapid response efforts.

Our objectives in this study were to develop and test an early detection system for *R. marina* and *P. leucomystax* on Iriomote Island, where local training data are unavailable. First, we trained a deep learning model to identify the two alien and six native frog species from their calls. During this process, the training data of the alien species was obtained from the nearby Ishigaki Island. We expected minimal domain shift in biogenic sounds due to their faunal similarity (e.g., all amphibian species are shared; Matsui and Maeda 2018). On the other hand, we expected more intense anthropogenic sounds such as traffic noise on Ishigaki Island, which is 20 times more populated than Iriomote Island (Okinawa Prefectural Office 2024). Next, we evaluated the model performance in detecting alien frogs on both Ishigaki and Iriomote Islands. On Iriomote Island, where these alien species had no established population, we tested the model using field playback of their calls. Third, we assessed the similarity between playback sounds and wild calls by mapping acoustic features (i.e., model embeddings). Finally, to assess the seasons for effective acoustic monitoring, we investigated the breeding phenology of *P. leucomystax* and *R. marina* by applying the trained model to acoustic data collected over one-year on Ishigaki Island. In doing so, we aimed to obtain a deep learning model that generalizes well across the invasion front for real-world deployment.

Materials and methods

Acoustic data collection

Acoustic data were collected from three sites on Ishigaki Island and three sites on Iriomote Island (Fig. 1). We selected habitats including roadside ponds (sites 1, 3, 4), mixture of sugar cane field and rice paddy (site 2), rice paddy (site 5), and roadside ditch (site 6), because *P. leucomystax* and *R. marina* are commonly found in disturbed environments on these islands (Matsui and Maeda 2018). We placed acoustic recorders (RR-XS455, Panasonic) at sites 1–4 from March 23 or September 23, 2023, to March 8, 2024. SD cards and batteries were replaced six months after deployment. For sites 5 and 6, we used a sound dataset initially obtained for long-term frog monitoring on Iriomote Island (K. Fukuyama, unpublished). Recordings were scheduled with a built-in timer for 15 minutes starting at 21:00 every night in MP3 format at 128 kbps.

To evaluate model performance on Iriomote Island, where the target alien frogs had no established populations, we obtained the field recordings of the alien frogs by playing back their calls. We asked a park ranger to conduct a playback survey monthly at site 4, sometime between 21:00 to 21:15. The playback sound of *P. leucomystax* was collected on Okinawajima Island and that of *R. marina* was collected on

Ishigaki Island (available on the website of Ministry of the Environment, Japan: https://kyushu.env.go.jp/okinawa/wildlife/mat/m_2_2.html). The playback duration varied day by day but mostly around 60 seconds (range 30–80) for each species.

Training and test datasets

We annotated wild and playback sounds using Deep Audio Segmenter software (Steinfath et al. 2021). For wild sounds, we selected 30 to 38 recordings per sites, totaling 190 minutes, to represent various seasons, weather conditions, and vocalizing species. The first 60 seconds of these recordings were inspected through sounds and spectrograms, and we identified the beginning and end of species-specific calls following Cañas et al. (2023). Faint calls hardly visible on spectrograms were not annotated, and calls with inter-call intervals less than 0.5 seconds were combined. The annotated recordings were segmented into 3-sec audio clips to fit the model analysis unit.

We created a wild sound test dataset by randomly selecting 20% of the annotated recordings, with the rest divided into training (80%) and validation (20%) subsets (Christin et al. 2021) (Table 1). The training dataset was used to train the model, the validation dataset was for model hyperparameter tuning, and the test dataset was to evaluate model performance. To prevent overly optimistic model evaluation due to consecutive (and hence non-independent) calls appearing in the both training and test datasets, we used the entire annotated recordings (60 seconds), instead of shorter 3-second clips, as the unit of splitting (Desprez et al. 2023. arXiv:2305.10472).

To create the playback sound test dataset, we extracted all playback recordings (totaling 27.8 minutes) and an approximately equal amount (28 minutes) of recordings without playback calls from site 4, which were randomly selected from 28 different days. These audio data were annotated following the same procedure as the wild sounds and used as a test dataset (Table 1). Note that no playback sounds were used for training to keep this test dataset as independent as possible from the training dataset.

We identified six native (*Buergeria choui*, *Fejervarya sakishimensis*, *Kurixalus eiffingeri*, *Microhyla kuramotoi*, *Nidirana okinavana*, *Zhangixalus owstoni*) and two alien frog species (*P. leucomystax*, *R. marina*) during the annotation procedure. For *P. leucomystax*, at least three types of calls were observed, but only the most common and conspicuous call type was annotated.

Model training and evaluation

We used BirdNET V2.4 for custom classification model training (Kahl et al. 2021). BirdNET V2.4 is an EfficientNetB0-based classifier optimized for animal sounds (Ghani et al. 2023). While the original BirdNET is trained chiefly on bird songs, it can be applied to other taxa by fine-tuning. The hyperparameters of BirdNET include model architecture settings (the number of hidden units and dropout rate) and training procedure options (e.g., learning rate, upsampling ratio of minority classes, and mixup data augmentation). These parameters can be automatically optimized through Bayesian hyperparameter tuning. We set the number of training epochs to 100 and the number of hyperparameter tuning trials to 100 (see BirdNET

documents for details: <https://github.com/kahst/BirdNET-Analyzer>).

