7-7-2022

NABat ML: Utilizing deep learning to enable crowdsourced development of automated, scalable solutions for documenting North American bat populations

Ali Khalighifar

Benjamin S. Gotthold

Erin Adams

Jenny Barnett

Laura O. Beard

See next page for additional authors

Authors

Ali Khalighifar, Benjamin S. Gotthold, Erin Adams, Jenny Barnett, Laura O. Beard, Eric R. Britzke, Paul A. Burger, Kimberly Chase, Zackary Cordes, Paul M. Cryan, Emily Emily, Christopher T. Fill, Scott E. Gibson, G. Scott Haulton, Kathryn M. Irvine, Lara S. Katz, William L. Kendall, Christen A. Long, Oisin Mac Aodha, Tessa McBurney, Sara McCarthy, Matthew W. McKown, Joy O'Keefe, Lucy D. Patterson, Kristopher A. Pitcher, Matthew Rustand, Jordi L. Segers, Kyle Seppanen, Jeremy L. Siemers, Christian Stratton, Bethany R. Straw, Theodore J. Weller, and Brian E. Reichert

Received: 17 January 2022 | Accepted: 7 July 2022

DOI: 10.1111/1365-2664.14280

Journal of Applied Ecology BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

NABat ML: Utilizing deep learning to enable crowdsourced development of automated, scalable solutions for documenting North American bat populations

Ali Khalighifar[1] | Benjamin S. Gotthold[2] | Erin Adams[3,4] | Jenny Barnett[5] | Laura O. Beard[6] | Eric R. Britzke[7] | Paul A. Burger[8] | Kimberly Chase[9] | Zackary Cordes[10] | Paul M. Cryan[2] | Emily Ferrall[11] | Christopher T. Fill[12] | Scott E. Gibson[13] | G. Scott Haulton[14] | Kathryn M. Irvine[15] | Lara S. Katz[16] | William L. Kendall[17] | Christen A. Long[18] | Oisin Mac Aodha[19] | Tessa McBurney[20] | Sara McCarthy[21] | Matthew W. McKown[22] | Joy O'Keefe[23] | Lucy D. Patterson[24] | Kristopher A. Pitcher[25] | Matthew Rustand[26] | Jordi L. Segers[27] | Kyle Seppanen[28] | Jeremy L. Siemers[29] | Christian Stratton[30] | Bethany R. Straw[2] | Theodore J. Weller[31] | Brian E. Reichert[2]

[1]Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO, USA; [2]Fort Collins Science Center, U.S. Geological Survey, Fort Collins, CO, USA; [3]U.S. Fish and Wildlife Service, Upper Mississippi River National Wildlife and Fish Refuge, Onalaska, WI, USA; [4]U.S. Fish and Wildlife Service, Washington Fish and Wildlife Office, Lacey, WA, USA; [5]U.S. Fish and Wildlife Service, Burbank, WA, USA; [6]Wyoming Game and Fish Department, Lander, WY, USA; [7]Environmental Laboratory, US Army Engineer Research & Development Center, Vicksburg, MS, USA; [8]National Park Service, Anchorage, AK, USA; [9]Florida Fish and Wildlife Conservation Commission, Tallahassee, FL, USA; [10]Ecological Services Section, Kansas Department of Wildlife and Parks, Pratt, KS, USA; [11]Wildlife Conservation Section, Georgia Department of Natural Resources, Social Circle, GA, USA; [12]Nebraska Cooperative Fish and Wildlife Research Unit, University of Nebraska Lincoln, Lincoln, NE, USA; [13]Utah Division of Wildlife Resources, Price, UT, USA; [14]Division of Forestry, Indiana Department of Natural Resources, Indianapolis, IN, USA; [15]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA; [16]Department of Wildlife, Fisheries, and Conservation Biology, University of Maine, Orono, ME, USA; [17]U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO, USA; [18]Bat Conservation International, Austin, TX, USA; [19]School of Informatics, University of Edinburgh, Edinburgh, UK; [20]Canadian Wildlife Health Cooperative, Atlantic Region, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada; [21]Wildlife Division, Fisheries, Forestry and Agriculture, Happy Valley-Goose Bay, Newfoundland and Labrador, Canada; [22]Conservation Metrics, Inc., Santa Cruz, CA, USA; [23]Natural Resources and Environmental Sciences, University of Illinois Urbana Champaign, Champaign, IL, USA; [24]Pukaskwa National Park, Parks Canada Agency, Heron Bay, Ontario, Canada; [25]U.S. Air Force, Avon Park Air Force Range, Avon Park, FL, USA; [26]Royal Gorge Field Office, Bureau of Land Management, Canon City, CO, USA; [27]Canadian Wildlife Health Cooperative, Charlottetown, Prince Edward Island, Canada; [28]Keweenaw Bay Indian Community Natural Resources Department, L'Anse, MI, USA; [29]Colorado Natural Heritage Program, Colorado State University, Fort Collins, CO, USA; [30]Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA and [31]U.S. Department of Agriculture Forest Service, Pacific Southwest Research Station, Arcata, CA, USA

Correspondence
Ali Khalighifar
Email: a.khalighifar@colostate.edu

Funding information
U.S. Geological Survey

Handling Editor: Robert Britton

Abstract

1. Bats play crucial ecological roles and provide valuable ecosystem services, yet many populations face serious threats from various ecological disturbances. The North American Bat Monitoring Program (NABat) aims to use its technology infrastructure to assess status and trends of bat populations, while developing innovative and community-driven conservation solutions.

2. Here, we present *NABat ML*, an automated machine-learning algorithm that improves the scalability and scientific transparency of NABat acoustic monitoring. This model combines signal processing techniques and convolutional neural networks (CNNs) to detect and classify recorded bat echolocation calls. We developed our CNN model with internet-based computing resources ('cloud environment'), and trained it on >600,000 spectrogram images. We also incorporated species range maps to improve the robustness and accuracy of the model for future 'unseen' data. We evaluated model performance using a comprehensive, independent, holdout dataset.
3. *NABat ML* successfully distinguished 31 classes (30 species and a noise class) with overall weighted-average accuracy and precision rates of 92%, and ≥90% classification accuracy for 19 of the bat species. Using a single cloud-environment computing instance, the entire model training process took <16 h.
4. *Synthesis and applications*. Our convolutional neural network (CNN)-based model, *NABat ML*, classifies 30 North American bat species using their recorded echolocation calls with an overall accuracy of 92%. In addition to providing highly accurate species-level classification, *NABat ML* and its outputs are compatible with Bayesian and other statistical techniques for measuring uncertainty in classification. Our model is open-source and reproducible, enabling future implementations as software on end-user devices and cloud-based web applications. These qualities make *NABat ML* highly suitable for applications ranging from grassroots community science initiatives to big-data methods developed and implemented by researchers and professional practitioners. We believe the transparency and accessibility of *NABat ML* will encourage broad-scale participation in bat monitoring, and enable development of innovative solutions needed to conserve North American bat species.

KEYWORDS
automatic identification, bat echolocation calls, bioacoustics monitoring, community scientists, machine learning, North America, quantitative ecology, signal and image processing

1 | INTRODUCTION

Anthropogenic modifications are increasing rates of habitat destruction and biodiversity loss (Ceballos et al., 2015; Nelson et al., 2006; Sugai & Llusia, 2019), exceeding the capacity of conservation biologists to track and protect biodiversity (Wilson, 2017). Of critical concern are bat species, which provide valuable ecosystem services including insect control, seed dispersal and plant pollination (Boyles et al., 2011; Ghanem & Voigt, 2012; Kalka et al., 2008; Medellin et al., 2017). Numerous cave-hibernating bats across the United States and Canada have experienced severe declines over significant portions of their ranges from white-nose syndrome (WNS; Cheng et al., 2021). Additionally, other factors including collisions with wind turbines (Frick et al., 2017; Thompson et al., 2017) and climate change (Hayes & Adams, 2017) may have significant negative effects on North American bat populations. More than half of bat species in the United States, Canada and Mexico are already of conservation concern (O'Shea & Bogan, 2003), yet we lack basic information on where they occur and how their populations are changing over time (Frick et al., 2020). Urgent efforts are needed to address these critical information gaps, such that we will be able to develop innovative solutions for bat conservation.

In order to address these information gaps, robust, scalable bat monitoring networks and technologies are needed. Advances in recording tools and passive acoustic monitoring (PAM) techniques show promise in meeting needs for innovative bat monitoring solutions (Beason et al., 2020; Kloepper et al., 2016). PAM techniques provide potentially low-cost and efficient ways of monitoring bats at scale because they are noninvasive, autonomous and can be remotely deployed (Sugai et al., 2019; Wood et al., 2021); yet, the abundance and rapid pace of acquired PAM data makes analysis challenging (Gibb et al., 2019). Additionally, survey standardization can be difficult, as recommendations regarding performance and biases of various sensors and sampling protocols across different habitats are lacking (Browning et al., 2017), but see Brigham et al. (2004) and Fraser et al. (2020). As such, data collected among various research and monitoring projects are often not directly comparable. Broad-scale monitoring networks like the North American Bat Monitoring Program (NABat; Loeb et al., 2015) offer solutions to problems of inconsistency and latency in the capture and analysis of bat calls by providing standard protocols for collecting PAM recordings, as well as a recording repository and platform for easy data sharing among collaborators. The data infrastructure created by the NABat program provides a crucial opportunity to develop standard, open-source platforms to process and analyse PAM recordings (www.nabatmonitoring.org).