BirdNET by default generated confidence scores between 0.1 and 1 as predictions for each 3-sec audio segment and species (Wood and Kahl 2024). For the wild sound test dataset, we optimized the confidence score threshold to determine species presence/absence by maximizing F1 score (the harmonic mean of precision and recall; see below), with thresholds incrementally adjusted from 0.1 to 0.99 in 0.01 steps. For the detection of the playback sounds at site 4, where playback survey dates were known, we experimented with thresholds from 0.1 to 0.9 in 0.1 steps to determine a value that effectively identified playback calls while minimizing false positives on the dates without the playback survey. Maximizing the F1 score was not used for the playback calls because it was impractical to prepare a test dataset reflecting the positive-to-negative sample ratio at site 4 (around 1:200).

Model classification performance was evaluated on the wild sound test dataset and the playback sound test dataset. Performance metrics included precision = $TP/(TP + FP)$ and recall = $TP/(TP + FN)$, where TP, FP, and FN represent true positives, false positives, and false negatives, respectively, assessed on each of the 3-sec audio clips in the test datasets.

Visualization of model embeddings

To assess the similarity between wild and playback sounds, feature embeddings of the two datasets were visualized using Uniform Manifold Approximation and Projection (UMAP) (McInnes et al. 2018, arXiv:1802.03426). Feature embeddings are vectors obtained from an intermediate layer of a machine learning model, which can be considered as a high-level representation of input audio. BirdNET V2.4 produce 1024-dimensional embeddings for a given 3-sec audio clip. These embeddings of the wild and playback datasets were plotted in 2D space with the umap package (Konopka 2023) in the R environment (R Core Team 2024). Points close in the embedding space tend to be plotted closely after applying UMAP (Coenen and Pearce 2019).

Breeding phenology

We applied the trained model to the entire recordings collected on Ishigaki Island (sites 1–3) to infer the breeding phenology of *P. leucomystax* and *R. marina*. Reproductive activity was quantified by counting the number of detections (Kimura and Sota 2023). Given that each recording was 15 minutes in length and the BirdNET analysis unit was 3 seconds, the number of detections per recording ranged from 0 (no detection) to 300 (continuous detections throughout the recording).

Results

Model performance on wild sounds

Automatic hyperparameter tuning resulted in the following values; 2048 hidden units, dropout rate of 0, batch size of 32, learning rate of 0.0001, up-sampling ratio of 0, and no application of mixup data

augmentation. Model precision on the wild sound test dataset ranged from 0.718 to 0.972, and recall ranged from 0.683 to 0.906, depending on species with the optimized detection thresholds (Table 2). For *P. leucomystax*, precision was 0.972 and recall was 0.906 with threshold set as 0.43. For *R. marina*, precision was 0.899 and recall was 0.705 with threshold set as 0.92.

Model performance on playback sounds

Model performance varied depending on the detection threshold, with higher thresholds resulted in higher precision but lower recall. Moderate threshold values around 0.5 provided a better balance of precision and recall compared to a stricter threshold of 0.9 (Table 3). However, a slight reduction of precision at the lower threshold made it difficult to identify playback survey dates (Fig. 2a). Setting the threshold as 0.9 resulted in the clearest peaks of detections on survey dates, except for *P. leucomystax* on December 28 (Fig. 2b). Additionally, relatively large discrepancies from the ground-truth detection counts were observed on July 29 for *P. leucomystax* and on May 25 for *R. marina*, although the survey dates were still identifiable (Fig. 2b). Inspection of the model predictions indicated that the most common mistakes were false negatives due to faint calls. Most false positives were misclassification of audio segments including *M. kuramotoi* as *P. leucomystax* or *R. marina*.

Visualization of model embeddings

Majority of the playback sound embeddings clustered near or overlapped with that of wild sounds (Fig. 3a). However, some audio clips containing playback calls were plotted distant from the wild sound examples (e.g., Fig. 3b) on May 25, July 29, and December 28. On May 25 and July 29, the native *Buergeria choui* called simultaneously with the playback sounds (Fig. 3c). On December 28, dense choruses of *Microhyla kuramotoi* masked *P. leucomystax* playback sounds (Fig. 3d). As mentioned above, detection rates of playback calls were low on these nights. In contrast, playback calls of *R. marina* were clearly visible on spectrograms on December 28 (Fig. 3e) and detected by the model (Fig. 2b). While UMAP visualizations were influenced by parameter settings, the close proximity of wild and playback sounds, as well as the distant plotting of sounds on December 28, remained consistent across parameter adjustments (Supplementary Information S1).

Breeding phenology

Applying the trained model to the recordings from Ishigaki Island revealed that *P. leucomystax* vocalized continuously from mid-April to early November, peaking from May to September (Fig. 4). *Rhinella marina* exhibited intermittent breeding activity without clear seasonality. Detection of *R. marina* was high in April to May 2023, September 2023, and February 2024 at site 2, and in mid-December at site 3. *Rhinella marina* was absent from site 1.

Discussion

Using deep learning to monitor invasive alien species is increasingly popular, but its application in bioacoustics is still few (Kelly et al. 2023; Bota et al. 2024), and the challenge of training deep learning classifiers at a pre-establishment area of invasion front has rarely been discussed. Our custom-trained BirdNET model clearly detected the playback calls on most survey dates when detection thresholds were properly adjusted (Fig. 2b), demonstrating its potential for alien species management on Iriomote Island. Despite some false negatives with short 3-sec audio clips as the detection unit (Table 3), the detection rate was sufficient to identify bouts of consecutive playback calls. The only exception was December 28, when intensive choruses of the native frogs masked the playback calls of *P. leucomystax*. Therefore, using *ex situ* training data from Ishigaki Island did not greatly matter in our case, possibly owing to the faunal similarity between the two islands and the inclusion of training samples from both locations. It should be noted that the model performed better for *R. marina* on Iriomote Island than Ishigaki Island (Table 2,3), but this may be due to the higher sound quality at the playback survey location compared to Ishigaki Island, where *R. marina* often call at distance. Importantly, the playback survey was conducted monthly, with each species' calls lasting only 30–80 seconds in this study. However, the alien frogs could vocalize more frequently during their breeding seasons and may therefore be more easily detected if introduced to the island. For example, in the wild, calls were often detected almost continuously during the 15-minute recording periods (Fig. 4), though the high population density on Ishigaki Island needs to be taken into account.

For managing invasive alien species, avoiding false negatives is crucial to ensure early detection, while excessive false positives are problematic due to the difficulty of manual review (Kelly et al. 2023). Since the introductions of alien frogs to Iriomote Island were occasional (Nakajima et al. 2005), most acoustic monitoring data will consist of negative samples. In this situation, even a slight reduction in precision can produce a large number of false positives and obscure actual detections, as observed at site 4 in our study (Fig. 2a). The prioritization of avoiding false negatives or false positives can be controlled by selecting a low or high detection threshold, respectively, and the threshold should be optimized for the dataset and the available resources for manual review (Kelly et al. 2023). Alternatively, detections could be screened starting from the highest confidence scores downward, as long as resources permit. This approach allows efficient manual review as the proportion of true positives tends to increase with higher confidence scores (Wood and Kahl 2024). BirdNET's 'segments' tool support this strategy by extracting audio segments containing potential detections and naming files with the associated confidence scores.

Our model has several potential applications for managing invasive species. First, it can efficiently screen passive acoustic monitoring (PAM) data for early detection of *P. leucomystax* and *R. marina*. Using PAM could reduce the cost and manpower of current manual surveillance, allowing for expansion of the survey area on Iriomote Island as well as surrounding islands that are difficult to survey frequently (e.g., Hatomajima Island; Ota et al. 2004). Second, it can inform effective surveillance schedules based on breeding phenology (Fig. 4). Third, if *P. leucomystax* and *R. marina* were introduced and established on Iriomote Island, the model could analyze PAM data to monitor its spread using site-occupancy models

(Wood et al. 2020) and to quantify the impacts on native frog communities (Taylor et al. 2017).

Our workflow for training and evaluating a deep learning model is simple and can be adapted to different locations and taxa by replacing the annotated datasets. Both *P. leucomystax* and *R. marina* have been introduced to other tropical regions (Invasive Species Specialist Group ISSG 2015) and possibly continue to expand their range. Given the significant impacts of *R. marina* on local biodiversity known from studies in Australia (Shine 2010), their spread should be carefully monitored at invasion fronts. Other sound-emitting alien species such as birds, insects, and mammals can also be detected through PAM (Ribeiro et al. 2022). Deep learning-based acoustic monitoring can aid early detection of these species.

An obvious limitation of our approach is that playback calls of the target species may not perfectly mimic the soundscape at their early invasion phase. Visualization of model embeddings suggested that the wild and playback sounds have similar features but are not identical (Fig. 3). Differences between the wild calls on Ishigaki Island and the playback sounds on Iriomote Island can be attributed to site-specific background noise in some part, but also to the playback sound property such as individual variation in call characteristics. Although the model effectively detected the playback sounds, similar accuracy is not guaranteed during actual applications.