Currently, NABat relies almost entirely on expert biologists and wildlife professionals for data collection efforts (Reichert et al., 2021). Reliance on highly skilled experts provides reliable and consistent data, particularly from often inaccessible geographical regions. However, expanding collection efforts by engaging community scientists and nonprofessionals has strong potential to increase the quantity and representativeness of monitoring data, thus increasing the chances of gathering more consistent, longer-term datasets spanning entire species ranges and decades of ecological disturbance. One major challenge to scaling up efforts for acoustic monitoring is providing timely and representative results to stakeholders and decision makers (Reichert et al., 2018). In their current forms, pipelines for collecting, processing and submitting data to centralized repositories is labour and time intensive, creating opportunities for data loss. This can be improved by increasing accessibility to open-source solutions and reducing timesteps with on-board processing (Figure 1). Automated machine learning and signal processing methods have shown promise in efficiently and accurately processing bat acoustic data, with high potential to meet the above needs. Multiple groups have developed frameworks that process recordings to detect and identify bat species utilizing both signal processing and machine-learning algorithms. These include commercial software packages such as SonoBat (www.sonobat.com), Kaleidoscope Pro (hereafter called Kaleidoscope; Wildlife Acoustics, www.wildlifeacoustics.com), Bat Call ID (BCID; Bat Call Identification, Inc. www.batcallid.com), and free programs such as EchoClass (Eric Britzke; https://www.fws.gov/media/echoclass-instructions-v3-and-software-files), and multiple published open-source machine-learning-based models (Britzke et al., 2011; Skowronski & Fenton, 2008; Skowronski & Harris, 2006). However, these packages remain beyond the means or technological prowess of many potential users. These packages also do not incorporate recent advances in deep learning techniques, which have enabled the application of convolutional neural networks (CNNs; Albawi et al., 2017) to bat signal identification (e.g. BatDetective (Mac Aodha et al., 2018); BatNet (Chen et al., 2020); Pettersson (2020); and Tabak et al. (2022)).

In addition to scalability and improved access, increased transparency in classification error rates and uncertainty from autoclassification software are needed to accelerate scientific efforts informing bat conservation. Autoclassification software programs tend to show lower-than-expected agreement in species classification (Lemen et al., 2015; Nocera et al., 2019) without providing transparency in uncertainty calculations. This may be especially problematic when combining acoustic monitoring data classified using a variety of autoclassification software packages to assess species distributions and related metrics. Current statistical models can help correct for known sources of bias in acoustic monitoring data (Stratton et al., 2022; Wright et al., 2020). However, transparency in the various architectures of automated classification programs will allow for development of new modelling approaches that account for and propagate uncertainty arising from acoustic classification software.

We present a scalable, open-source and fully automated model to detect and identify echolocation calls of North American bats. Using field-recorded, full-spectrum audio files from the NABat community-driven database, we demonstrate how our model combines advanced signal-processing techniques to isolate bat pulses (single sound made by an echolocating bat; Figure 2), and then correctly classifies them to the labelled species using a CNN-based approach (Abadi et al., 2016). By incorporating species distributional information, we further demonstrate how to mitigate the effects of 'unseen' data on model robustness and accuracy. This cloud-based model development can be uniquely tailored to the needs of end-users and was designed to satisfy the assumptions of statistical methods of strong interest to bat researchers, such as Bayesian hierarchical models that link acoustic data with location-specific covariates (e.g. topography, land cover, disease, etc.).

2 | MATERIALS AND METHODS

NABat is an international, multiagency, long-term North American bat monitoring program that seeks to provide reliable data to guide effective conservation solutions at local and continental scales, as well as regularly assess the status and trends in bat species abundance and distribution (Loeb et al., 2015; Reichert et al., 2021). Most bats use echolocation to understand their surrounding environment, search for food, and avoid threats and obstacles (Schnitzler et al., 2003). Bats emit three general types of echolocation calls: search-phase calls to search for and localize prey (e.g. insects), approach-phase calls to track and pursue prey, and terminal-buzz calls that represent the last phase of an echolocation call sequence used immediately preceding and during prey capture (Kalko & Schnitzler, 1989). Here, we developed our species identification model using any types of echolocation calls that pass our signal processing/filtering steps (see Section 2.2.2 for more details), and excluded all social calls from these analyses.

2.1 | Data collection

Expert users and data contributors manually labelled digital audio files of bat echolocation calls in Waveform Audio File format (wav) and submitted them to the NABat database through an online interface (NABat Partner Portal), cloud-to-cloud transfer, or by mailing external hard drives from 2016 to 2021. In total, over 20 experts from 35 NABat community projects provided manually vetted audio files to the NABat database. These data were collected via the NABat community independent of this study, so no ethical approval or licences were required.

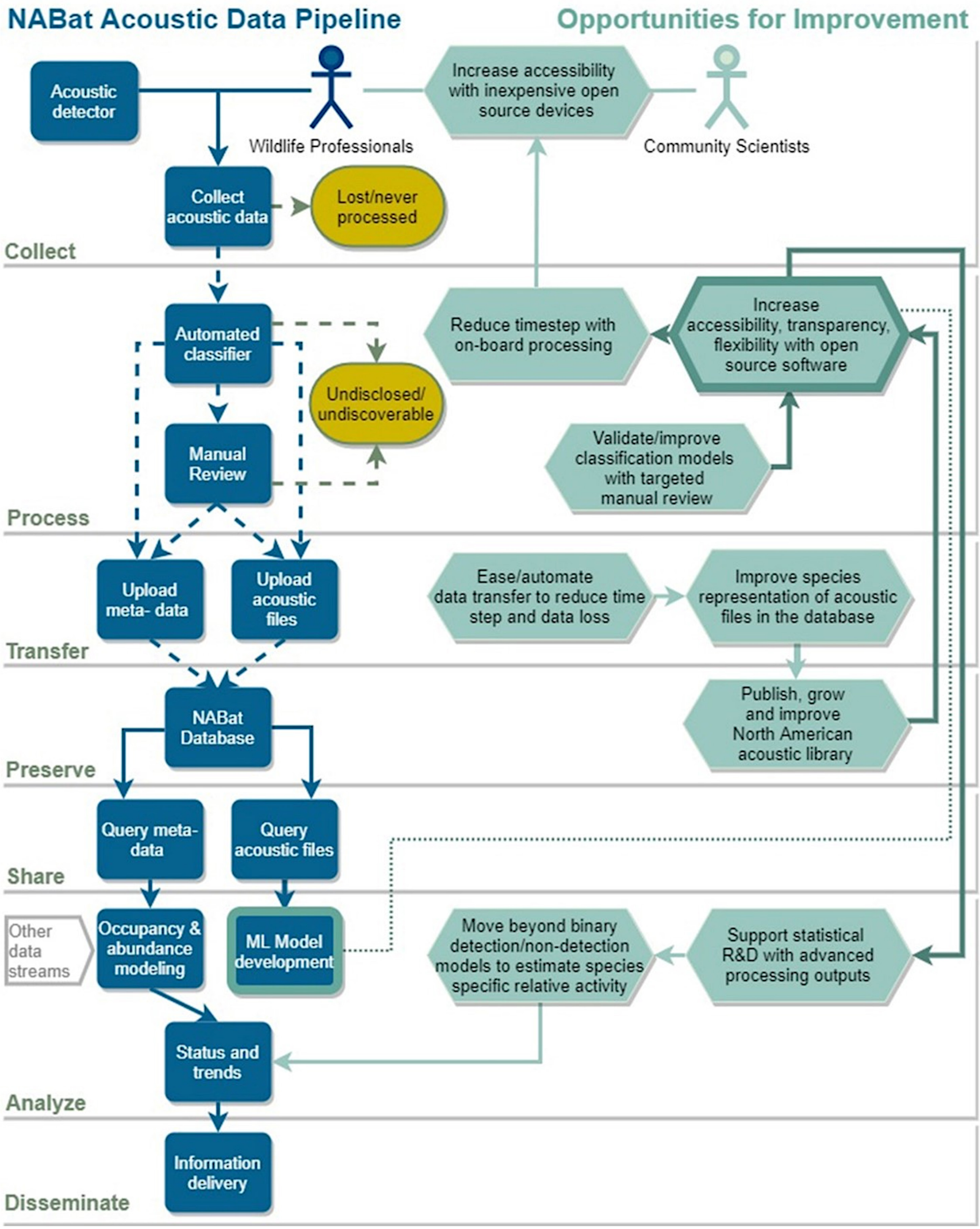
NABat Acoustic Data Pipeline
Opportunities for Improvement
Acoustic detector
Wildlife Professionals
Increase accessibility with inexpensive open source devices
Community Scientists
Collect acoustic data
Lost/never processed
Collect
Automated classifier
Reduce timestep with on-board processing
Increase accessibility, transparency, flexibility with open source software
Undisclosed/ undiscoverable
Manual Review
Validate/improve classification models with targeted manual review
Process
Upload meta- data
Upload acoustic files
Ease/automate data transfer to reduce time step and data loss
Improve species representation of acoustic files in the database
Transfer
NABat Database
Publish, grow and improve North American acoustic library
Preserve
Query meta- data
Query acoustic files
Share
Other data streams
Occupancy & abundance modeling
ML Model development
Move beyond binary detection/non-detection models to estimate species specific relative activity
Support statistical R&D with advanced processing outputs
Status and trends
Analyze
Information delivery
Disseminate