This study is the first to report the breeding phenology of *R. marina* throughout a year in Japan. Our passive acoustic survey detected their calls in all months, with peaks in spring and late summer at site 2 and in December at site 3 (Fig. 4). *Rhinella marina* is likely a year-round breeder on this island, as evidenced by previous observations of eggs and hatchlings in January, March, and September, which span from winter to summer (Matsui and Seki 2016; Takeuchi and Haramura 2016). Breeding activity of this species is stimulated by rainfall and moderate temperature (Brodie et al. 2021). Additionally, water management of rice paddies likely influences frog reproductive activity at site 2, which was surrounded by the mixture of sugar cane fields and rice paddies. For *P. leucomystax*, the reproductive phenology on Ishigaki Island was similar to that on Okinawajima Island, another introduced population in Japan (Tanaka 2021). This species seems to be reproductively inactive during winter in Japan, unlike in Singapore (Berry 1964). These findings suggest that monitoring during spring to fall is sufficient for detecting *P. leucomystax* in this region, while year-round surveillance potentially enable early detection of *R. marina*.

In summary, this study presented a rapid prototyping approach for developing a deep learning-based acoustic detector targeting two alien frog species and for testing its performance with playback calls. Although training data were collected from Ishigaki Island, the resulting model successfully detected bouts of playback calls on Iriomote Island. These results highlighted the capacity of deep learning models to help surveillance efforts for alien frogs in pre-establishment areas of invasion fronts, facilitating continuous monitoring across broader areas than currently feasible due to resources and manpower limitations. Our observations on the breeding phenology of *P. leucomystax* and *R. marina* on Ishigaki Island further informed effective monitoring strategies. With just 2.6 hours of annotated data—a practical amount even with limited annotation resources—we successfully trained an accurate detector using transfer learning.

This demonstrates the approach's adaptability to other locations and taxa, as it can be fine-tuned for different species or environments with minimal effort, encouraging the development of automated detectors for invasive alien species in other systems.

Data availability

Code and data to produce figures are available at our GitHub repository (<https://github.com/kaede-kmr/Iriomote-IAS-Detect>). Larger audio datasets and the trained model are deposited in Open Science Framework data repository (<https://osf.io/ju594>).

Acknowledgements

We thank local landowners and Banna Park (Okinawa prefecture) for kindly allowing us to place the acoustic recorders; Fumiko Inoue for conducting the playback survey; Chieko Matsumoto, Mitsugi Matsumoto, and the staff of Iriomote Ranger Office and Ishigaki Ranger Office for their help to conduct field surveys; Katsutoshi Watanabe and two anonymous reviewers for their helpful comments to our manuscript. This work was supported by ESPEC Foundation for Global Environment Research and Technology (Charitable Trust).

References

- Aota T, Ashizawa K, Mori H, Toda M, Chiba S (2021) Detection of *Anolis carolinensis* using drone images and a deep neural network: an effective tool for controlling invasive species. *Biol Invasions* 23:1321–1327. <https://doi.org/10.1007/s10530-020-02434-y>
- Beery S, Van Horn G, Perona P (2018) Recognition in Terra Incognita. In: *Computer Vision – ECCV 2018*. Springer International Publishing, pp 472–489
- Berry PY (1964) The breeding patterns of seven species of Singapore anura. *J Anim Ecol* 33:227–243. <https://doi.org/10.2307/2629>
- Bota G, Manzano-Rubio R, Fanlo H, Franch N, Brotons L, Villero D, Devisscher S, Pavesi A, Cavaletti E, Pérez-Granados C (2024) Passive acoustic monitoring and automated detection of the American bullfrog. *Biol Invasions* 26:1269–1279. <https://doi.org/10.1007/s10530-023-03244-8>
- Brodie S, Yasumiba K, Towsey M, Roe P, Schwarzkopf L (2021) Acoustic monitoring reveals year-round calling by invasive toads in tropical Australia. *Bioacoustics* 30:125–141. <https://doi.org/10.1080/09524622.2019.1705183>
- Cañas JS, Toro-Gómez MP, Sugai LSM, Benítez Restrepo HD, Rudas J, Posso Bautista B, Toledo LF, Dena S, Domingos AHR, de Souza FL, Neckel-Oliveira S, da Rosa A, Carvalho-Rocha V, Bernardy JV,

Sugai JLMM, Dos Santos CE, Bastos RP, Llusia D, Ulloa JS (2023) A dataset for benchmarking Neotropical anuran calls identification in passive acoustic monitoring. *Sci Data* 10:771. <https://doi.org/10.1038/s41597-023-02666-2>

Christin S, Hervet É, Lecomte N (2021) Going further with model verification and deep learning. *Methods Ecol Evol* 12:130–134. <https://doi.org/10.1111/2041-210x.13494>

Coenen A, Pearce A (2019) Understanding UMAP. <https://pair-code.github.io/understanding-umap/>. Accessed 16 May 2024

Dorcas ME, Price SJ, Walls SC, Barichivich WJ (2009) Auditory monitoring of anuran populations. In: Dodd CK (ed) *Amphibian ecology and conservation: a handbook of techniques*. Oxford University Press Oxford, pp 281–298

Fricke RM, Olden JD (2023) Technological innovations enhance invasive species management in the anthropocene. *Bioscience* 73:261–279. <https://doi.org/10.1093/biosci/biad018>

Ghani B, Denton T, Kahl S, Klinck H (2023) Global birdsong embeddings enable superior transfer learning for bioacoustic classification. *Sci Rep* 13:22876. <https://doi.org/10.1038/s41598-023-49989-z>

Hasegawa H, Ota H (2017) Parasitic helminths found from *Polypedates leucomystax* (Amphibia: Rhacophoridae) on Miyakojima Island, Ryukyu Archipelago, Japan. *Curr Herpetol* 36:1–10. <https://doi.org/10.5358/hcj.36.1>

Invasive Species Specialist Group ISSG (2015) The Global Invasive Species Database. In: The Global Invasive Species Database. <https://www.iucngisd.org/gisd/>. Accessed 21 Jun 2024

IPBES (2023) Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. In: Roy HE, Pauchard A, Stoett P, Renard Truong T, Bacher S, Galil BS, Hulme PE, Ikeda T, Sankaran KV, McGeoch MA, Meyerson LA, Nuñez MA, Ordonez A, Rahlao SJ, Schwindt E, Seebens H, Sheppard AW, Vandvik V (eds). IPBES Secretariat, Bonn, Germany

Kahl S, Wood CM, Eibl M, Klinck H (2021) BirdNET: A deep learning solution for avian diversity monitoring. *Ecol Inform* 61:101236. <https://doi.org/10.1016/j.ecoinf.2021.101236>

Kelly KG, Wood CM, McGinn K, Kramer HA, Sawyer SC, Whitmore S, Reid D, Kahl S, Reiss A, Eiseman J, Berigan W, Keane JJ, Shaklee P, Gallagher L, Munton TE, Klinck H, Gutiérrez RJ, Peery MZ (2023) Estimating population size for California spotted owls and barred owls across the Sierra Nevada ecosystem with bioacoustics. *Ecol Indic* 154:110851. <https://doi.org/10.1016/j.ecolind.2023.110851>

Kimura K, Sota T (2023) Evaluation of deep learning-based monitoring of frog reproductive phenology. *Ichthyology & Herpetology* 111:563–570. <https://doi.org/10.1643/h2023018>

Konopka T (2023) umap: Uniform Manifold Approximation and Projection. R package version 0.2.10.0. <https://doi.org/10.32614/CRAN.package.umap>

Martínez-González ÁT, Ramírez-Rivera VM, Caballero-Vázquez JA, Jáuregui DAG (2021) Deep learning algorithm as a strategy for detection an invasive species in uncontrolled environment. *Rev Fish*

Biol Fish 31:909–922. <https://doi.org/10.1007/s11160-021-09667-7>

Matsui M, Maeda N (2018) Encyclopedia of Japanese Frogs. Bun-ichi Sogo Shuppan, Tokyo, Japan

Matsui M, Seki S (2016) Nihon-no Kaeru [Frogs of Japan]. Seibundo Shinko Sha, Tokyo, Japan

Nakajima T, Toda M, Aoki M, Tatara M (2005) The project for control of the cane toad *Bufo marinus* on Iriomote Island, Okinawa prefecture. Bulletin of the Herpetological Society of Japan 2005:179–186

National Institute for Environmental Studies (2024) NIES Invasive Species of Japan --A portal site of alien species of Japan--. https://www.nies.go.jp/biodiversity/invasive/index_en.html. Accessed 10 May 2024

Norman DL, Bischoff PH, Wearn OR, Ewers RM, Rowcliffe JM, Evans B, Sethi S, Chapman PM, Freeman R (2023) Can CNN-based species classification generalise across variation in habitat within a camera trap survey? Methods Ecol Evol 14:242–251. <https://doi.org/10.1111/2041-210x.14031>

Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer MS, Packer C, Clune J (2018) Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc Natl Acad Sci USA 115:E5716–E5725. <https://doi.org/10.1073/pnas.1719367115>

Okinawa Prefectural Office (2024) Isolated island related material. In: Okinawa official website. <https://www.pref.okinawa.lg.jp/kensei/tokei/1016451/index.html>. Accessed 25 Nov 2024

O'Shea-Wheller TA, Corbett A, Osborne JL, Recker M, Kennedy PJ (2024) VespAI: a deep learning-based system for the detection of invasive hornets. Commun Biol 7:354. <https://doi.org/10.1038/s42003-024-05979-z>

Ota H, Toda M, Masunaga G, Kikukawa A, Toda M (2004) Feral populations of amphibians and reptiles in the Ryukyu Archipelago, Japan. Glob Environ Res 8:133–143

R Core Team (2024) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Ribeiro JW, Harmon K, Leite GA, de Melo TN, LeBien J, Campos-Cerqueira M (2022) Passive acoustic monitoring as a tool to investigate the spatial distribution of invasive alien species. Remote Sensing 14:4565. <https://doi.org/10.3390/rs14184565>

Schneider S, Greenberg S, Taylor GW, Kremer SC (2020) Three critical factors affecting automated image species recognition performance for camera traps. Ecol Evol 10:3503–3517. <https://doi.org/10.1002/ece3.6147>