FIGURE 1 North American Bat Monitoring Program (NABat) current acoustic data pipeline (blue) and future strategies (green) for inferring and reporting on status and trends of populations through remote detection of bat calls. Wildlife professionals that record and classify ultrasonic calls of bats with specialized equipment and software form the core of the current NABat acoustic monitoring strategy (left, blue; Loeb et al., 2015). Current protocols suggest a human expert review acoustic files after being processed via an automated classifier to confirm or overturn species identification. Uploading meta-data (e.g. species detections and information about the survey including location, date, time, survey design, etc.) and acoustic files to a common database make them discoverable by researchers who may access them in support of various research objectives. System improvements (right, green) aim to decrease burden on wildlife professionals, lower barriers to community scientist participation, reduce data loss and expand research capabilities. Accessible, transparent and flexible open-source machine learning models for identifying species from acoustics (bolded lines) support other system advancements that will increase the quantity, quality, and throughput of data to ultimately improve and accelerate the delivery of information on the status and trends of North America's bat populations.

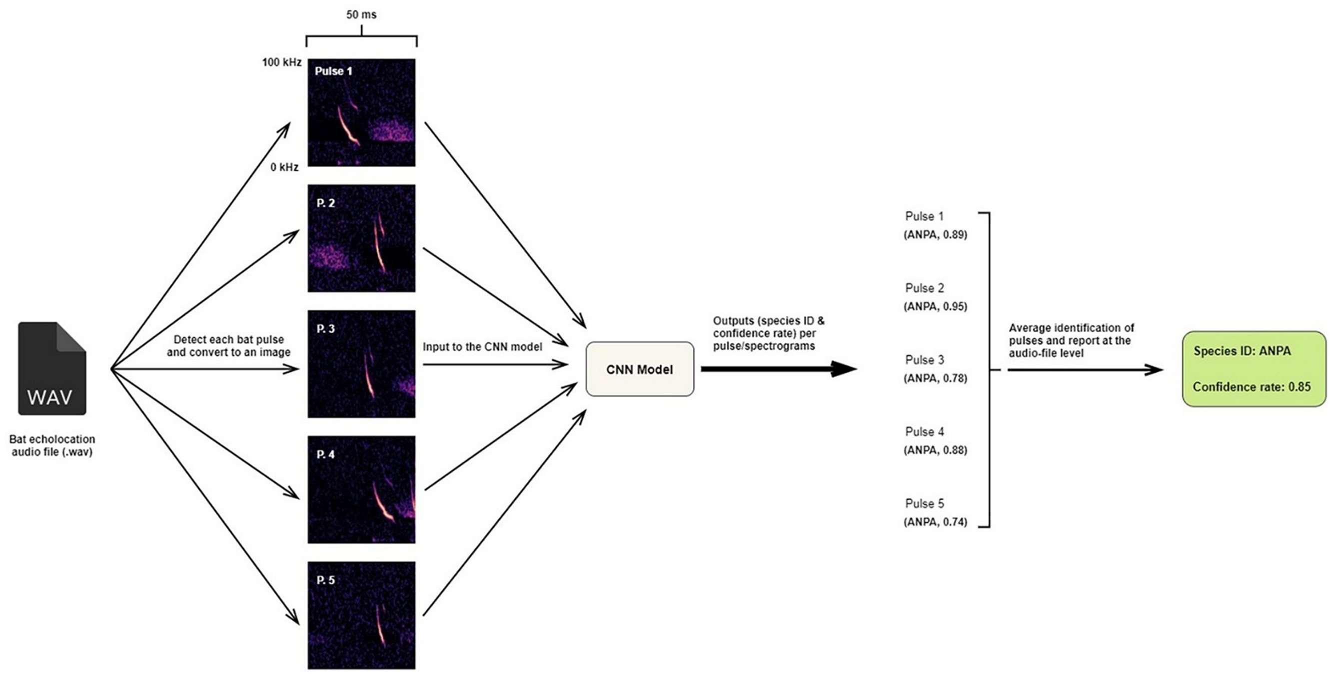

FIGURE 2 Workflow of the fully automated *NABat ML* data processing steps, using an audio file with 5 pulses of pallid bat (*Antrozous pallidus*, or ANPA) as an example. We detect single bat echolocation pulses from the raw digital audio files (WAV files) and convert to images (PNG). Then, we use spectrograms as input to the convolutional neural network (CNN) model. After identifying each spectrogram, we summarize/average the identification outputs (species ID and confidence rate) and report at the audio-file level.

2.2 | Data processing

2.2.1 | Sound reference library

We first created a reference library of bat calls by grouping manually vetted, full-spectrum Waveform Audio File format (.wav) files stored in the NABat database by species and location (i.e. NABat master sample grid cell; Reichert et al., 2021). This resulted in a list of 2459 unique species/NABat grid-cell combinations (Figure 3). To represent geographic variation in our reference library, we randomly drew recordings from each species/NABat grid cell combination in round-robin style until either all recordings associated with a given species were used, or the number of recordings reached 1250. This process resulted in 23,835 recordings for 30 species, and a 'noise' class of ambient, anthropogenic and nonbat noises (see Table 1). We then randomly split our dataset into three parts: 80% of audio files were used to create a training set, 10% for a validation set and 10% for a holdout test set.

2.2.2 | Spectrogram reference library

After randomly assigning all 23,835 recordings into the three datasets, we processed each recording file using Librosa (McFee et al., 2015) library implemented in Python 3.6 (Python Software Foundation. Python Language Reference, version 3.6; http://www.python.org). First, we subsampled each recording with a 50-millisecond sliding window to detect bat call pulses in the recording, regardless of whether it belonged to the training, validation or test set. We specifically chose a 50-ms audio sampling duration because it exceeds that of call pulses emitted by all known species of echolocating bats in North America. Second, we applied a fast