Shine R (2010) The ecological impact of invasive cane toads (*Bufo marinus*) in Australia. Q Rev Biol 85:253–291. <https://doi.org/10.1086/655116>

Steinfath E, Palacios-Muñoz A, Rottschäfer JR, Yuezak D, Clemens J (2021) Fast and accurate annotation of acoustic signals with deep neural networks. eLife 10:e68837. <https://doi.org/10.7554/eLife.68837>

Takeuchi H, Haramura T (2016) Tadpoles of the cane toad, *Rhinella marina*, observed in estuarine region. Akamata 26:8–10

Tanaka S (2021) Egg laying season and foam nesting sites of the introduced population of white-lipped tree frog *Polypedates leucomystax* in Okinawajima Island, Ryukyu Archipelago, Japan. Bull Mus Okinawa Pref Mus Art Mus 5:1–10

Taylor A, McCallum HI, Watson G, Grigg GC (2017) Impact of cane toads on a community of Australian native frogs, determined by 10 years of automated identification and logging of calling behaviour. J Appl Ecol 54:2000–2010. <https://doi.org/10.1111/1365-2664.12859>

Wood CM, Gutiérrez RJ, Keane JJ, Peery MZ (2020) Early detection of rapid Barred Owl population growth within the range of the California Spotted Owl advises the Precautionary Principle. Condor 122:1–10. <https://doi.org/10.1093/condor/duz058>

Wood CM, Kahl S (2024) Guidelines for appropriate use of BirdNET scores and other detector outputs. J Ornithol. <https://doi.org/10.1007/s10336-024-02144-5>

Statements and Declarations

Funding

This work was supported by ESPEC Foundation for Global Environment Research and Technology (Charitable Trust).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

All authors contributed to the study conception and design. Data collection was performed by Kaede Kimura and Ibuki Fukuyama. Analysis was performed by Kaede Kimura. The first draft of the manuscript was written by Kaede Kimura, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Table 1 Summary of datasets used in this study. Single sample corresponds to a 3-second audio clip.

Dataset	Sample size	Recording site
Training + validation	3080	Sites 1–6
Wild sound test dataset	720	Sites 1–6
Playback sound test dataset	1041	Site 4

Table 2 Model performance evaluated on wild sound test dataset. TP, FP, FN are the number of true positives, false positives, and false negatives, respectively. Detection thresholds of confidence scores were determined to maximize the F1 score for each species. Bold font are alien species and the others are native species.

Species	TP	FP	FN	Precision	Recall	Threshold
<i>Polypedates leucomystax</i>	242	7	25	0.972	0.906	0.43
<i>Rhinella marina</i>	62	7	26	0.899	0.705	0.92
<i>Buergeria choui</i>	131	5	25	0.963	0.840	0.61
<i>Fejervarya sakishimensis</i>	93	28	41	0.769	0.694	0.47
<i>Kurixalus eiffingeri</i>	79	31	36	0.718	0.687	0.30
<i>Microhyla kuramotoi</i>	418	22	69	0.950	0.858	0.48
<i>Nidirana okinavana</i>	43	3	20	0.935	0.683	0.78
<i>Zhangixalus owstoni</i>	191	35	40	0.845	0.827	0.31

Table 3 Model performance evaluated on playback sound test dataset for two different detection thresholds (0.5 and 0.9). TP, FP, FN are the number of true positives, false positives, and false negatives, respectively.

Species	TP	FP	FN	Precision	Recall	Threshold
<i>Polypedates leucomystax</i>	193	1	44	0.995	0.814	0.5
<i>Rhinella marina</i>	206	5	6	0.976	0.971	0.5
<i>Polypedates leucomystax</i>	147	0	90	1.00	0.620	0.9
<i>Rhinella marina</i>	183	0	29	1.00	0.863	0.9