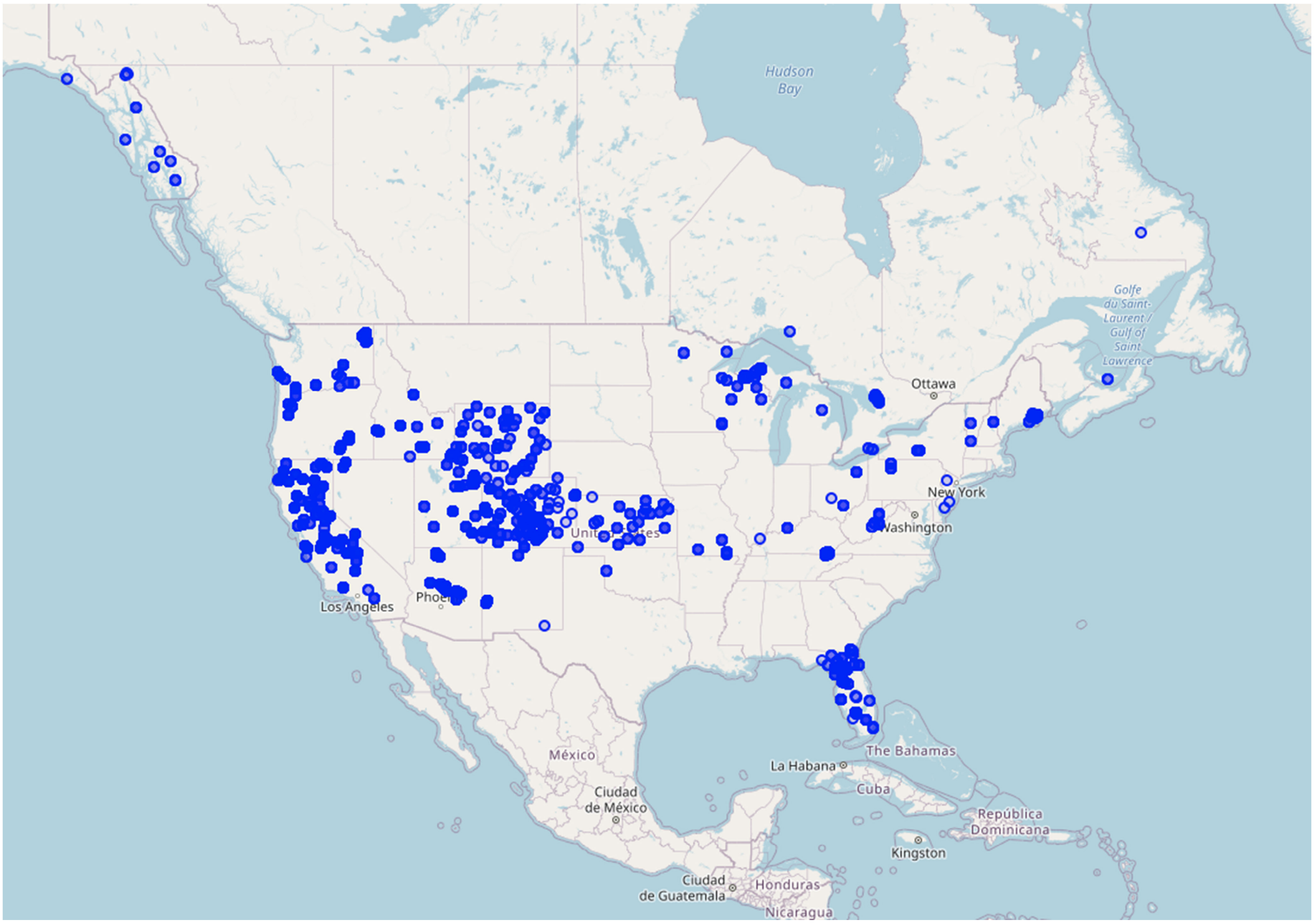

FIGURE 3 The distribution map of 2459 unique species/NABat master sample grid cell combinations (blue dots; Reichert et al., 2021), from which representative recordings of bat calls from 30 different species were selected to train a convolutional neural network (CNN) for an automatic detection and classification system.

Fourier transformation (FFT; Blackman & Tukey, 1958; Heideman et al., 1985), using an FFT window size of 1 ms, an overlap of 25%, with 'hamming' function to all 50-ms samples. Third, we ran all audio samples through a band-pass filter to remove any noise outside of the 5–100 kHz frequency range (see Figure S1 for more details) within which most North American bat species emit echolocation calls. Fourth, we extracted signal features such as frequency, peak amplitude and the timepoint within the recording from which each sample was taken, to be used in denoising and calculation of the signal-to-noise ratio (SNR). Fifth, we applied a denoising technique to each 50-ms sample by subtracting the median amplitude from each single value of the sample matrix. Finally, we calculated a SNR for each sample after denoising.

We filtered and excluded any 50-ms audio samples that met one or more of the following criteria from analysis: (1) the focal bat pulse occurred within the first or last 10 ms of the sample, in order to prevent bat pulses from getting cut off at the edges of the sample; (2) the greatest frequency magnitude was equal to or <21 (short-term Fourier transform in Librosa; McFee et al., 2015); or (3) the SNR was equal to or <7. After excluding any samples that met the above criteria, we converted the remaining samples to spectrogram images in portable graphics format (PNG) format. Each 50-ms image was 100 × 100 pixels, with a 95 kHz frequency range on the *y*-axis. *y*-axis lower and upper bounds were set at 5 and 100, respectively. We used these spectrograms as inputs to our CNN model (Figure 2).

2.2.3 | Model architecture and training

We applied a CNN modelling approach to detect and classify bat pulses in spectrograms associated with call recordings from the 30 bat species and a noise class. CNN models are comprised of three types of layers: (1) convolutional layers, which apply hierarchical feature extraction and decomposition to input images; (2) pooling layers, which carry out operations to reduce the number of parameters and computations in the network; and (3) fully connected layers, which perform classification at the end of the process, where each input is connected to all nodes of that layer. A main advantage of CNN models over traditional machine-learning classifiers is their automated image feature extraction, which eliminates the need for users to determine a priori which image features will be important for accomplishing the desired detection and classification tasks. We

TABLE 1 Summary of the 31 classes analysed, correct identification rates and sample sizes of validation set (pulse-level identification) and test set (audio-file level identification) for 30 North American bat species and a noise class

			Validation set (pulses)		Test set (audio files)	
Scientific name	Common name	Label	Identification rate	Sample size	Identification rate	Sample size
Antrozous pallidus	Pallid bat	ANPA	0.69	773	0.78	23
Corynorhinus townsendii	Townsend's big-eared bat	COTO	0.32	642	0.91	34
Eptesicus fuscus	Big brown bat	EPFU	0.89	4190	0.91	117
Euderma maculatum	Spotted bat	EUMA	0.78	1029	0.97	38
Eumops perotis	Greater bonneted bat	EUPE	0.83	287	0.72	25
Idionycteris phyllotis	Allen's big-eared bat	IDPH	0.43	282	0.25	4
Lasionycteris noctivagans	Silver-haired bat	LANO	0.68	2541	0.92	122
Lasiurus blossevillii	Western red bat	LABL	0.63	269	0.53	17
Lasiurus borealis	Eastern red bat	LABO	0.72	4519	0.84	125
Lasiurus cinereus	Hoary bat	LACI	0.71	1447	0.84	116
Lasiurus intermedius	Northern yellow bat	LAIN	0.93	2824	0.96	101
Lasiurus seminolus	Seminole bat	LASE	0.71	2663	0.89	114
Myotis austroriparius	Southeastern myotis	MYAU	0.88	586	1.00	9
Myotis californicus	California myotis	MYCA	0.84	4126	0.91	101
Myotis ciliolabrum	Western small-footed myotis	MYCI	0.85	3986	0.95	127
Myotis evotis	Long-eared myotis	MYEV	0.89	3919	0.97	119
Myotis grisescens	Grey bat	MYGR	0.97	1248	1.00	42
Myotis leibii	Eastern small-footed myotis	MYLE	0.94	2495	0.97	96
Myotis lucifugus	Little brown bat	MYLU	0.82	3734	0.83	127
Myotis septentrionalis	Northern myotis	MYSE	0.81	1214	0.76	46
Myotis sodalis	Indiana bat	MYSO	0.96	2977	0.98	85
Myotis thysanodes	Fringed myotis	MYTH	0.90	3609	0.94	110
Myotis velifer	Cave bat	MYVE	0.14	175	0.25	4
Myotis volans	Long-legged myotis	MYVO	0.76	3019	0.97	76
Myotis yumanensis	Yuma myotis	MYYU	0.84	3442	0.98	121
Nycticeius humeralis	Evening bat	NYHU	0.79	5221	0.93	136
Nyctinomops macrotis	Big free-tailed bat	NYMA	0.69	310	0.87	23
Parastrellus hesperus	Canyon bat	PAHE	0.96	3704	0.99	83
Perimyotis subflavus	Tri-coloured bat	PESU	0.89	3480	0.93	117
Tadarida brasiliensis	Brazilian free-tailed bat	TABR	0.74	1661	0.91	109
—	Noise	NOISE	0.77	56	1.00	5

specifically chose to use a CNN because they show reliably high performance with image classification tasks and pattern recognition (Gu et al., 2018; Krizhevsky et al., 2017).

We developed a CNN model with 10 hidden layers, using the Keras (https://keras.io/api/) and TensorFlow 2 (Abadi et al., 2016) software libraries in Python 3.6 (see Table S1), and used Amazon Web Services (AWS) SageMaker for training and evaluation. Keras is a high-level, deep-learning application programming interface (API) running on TensorFlow, an open-source, end-to-end machine learning platform. We used a cloud-based modelling environment (ml.p3.2xlarge AWS instance) with 1 Graphic Processing Unit (GPU), 61 GB of memory, and 8 virtual Central Processing Units (vCPUs). We trained our CNN model on 611,637 spectrograms associated with the 31 classes (30 species and 1 noise class) in the training set using a batch size of 32. We considered the validation loss value (the sparse categorical cross-entropy function) as a factor to find optima in the training process. The loss function is part of the model optimization, which was calculated repeatedly to measure error of the model's current status (Akbari et al., 2021). To achieve the highest performance possible, we completed the training process in two separate rounds. First, we initiated the training process with a learning rate of 1×10^{-4} using the Adam optimizer and rectified linear unit activation function (ReLU). We used dropout regularization layers to avoid over-fitting during the training process, and the Early Stopping method in Keras to monitor validation loss values, which automatically stopped the training process once model prediction performance failed to increase on the holdout validation dataset. After completing the first round, we reduced the learning rate to 1×10^{-6} and kept all other configurations the same before beginning the second training round. We ultimately chose this process for training our model because it resulted in the best performance (with lowest validation loss) on the validation set.

2.2.4 | Model evaluation

Our validation set consisted of 74,773 spectrograms from 2396 recordings. Throughout the training process, we calculated overall weighted accuracy, precision, recall (hereafter, correct identification rate) and F1-score to measure and report model performance using this set. We used a softmax function as the final layer of our CNN network to normalize model output and provide a confidence rate (between 0 and 1) to each classification label assignment. We measured the distribution of these confidence rates to identify a threshold confidence rate for filtering false positives at the evaluation step (see below).