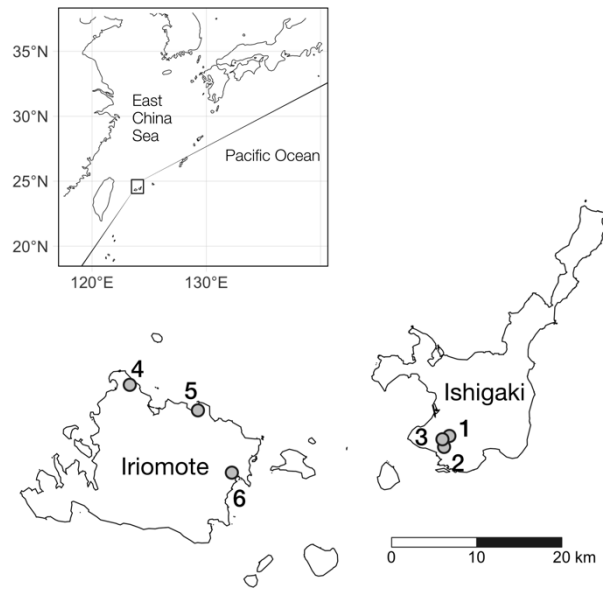
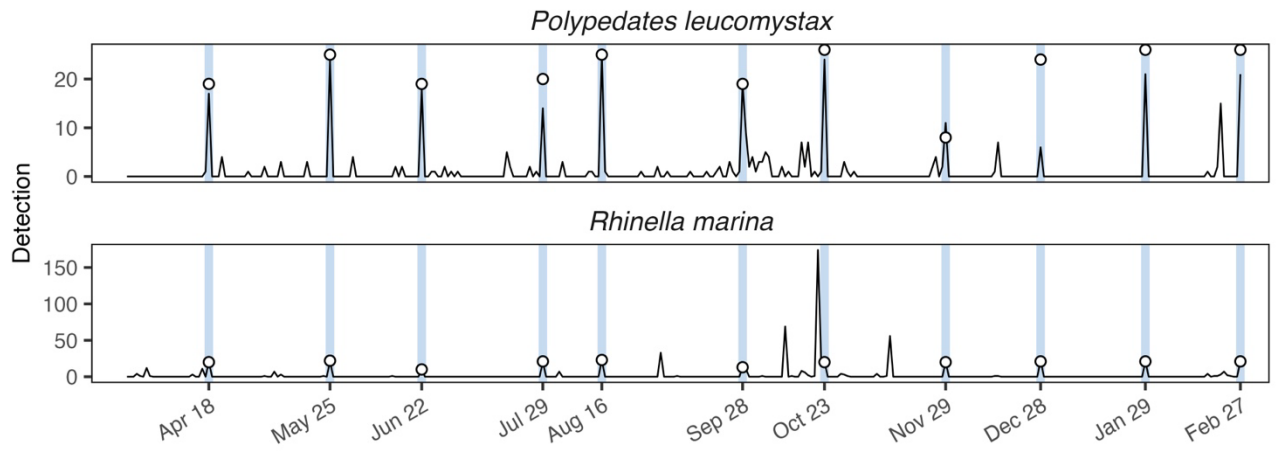


Fig. 1 Map of study sites. Points indicate recording sites. Two alien frog species, *Polypedates leucomystax* and *Rhinella marina*, are widespread on Ishigaki Is. and have been transported to nearby islands including Iriomote Is. Map is modified from digital national land information (MLIT, Japan) (<https://nlftp.mlit.go.jp>)

(a) Threshold 0.5



(b) Threshold 0.9

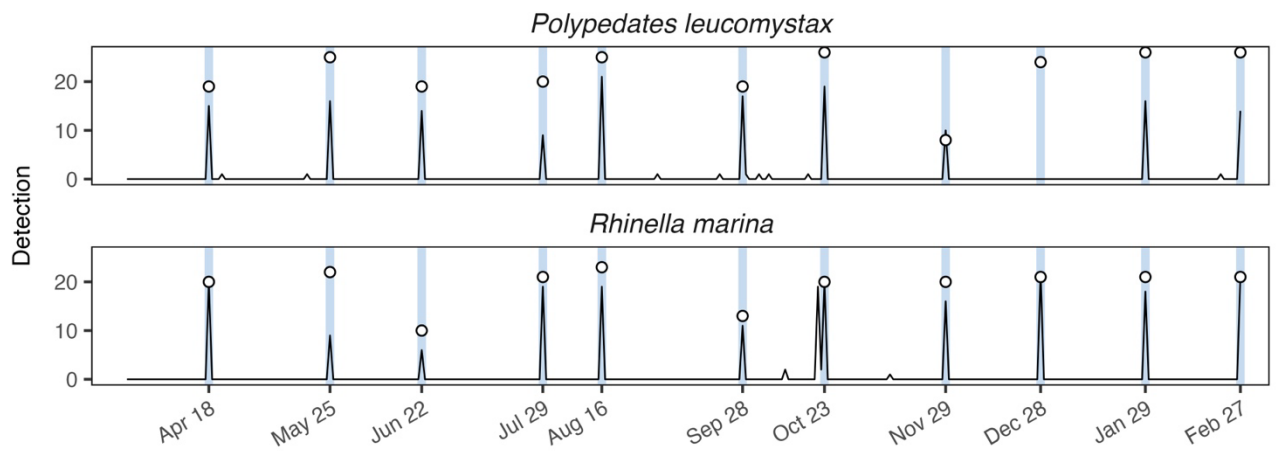


Fig. 2 Detection of playback calls using the custom-trained BirdNET model with the detection threshold set at 0.5 (a) or 0.9 (b). Black lines are the model output, and open circles are the ground-truth values. The closer the black lines are to the open circles, the more accurate the model predictions are. Playback survey dates are highlighted in light blue.