The remaining 10% of recordings in our sound reference library constituted an independent test set of 2426 recordings resulting in 82,103 spectrograms. To evaluate model performance after completing the training process, we calculated the same metrics for measuring and reporting model performance detailed above. However, since the test set represents a simulation of the 'real-world', we measured model performance at the audio-file level. To report metrics of model performance at the audio-file level, we set three conditions:

1. *Number of detected pulses*: In order to receive an identification, each recording needed to contain at least one detectable bat pulse. Lack of a detectable bat pulse resulted in returning "NoID" for the audio file.
2. *Mean confidence rate of pulses*: If the confidence rate for classification of every detected pulse within an audio file was below 0.57 (i.e. the confidence rate threshold calculated using the validation set), the model returned 'NoID' for the audio file. We calculated this threshold using an optimal cut-point value with the receiver operating characteristic (ROC) technique (optimal AUC), and used it to reduce our false positive rate (see Figure S2).
3. *Species distributional information*: We incorporated species distributional data as secondary information to improve our identification rate and further reduce our false positive rate. To do so, we used species range maps (North American Bat Species Distribution, 2014), and added a 300 km buffer to each range. After assigning a species label to each audio file, we compared the recording's location with the species range. If no overlap was found between the buffered species range and the recording's location, the model returned the next species label for the audio file. This process continued until the recording's location overlapped with the species range map.

All codes for processing audio files, training and evaluating the CNN model, as well as the image and sound reference libraries are made available via U.S. Geological Survey data and software release (https://doi.org/10.5066/P969TX8F; https://doi.org/10.5066/P9XJRJZX) to facilitate future training and validation of our classifier in applied settings.

3 | RESULTS

We calibrated our bat classification model through two separate training rounds using 611,637 spectrograms. In the first round, the early stopping method terminated the training process after 19 epochs, as no further improvement was observed in model performance. Using the same method, the second round lasted 18 epochs, and ultimately achieved a validation loss score of 0.75. This was the lowest validation loss value achieved across all tested approaches. Each epoch took an average of 25 min, resulting in ~15.5 h of total training process time across both rounds. We then used independent validation and test sets to evaluate our model's performance at the pulse- and audio-file levels of identification, respectively.

Validation set: We applied our fully automated classifier to the 2396 audio files of the validation dataset, resulting in 74,773 spectrograms for identification (see Table 1). We created a confusion matrix to depict model results at the pulse level for 31 classes (Figure 4). The overall accuracy and precision rates for pulse-level identification were 83 and 80%, respectively. We achieved >80% correct identification rate for 16 classes. Of these, correct identification rates were highest (>95%) for grey bat (*Myotis grisescens*, MYGR), Indiana bat (*Myotis sodalis*, MYSO), and canyon bat (*Parastrellus hesperus*, PAHE). The lowest correct identification rates were for cave bat (*Myotis velifer*, MYVE), Townsend's big-eared bat (*Corynorhinus townsendii*, COTO), Allen's big-eared bat (*Idionycteris phyllotis*, IDPH), and western red bat (*Lasiurus blossevillii*, LABL), with 14%, 32%, 43% and 63%, respectively. Notably, species with the lowest correct identification rates were represented by the fewest audio samples among the 30 species of bats tested (MYVE, $n = 175$; LABL, $n = 269$; IDPH, $n = 282$).

Test set: We used our test dataset to evaluate model performance by applying our classifier to 2426 audio files, generating 82,103 spectrograms (see Table 1, Figure 5). After classifying each spectrogram within an audio file, we summed their confidence rates per species and labelled the file with the species name that had the highest cumulative value, provided all three recording criteria were met (see Methods). Applying these conditions resulted in 'NoID' returns, or change of species labels, for 93 audio files (i.e.

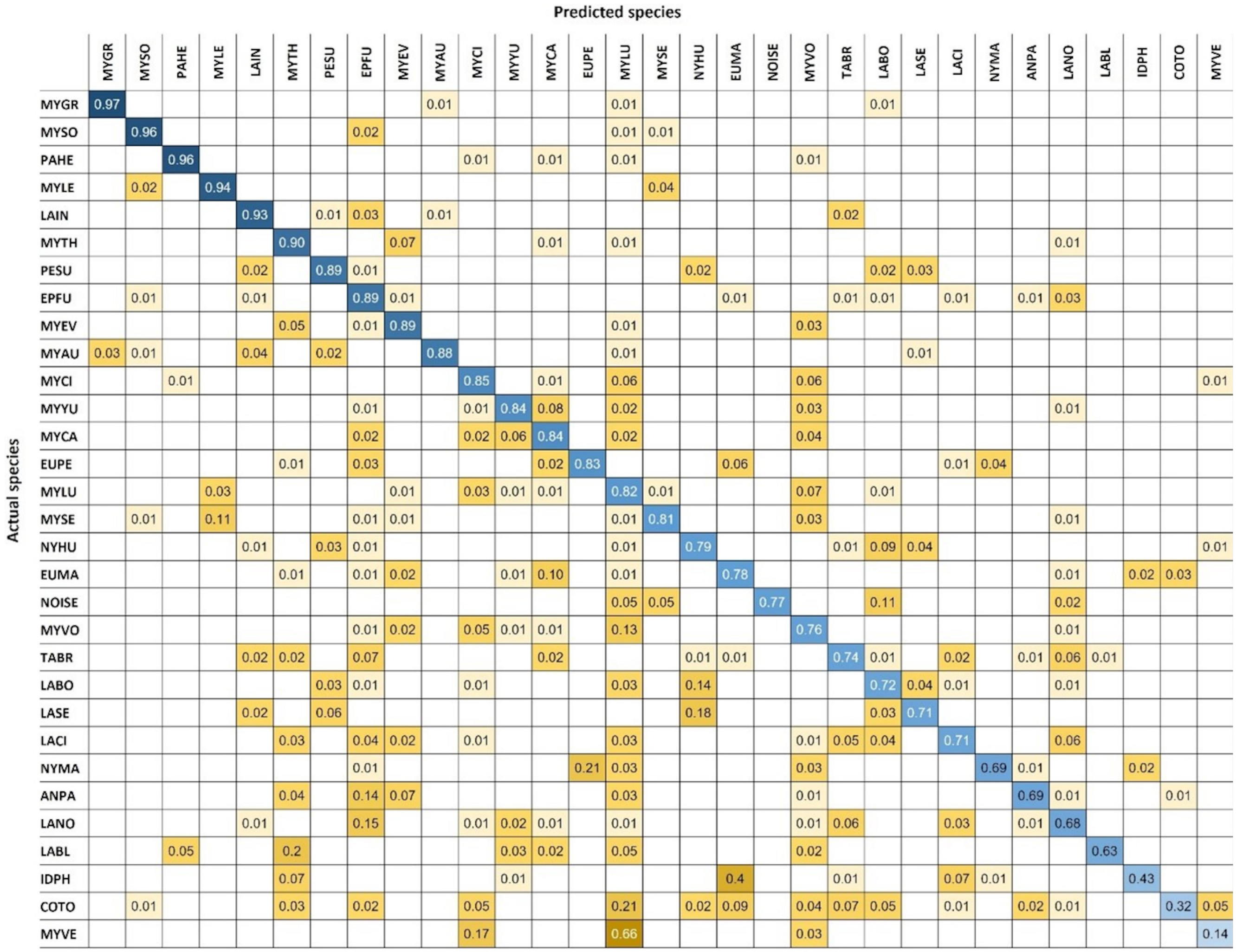

FIGURE 4 Confusion matrix for the 31 classes of North American bat species (using species codes) and the noise class in the validation set (pulse-level identification). Blue, correct identifications; brown, misidentifications. All values of zero or <1% are removed for ease of visualization.

3.8% of recordings in test dataset). Recording locations for 39 files (1.6%) did not coincide with the range of the top identified species: 29 files (1.2%) had a mean confidence rate below 0.57, and 25 (1%) had <1 detectable bat pulse. This process resulted in a 92% overall weighted-average classification accuracy and precision at the recordings level. Incorporating species distribution information and confidence rate criteria improved our overall classification accuracy by 2%.

At the audio-file level, we achieved ≥90% correct identification rates for 19 of the bat species classes, out of a total 31 classes (Figure 5). The highest correct identification rates were for calls of southeastern myotis (*Myotis austroriparius*, MYAU), MYGR, and the noise class, each at 100%, followed by PAHE at 99%. Notably, MYAU and the noise class were represented by only 9 and 5 samples in the test set, respectively (Table 1). The lowest correct identification rates were for MYVE (25%), IDPH (25%), and LABL (53%); these species were represented by correspondingly sparse representation in the test set, with 4, 4, and 17 samples, respectively. Our model performed poorly when classifying hoary bat calls (*Lasiurus cinereus*, LACI); known calls of that species were confused with 8 other species: 4% of recordings associated with LACI were misidentified as Brazilian free-tailed bat (*Tadarida brasiliensis*, TABR), 2% each as evening bat (*Nycticeius humeralis*, NYHU), silver-haired bat (*Lasionycteris noctivagans*, LANO), big brown bat (*Eptesicus fuscus*, EPFU), eastern red bat (*Lasiurus borealis*, LABO) and little brown bat (*Myotis lucifugus*, MYLU), and 1% each as MYSO and spotted bat (*Euderma maculatum*, EUMA). Additional model performance details, including precision and F-1 scores for all 31 classes at both the pulse and audio-file levels, are located in Table S2.