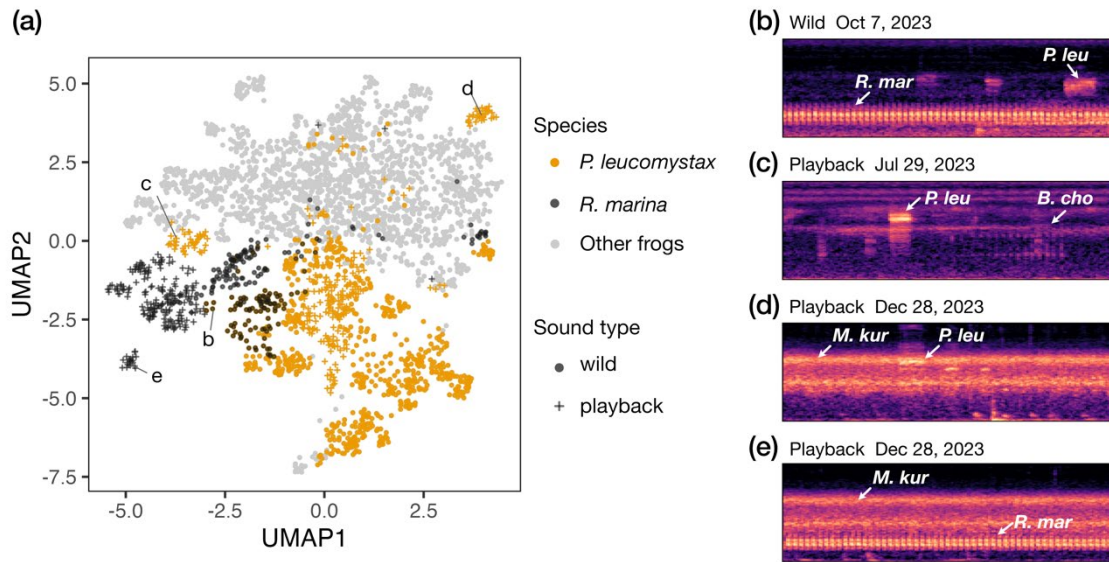


Fig. 3 Visualization of wild and playback sound embeddings by UMAP projection (a), and examples of the corresponding mel-scaled spectrograms (b–e). Spectrograms are converted from 3 sec audio clips with the frequency range of 0 to 8000 Hz. White arrows point to some calls of *Rhinella marina* (*R. mar*), *Polypedates leucomystax* (*P. leu*), *Buergeria choui* (*B. cho*) and *Microhyla kuramotoi* (*M. kur*).

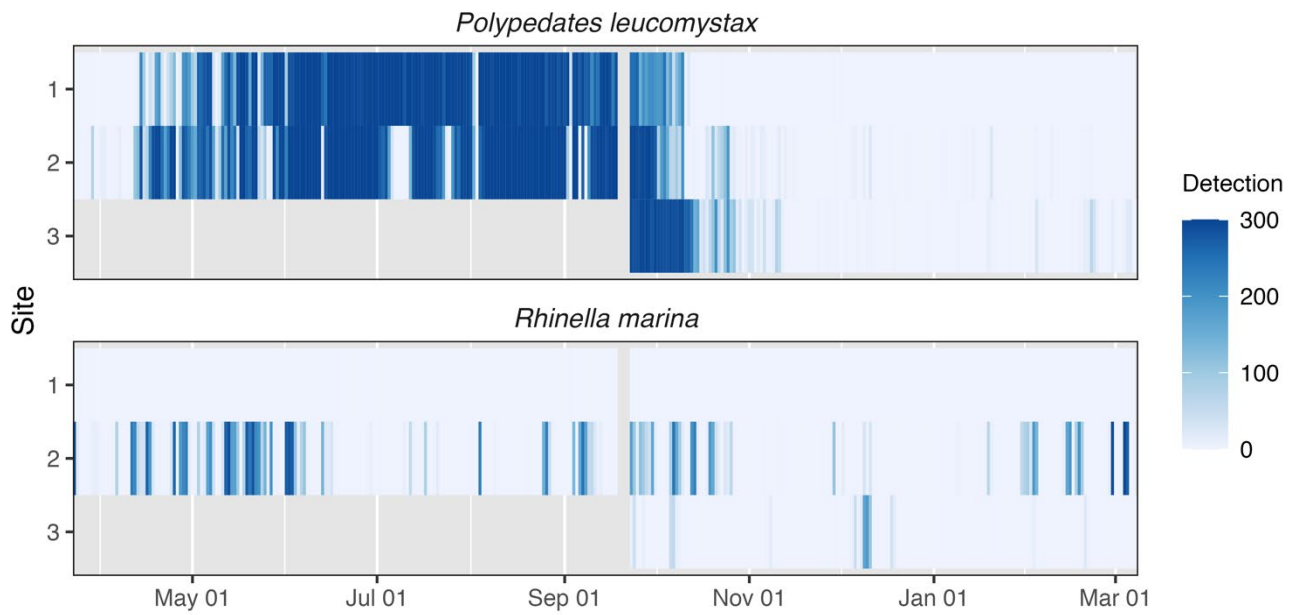


Fig. 4 Breeding phenology of *Polypedates leucomystax* and *Rhinella marina* on Ishigaki Is. inferred from the custom trained BirdNET model. Grey color indicates missing recordings. The recording date starts from March 23, 2023, and end with March 8, 2024.