4 | DISCUSSION

We present *NABat ML*, a fully automated, scalable and open-source model to detect and classify bat pulses from field recordings. Our model converts these pulses to images, identifies each pulse spectrogram at the species level using an advanced CNN-based model and reports the final identification at the audio-file level. After testing on

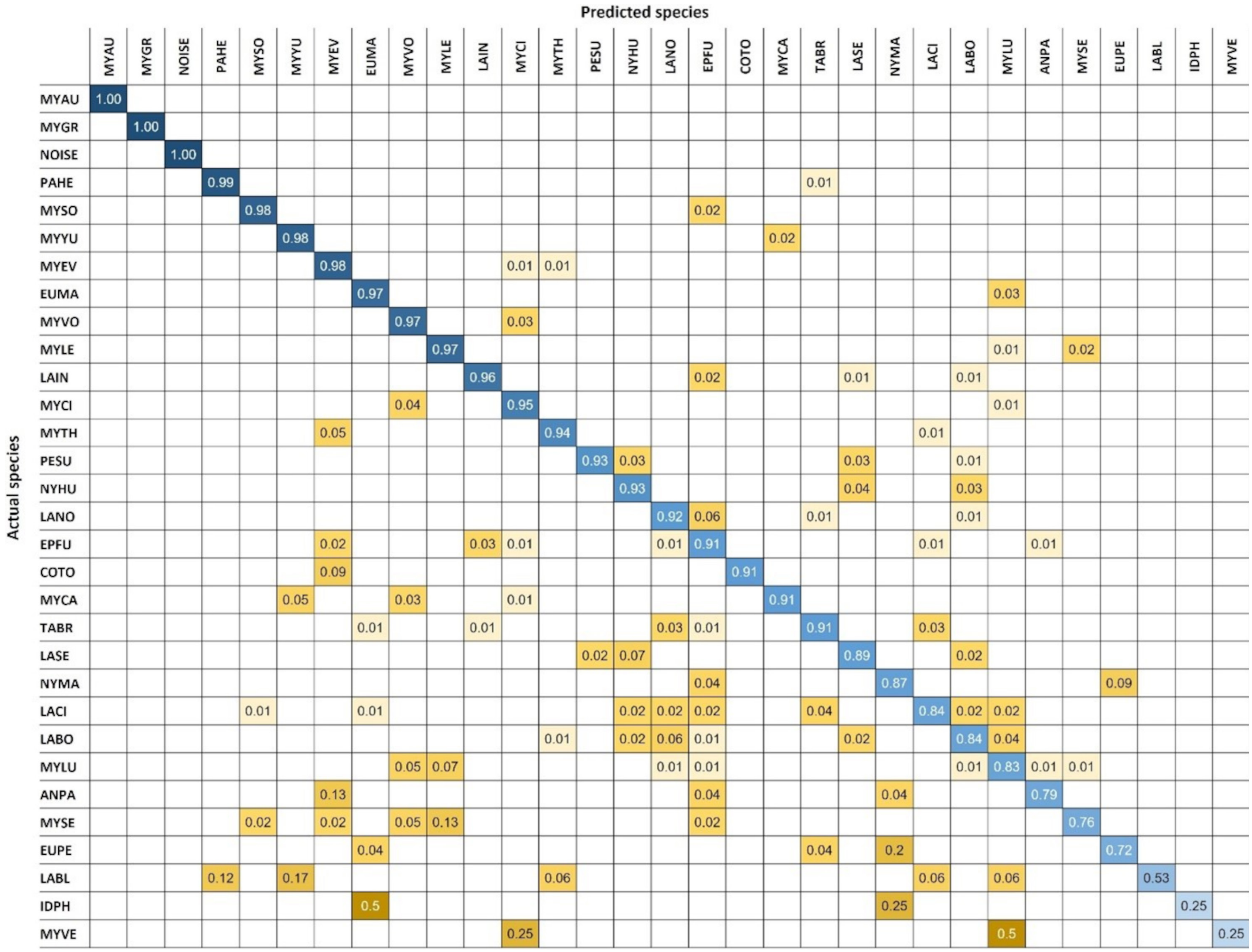

FIGURE 5 Confusion matrix for the 31 classes of North American bat species (using species codes) and the noise class in the test set (audio-file level identification). Blue, correct identifications; brown, misidentifications. All values of zero are removed for ease of visualization.

an independent dataset, our model successfully identified recordings of 31 classes (30 species of bats and a noise class) with overall weighted-average accuracy and precision rates of 92%. We achieved correct identification rates of ≥90% for 20 of these classes, including the noise class. Notably, three of the 20 classes represent federally endangered species in the United States (MYGR and MYSO; U.S. Fish and Wildlife Service, 1991), or species of Special Concern in Canada (EUMA; COSEWIC, 2004). Additionally, the tri-coloured bat (*Perimyotis subflavus*, PESU) is federally listed as Endangered under the Species at Risk Act in Canada (COSEWIC, 2012), and is currently under review for protection under the U.S. Endangered Species Act (U.S. Fish and Wildlife Service, 2017). We also achieved >75% precision across all classes, indicating that our model has a low rate of false positives at the audio-file level. To the best of our knowledge, the model we presented is the most comprehensive deep learning-based bat echolocation classifier available, benefiting from a combination of advanced signal and image processing techniques. Our model is reproducible, fully automated, performs at analysis speeds faster than real-time, and can be used to process enormous quantities of data generated by PAM techniques. As an open-source software application with encouraging classification accuracy, this model shows promise for improving large-scale bat monitoring efforts in North America.

CNNs are increasingly developed to analyse data collected through PAM techniques, camera traps, and other autonomous monitoring devices because of their proven effectiveness at automated feature extraction, image decomposition and information gathering from huge volumes of data (Allen et al., 2021; Gray et al., 2019; Norouzzadeh et al., 2018). In particular, multiple studies have successfully applied CNNs to bat echolocation calls to detect and/or classify species (Chen et al., 2020; Mac Aodha et al., 2018; Pettersson, 2020; Tabak et al., 2022). However, these foundational prior efforts fell short of meeting NABat's goals to assess status and trends of bat populations well into the future. For example, although BatDetective (Mac Aodha et al., 2018) successfully detects bat calls in recordings, the CNN-based model does not classify species producing those calls. BatNet (Chen et al., 2020) classifies species with 86% overall accuracy, but is limited to Asian bats and requires manual extraction of signals from audio files. Tabak et al.'s (2022) CNN-based model classifies 10 North American species of bats with 90% overall accuracy, but is not compatible with cloud environments like the NABat database. Additionally, this model was trained on acoustic data recorded in zero-crossing (ZC) format, which makes the model rely on a reduced set of information that could be critical for bat call identification. As such, our model adds to a pool of CNN-based bat identification models that is capable of identifying echolocation calls with high accuracy, increases equitable access and offers maximum transparency, while addressing the needs of the NABat and bat research communities at large.

We also incorporated species distributional information to further improve our model's robustness to future unseen data. However, our method differs from the classification process used in other common software packages, which incorporate species lists based on geographic information prior to classification. In early tests, we experimented with incorporating species distributional information into our model prior to pulse classification, and experimentally removed species that were not known to occur in a specific area (see Figure S3 and S4). For example, if one species was not known to occur where a call was recorded, that species was not considered a potential classification candidate for that recording. Although this experimental test resulted in a slight improvement in overall weighted-average classification accuracy (93% overall accuracy), we ultimately decided to incorporate distributional information after classification to account for disagreement and uncertainty in existing species ranges, as well as any needed model adjustments as ranges shift over time.

While CNN models can detect and classify echolocation calls of bats with high accuracy, several model performance challenges remain. As exemplified in our analysis by MYVE, IDPH and LABL, this is particularly true when classifying species represented by few training examples, resulting in lower correct classification rates. This can be addressed by incorporating more training data containing a broader sampling of intra-specific variation, which can create models that are more generalizable to the true variation present in echolocation calls across the ranges of widely distributed species (Murray et al., 2001; Russo et al., 2018). Additionally, our model assigns only one label to each spectrogram, even though a single audio file may be representative of multiple species. Future steps may include utilizing object detection CNNs, which would allow separate detection of pulses within a single spectrogram/audio file, or recurrent neural networks that better handle time-series data.

Previous studies using machine learning algorithms to identify bat calls that report high test set accuracy sometimes have reduced performance when applied to field recordings (Rydell et al., 2017). Determining the degree of truth of call labels used in model training remains a major challenge. The NABat program and community recording database benefit from many skilled experts manually vetting calls, most of which were initially recorded and flagged by existing software programs. Human error in species identification is inevitable and hard to detect, especially because many existing algorithms were designed to discriminate on human-recognizable patterns. When training models to classify bat species from calls, a required assumption is that species labels are correct, and that call-structure variability is representative of each species across its range. Because our model was trained on only a subset of representative calls, as well as calls that were selected and labelled by humans and algorithms, it likely incorporated bias toward certain features of search-phase calls that were previously identified by experts as diagnostic of each species. It is possible that the classification accuracy we observed was high because the test set included few calls deemed ambiguous by humans. To increase our model's applicability to field recordings in the future, we plan to create a fully independent and more representative call library (e.g. more call types and validated raw species recordings) to reduce human-introduced bias in labelled calls used to train classification algorithms.

We built *NABat ML* using an open-source framework to make it possible for anyone to participate and develop bat acoustic monitoring solutions, including software packages, cloud-based web applications, and single-board computer and microcontroller hardware sensor platforms (e.g. AudioMoth, Arduino, and Raspberry Pi). Alone or in combination, open technologies could greatly expand options for in-field, real-time bat call detection and classification, benefiting the goals of community scientists, expert bat biologists, wildlife professionals, and other interested stakeholders (Figure 1). For example, outputs from *NABat ML* can be used to build Bayesian hierarchical models that link acoustic data with location-specific covariates of bat occurrence and activity (Stratton et al., 2022; Wright et al., 2020), applied to bat monitoring and conservation programs, or utilized in education and outreach programs that encourage citizen scientists and community members to engage in local bat research. We can also apply various statistical procedures to account for uncertainty and bias in our species detection and classification process (Barré et al., 2019), providing the NABat community and other researchers with transparency about our current model's ability to perform classification tasks. We hope this new framework encourages users and data contributors with diverse expertise and backgrounds to collect bat acoustic data to help expand the NABat database to all species of North American bats and improve the current representation of their acoustic characteristics in the NABat database, which can be used regularly to update *NABat ML*. Ultimately, this will further improve the technology infrastructure necessary for effectively assessing status and trends of bat populations, as well as addressing important ecological questions involving bats.

In conclusion, North American bats are facing serious threats across the continent (Cheng et al., 2021; Hayes & Adams, 2017; Thompson et al., 2017), and due to the crucial ecological roles they play (Boyles et al., 2011; Ghanem & Voigt, 2012; Kalka et al., 2008; Medellin et al., 2017), there is great need for effective population monitoring to support conservation efforts. The NABat program aims to measure the impacts of range-wide threats to bats and offer innovative conservation solutions for bat populations through the development of *NABat ML*, a fully automated, reproducible CNN-based model for detecting and classifying bat species calls. This model can be implemented at local, national and international scales to address various ecological questions about bats. It is also fully accessible and scientifically transparent, such that it can promote community awareness, education, and participation in critical bat research by encouraging individuals with diverse backgrounds to engage in monitoring efforts. We believe that the methodological transparency and increased accessibility of our model will ultimately enable the development of the innovative conservation solutions needed to ensure the health and vitality of North American bat populations.

AUTHOR CONTRIBUTIONS

Ali Khalighifar, Benjamin Gotthold, Paul Cryan, Kathryn Irvine, William Kendall, Oisin Mac Aodha, Christian Stratton, Bethany Straw and Brian Reichert conceived the ideas and designed the methodology. Erin Adams, Jenny Barnett, Laura Beard, Eric Britzke, Paul Burger, Kimberly Chase, Zackary Cordes, Emily Ferrall, Christopher Fill, Scott Gibson, Scott Haulton, Lara Katz, Christen Long, Tessa McBurney, Sara McCarthy, Matthew McKown, Joy O'Keefe, Lucy Patterson, Kristopher Pitcher, Matthew Rustand, Jordi Segers, Kyle Seppanen, Jeremy Siemers and Theodore Weller collected the data. Ali Khalighifar and Benjamin Gotthold analysed the data. Ali Khalighifar, Paul Cryan and Brian Reichert led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We are grateful to Winifred F. Frick, as well as the NABat community for supporting this research. We also thank Nat Goodby and Donald Solick for reviewing and helping improve an earlier draft of this manuscript. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT

The software and data presented in this study will be openly available via: (1) North American Bat Monitoring Program: NABat Acoustic ML, Version 1.0.1: U.S. Geological Survey software release, https://doi.org/10.5066/P9XJRJZX (Gotthold et al., 2022a) and (2) Training dataset for NABat Machine Learning V1.0: U.S. Geological Survey data release, https://doi.org/10.5066/P969TX8F (Gotthold et al., 2022b).

ORCID

Ali Khalighifar https://orcid.org/0000-0002-2949-8143
Paul M. Cryan https://orcid.org/0000-0002-2915-8894
Kathryn M. Irvine https://orcid.org/0000-0002-6426-940X
William L. Kendall https://orcid.org/0000-0003-0084-9891
Kristopher A. Pitcher https://orcid.org/0000-0002-1687-244X
Brian E. Reichert https://orcid.org/0000-0002-9640-0695

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., & Isard, M. (2016). Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation* (pp. 265–283). https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf

Akbari, A., Awais, M., Bashar, M., & Kittler, J. (2021). How does loss function affect generalization performance of deep learning? Application to human age estimation. In *38th International Conference on Machine Learning* (pp. 141–151). Proceedings of Machine Learning Research, Virtual Conference.

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *International Conference on Engineering and Technology* (pp. 1–6). IEEE.

Allen, A. N., Harvey, M., Harrell, L., Jansen, A., Merkens, K. P., Wall, C. C., Cattiau, J., & Oleson, E. M. (2021). A convolutional neural network for automated detection of humpback whale song in a diverse, long-term passive acoustic dataset. *Frontiers in Marine Science*, *8*, 607321.

Barré, K., Le Viol, I., Julliard, R., Pauwels, J., Newson, S. E., Julien, J. F., Claireau, F., Kerbiriou, C., & Bas, Y. (2019). Accounting for automated identification errors in acoustic surveys. *Methods in Ecology and Evolution*, *10*, 1171–1188.

Beason, R. D., Riesch, R., & Koricheva, J. (2020). Temporal pass plots: An intuitive method for visualising activity patterns of bats and other vocalising animals. *Ecological Indicators*, *113*, 106202.

Blackman, R. B., & Tukey, J. W. (1958). The measurement of power spectra from the point of view of communications engineering—Part I. *Bell System Technical Journal*, *37*, 185–282.

Boyles, J. G., Cryan, P. M., McCracken, G. F., & Kunz, T. H. (2011). Economic importance of bats in agriculture. *Science*, *332*, 41–42.

Brigham, R., Kalko, E., Jones, G., Parsons, S., & Limpens, H. (2004). *Bat echolocation research: Tools, techniques and analysis*. Bat Conservation International.

Britzke, E. R., Duchamp, J. E., Murray, K. L., Swihart, R. K., & Robbins, L. W. (2011). Acoustic identification of bats in the eastern United States: A comparison of parametric and nonparametric methods. *The Journal of Wildlife Management*, *75*, 660–667.

Browning, E., Gibb, R., Glover-Kapfer, P., & Jones, K.E. (2017) Passive acoustic monitoring in ecology and conservation. *WWF conservation technology series*, pp. 76. Woking, UK, WWF-UK.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*, e1400253.

Chen, X., Zhao, J., Chen, Y.-H., Zhou, W., & Hughes, A. C. (2020). Automatic standardized processing and identification of tropical bat calls using deep learning approaches. *Biological Conservation*, *241*, 108269.

Cheng, T. L., Reichard, J. D., Coleman, J. T., Weller, T. J., Thogmartin, W. E., Reichert, B. E., Bennett, A. B., Broders, H. G., Campbell, J., & Etchison, K. (2021). The scope and severity of white-nose syndrome on hibernating bats in North America. *Conservation Biology.*, *35*, 1586–1597.

COSEWIC. (2004). *Committee on the status of endangered wildlife in Canada*. https://sararegistry.gc.ca/virtual_sara/files/cosewic/sr_spotted_bat_e.pdf

COSEWIC. (2012). Committee on the Status of Endangered Wildlife in Canada. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments/tri-colored-bat-technical-summary-2012.html

Fraser, E. E., Silvis, A., Brigham, R. M., & Czenze, Z. J. (2020). *Bat echolocation research: A handbook for planning and conducting acoustic studies* (2nd ed.). Bat Conservation International.

Frick, W. F., Baerwald, E. F., Pollock, J. F., Barclay, R. M., Szymanski, J., Weller, T., Russell, A., Loeb, S., Medellin, R., & McGuire, L. (2017). Fatalities at wind turbines may threaten population viability of a migratory bat. *Biological Conservation*, *209*, 172–177.

Frick, W. F., Kingston, T., & Flanders, J. (2020). A review of the major threats and challenges to global bat conservation. *Annals of the New York Academy of Sciences*, *1469*, 5–25.

Ghanem, S. J., & Voigt, C. C. (2012). Increasing awareness of ecosystem services provided by bats. *Advances in the Study of Behavior*, *44*, 279–302.

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2019). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, *10*, 169–185.

Gotthold, B. S., Khalighifar, A., Straw, B. R., & Reichert, B. E. (2022a). *North American bat monitoring program: NABat acoustic ML, version 1.0.1*. U.S. Geological Survey Software Release. https://doi.org/10.5066/P9XJRJZX

Gotthold, B. S., Khalighifar, A., Straw, B. R., & Reichert, B. E. (2022b). *Training dataset for NABat machine learning V1.0*. U.S. Geological Survey Data Release. https://doi.org/10.5066/P969TX8F

Gray, P. C., Bierlich, K. C., Mantell, S. A., Friedlaender, A. S., Goldbogen, J. A., & Johnston, D. W. (2019). Drones and convolutional neural networks facilitate automated and accurate cetacean species identification and photogrammetry. *Methods in Ecology and Evolution*, *10*, 1490–1500.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., & Cai, J. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354–377.

Hayes, M. A., & Adams, R. A. (2017). Simulated bat populations erode when exposed to climate change projections for western North America. *PLoS ONE*, *12*, e0180693.

Heideman, M. T., Johnson, D. H., & Burrus, C. S. (1985). Gauss and the history of the fast Fourier transform. *Archive for History of Exact Sciences*, *34*, 265–277.

Kalka, M. B., Smith, A. R., & Kalko, E. K. (2008). Bats limit arthropods and herbivory in a tropical forest. *Science*, *320*, 71.

Kalko, E. K., & Schnitzler, H.-U. (1989). The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*. *Behavioral Ecology and Sociobiology*, *24*, 225–238.

Kloepper, L. N., Linnenschmidt, M., Blowers, Z., Branstetter, B., Ralston, J., & Simmons, J. A. (2016). Estimating colony sizes of emerging bats using acoustic recordings. *Royal Society Open Science*, *3*, 160022.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the Association for Computing Machinery (ACM)*, *60*, 84–90.

Lemen, C., Freeman, P. W., White, J. A., & Andersen, B. R. (2015). The problem of low agreement among automated identification programs for acoustical surveys of bats. *Western North American Naturalist*, *75*, 218–225.

Loeb, S. C., Rodhouse, T. J., Ellison, L. E., Lausen, C. L., Reichard, J. D., Irvine, K. M., Ingersoll, T. E., Coleman, J. T., Thogmartin, W. E., & Sauer, J. R. (2015). *A plan for the north American bat monitoring program (NABat)* (pp. 1–100). United States Department of Agriculture Forest Service.

Mac Aodha, O., Gibb, R., Barlow, K. E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G. R., Newson, S. E., Pandourski, I., Parsons, S., Ross, J., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Girolami, M., Brostow, G., & Jones, K. E. (2018). Bat detective—Deep learning tools for bat acoustic signal detection. *PLoS Computational Biology*, *14*, e1005995.

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and music signal analysis in python. In *Proceedings of the 14th Python in Science Conference* (pp. 18–25). Citeseer.

Medellin, R. A., Wiederholt, R., & Lopez-Hoffman, L. (2017). Conservation relevance of bat caves for biodiversity and ecosystem services. *Biological Conservation*, *211*, 45–50.

Murray, K. L., Britzke, E. R., & Robbins, L. W. (2001). Variation in search-phase calls of bats. *Journal of Mammalogy*, *82*, 728–737.

Nelson, G. C., Bennett, E., Berhe, A. A., Cassman, K., DeFries, R., Dietz, T., Dobermann, A., Dobson, A., Janetos, A., & Levy, M. (2006). Anthropogenic drivers of ecosystem change: An overview. *Ecology and Society*, *11*, 29.

Nocera, T., Ford, W. M., Silvis, A., & Dobony, C. A. (2019). Let's agree to disagree: Comparing auto-acoustic identification programs for northeastern bats. *Journal of Fish and Wildlife Management*, *10*, 346–361.

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, E5716–E5725.

North American Bat Species Distributions. (2014). https://www.sciencebase.gov/catalog/item/59de9ba6e4b05fe04ccd3b2e.

O'Shea, T. J., & Bogan, M. A. (2003). Monitoring trends in bat populations of the United States and territories: Problems and prospects. In *Information and technology report* (p. 274). U.S. Geological Survey, Fort Collins Science Center.

Pettersson, S. (2020). *Classifying bats to species based on ultrasound recordings: Systems using artificial neural networks and random forests*. Uppsala University.

Reichert, B., Lausen, C., Loeb, S., Weller, T., Allen, R., Britzke, E., Hohoff, T., Siemers, J., Burkholder, B. & Herzog, C. (2018). A guide to processing bat acoustic data for the North American Bat Monitoring Program (NABat). *U.S. Geological Survey Open-File Report*, 2018-1068, 33 p.

Reichert, B. E., Bayless, M., Cheng, T. L., Coleman, J. T., Francis, C. M., Frick, W. F., Gotthold, B. S., Irvine, K. M., Lausen, C., & Li, H. (2021). NABat: A top-down, bottom-up solution to collaborative continental-scale monitoring. *Ambio*, *50*, 901–913.

Russo, D., Ancillotto, L., & Jones, G. (2018). Bats are still not birds in the digital era: Echolocation call variation and why it matters for bat species identification. *Canadian Journal of Zoology*, *96*, 63–78.

Rydell, J., Nyman, S., Eklöf, J., Jones, G., & Russo, D. (2017). Testing the performances of automated identification of bat echolocation calls: A request for prudence. *Ecological Indicators*, *78*, 416–420.

Schnitzler, H.-U., Moss, C. F., & Denzinger, A. (2003). From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology & Evolution*, *18*, 386–394.

Skowronski, M. D., & Fenton, M. B. (2008). Model-based detection of synthetic bat echolocation calls using an energy threshold detector for initialization. *The Journal of the Acoustical Society of America*, *123*, 2643–2650.

Skowronski, M. D., & Harris, J. G. (2006). Acoustic detection and classification of microchiroptera using machine learning: Lessons learned from automatic speech recognition. *The Journal of the Acoustical Society of America*, *119*, 1817–1833.

Stratton, C., Irvine, K. M., Banner, K. M., Wright, W. J., Lausen, C., & Rae, J. (2022). Coupling validation effort with in situ bioacoustic data improves estimating relative activity and occupancy for multiple species with cross-species misclassifications. *Methods in Ecology and Evolution*, *13*, 1–16.

Sugai, L. S. M., & Llusia, D. (2019). Bioacoustic time capsules: Using acoustic monitoring to document biodiversity. *Ecological Indicators*, *99*, 149–152.

Sugai, L. S. M., Silva, T. S. F., Ribeiro, J. W., Jr., & Llusia, D. (2019). Terrestrial passive acoustic monitoring: Review and perspectives. *Bioscience*, *69*, 15–25.

Tabak, M. A., Murray, K. L., Reed, A. M., Lombardi, J. A., & Bay, K. J. (2022). Automated classification of bat echolocation call recordings with artificial intelligence. *Ecological Informatics*, *68*, 101526.

Thompson, M., Beston, J. A., Etterson, M., Diffendorfer, J. E., & Loss, S. R. (2017). Factors associated with bat mortality at wind energy facilities in the United States. *Biological Conservation*, *215*, 241–245.

U.S. Fish and Wildlife Service (USFWS). (1991). 50 CFR part 17, endangered and threatened wildlife and plants: 5-year review of listed species. *Federal Register*, *56*, 56882–56900.

U.S. Fish and Wildlife Service (USFWS). (2017). 50 CFR part 17, endangered and threatened wildlife and plants: 90-day findings for five species. *Federal Register*, *82*, 60362–60363.

Wilson, E. O. (2017). Biodiversity research requires more boots on the ground. *Nature Ecology & Evolution*, *1*, 1590–1591.

Wood, C. M., Klinck, H., Gustafson, M., Keane, J. J., Sawyer, S. C., Gutiérrez, R., & Peery, M. Z. (2021). Using the ecological significance of animal vocalizations to improve inference in acoustic monitoring programs. *Conservation Biology*, *35*, 336–345.

Wright, W. J., Irvine, K. M., Almberg, E. S., & Litt, A. R. (2020). Modelling misclassification in multi-species acoustic data when estimating occupancy and relative activity. *Methods in Ecology and Evolution*, *11*, 71–81.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Khalighifar, A., Gotthold, B. S., Adams, E., Barnett, J., Beard, L. O., Britzke, E. R., Burger, P. A., Chase, K., Cordes, Z., Cryan, P. M., Ferrall, E., Fill, C. T., Gibson, S. E., Haulton, G. S., Irvine, K. M., Katz, L. S., Kendall, W. L., Long, C. A., Mac Aodha, O. … Reichert, B. E. (2022). *NABat ML*: Utilizing deep learning to enable crowdsourced development of automated, scalable solutions for documenting North American bat populations. *Journal of Applied Ecology*, *00*, 1–14. https://doi.org/10.1111/1365-2664.14280