ACOUSTICS 2025
JOONDALUP
SOUNDS OF THE SUNSET COAST

Neural network detection and unsupervised clustering reveal song variability in a cryptic passerine

Lance De Vine* (1), Callan Alexander* (1), Shannon Maguire (2), Daniella Teixeira (1), Susan Fuller (1)

(1) School of Biology and Environmental Science, Queensland University of Technology, Brisbane, Australia
(2) Currumbin Wildlife Sanctuary

* Authors contributed equally to this work

ABSTRACT

Automated birdsong detection models are becoming essential tools for surveying cryptic and threatened species, yet species with highly variable vocalisations and complex repertoires can present significant classification challenges. This study presents an overview of a machine learning approach applied to the critically endangered Australian species Eastern Bristlebird (*Dasyornis brachypterus*). The Eastern Bristlebird's northern population has fewer than 50 individuals remaining in the wild following decades of habitat loss and altered fire regimes. These ground-dwelling birds inhabit dense grassy forest understorey, making visual detection extremely challenging, but they have a diverse and complex repertoire with highly variable song. Our methodology centres on a shallow neural network architecture designed to identify vocal classes and generalise to differing song types with minimal training data requirements, coupled with unsupervised feature analysis for repertoire investigation. Despite their architectural simplicity, our shallow networks produce effective results with minimal training data. Clustering analysis using global birdsong embeddings was also performed enabling repertoire characterisation, investigation of call type variations, and site-specific vocal patterns. These methods offer promising avenues for automated monitoring of species with complex vocal repertoires, with the potential for improved conservation management and population assessment.

1 INTRODUCTION

The northern population of Eastern Bristlebird (NEBB) is one of Australia’s most at risk avian populations with less than 50 known birds remaining in the wild (Charley et al., 2021). This population has experienced a substantial decline in the last 30 years, in both range and population size (Holmes, 1989). Reduced fire frequency over the last three decades has contributed to a significant loss of habitat (Stone et al., 2022). The key habitat requirements for the species (excluding heath habitats) are large patches (>40Ha) of contiguous grassy understorey, as well as tall tussocks and a high mean grass height (Stone et al., 2018). Territories typically exist close to rainforest, which it is presumed that the birds use as a refuge (Holmes, 1989). The species is known for its complex and diverse vocal repertoire and is ground-dwelling. They tend to inhabit dense foliage, which means that they have developed a reputation for being difficult to detect using traditional survey methodologies. The highly vocal nature of this species makes them an ideal candidate for acoustic monitoring, but the complex nature of their repertoire poses significant challenges for automated detection.Their songs are thought to vary even at the individual level (Baker, 1998).

In recent years deep learning based acoustic classifiers have been successfully used to detect vocalizations for many of the world's bird species (Borowiec et al., 2022; Ghani et al., 2023; Huus et al., 2025; Kahl et al., 2021) Such classifiers enable the automated analysis of large acoustic datasets. As the use of acoustic monitoring expands accordingly, unsupervised methodologies provide useful utility for bioacoustics tasks at large scales, particularly when applied to deep learning embeddings or acoustic features (Alexander et al., 2025; Bravo Sanchez et al., 2024; McGinn et al., 2023). In this study we combine both supervised and unsupervised approaches to investigate the repertoire of the northern population of the Eastern Bristlebird with the view to discern the level of song variability within and across the population. This is of particular importance as captive-bred birds are being released into the wild population, and acoustic monitoring may provide an opportunity to observe calling behaviour changes over time and also potentially aid in the identification of populations or individuals.

2 METHODOLOGY

2.1 Data Collection

12 Audiomoths (Hill et al., 2018) were deployed across 3 locations in northern NSW where wild Eastern Bristlebird were known to occur. 3 Audiomoths were deployed at Garima Conservation Reserve, adjacent to cages containing captive-bred and translocated Eastern Bristlebirds from Currumbin Wildlife Sanctuary (see Table 1). All Audiomoths were set to record for two hours shortly after dawn, at 48kHz sample rate and medium-high gain.

Table 1: Data Sources used in this Study

Deployments	Hours	Location
A	605	Northern NSW
B	451	Northern NSW
C	338	Northern NSW
Captive	486	Garima Conservation Reserve

2.2 Model architecture

A custom convolutional neural network (CNN) classifier was developed to detect vocalisations in passive audio recordings. Like many birdsong models, our classifier processes sequential spectrogram patches and reports call occurrences within temporal windows. However, our model is deliberately smaller and faster, using only 10 CNN layers compared to larger general-purpose recognisers. This compact design enables rapid processing and easy customisation for calls with unusual features, though it can be more challenging to train than transfer-learning approaches based on the use of large pre-trained models like BirdNET or Perch. The classifier generates spectrograms (0-8 KHz frequency range) using a 512-sample FFT window with 50% overlap and a 2.5-second classification window. The neural network architecture resembles ResNet-10, comprising 4 standard CNN layers, 3 residual connection CNN layers, and 3 fully connected classification layers. Initially, we attempted to separate EBB calls into approximately 10 classes based on manually verified detections, but subsequently simplified the approach to a binary classifier with all calls combined into a single class. The model was implemented using the PyTorch Python library.Data preparation and Training of the Neural Network.Instances of EBB calls were manually labelled using Raven software.

The initial training dataset was build by labelling approximately 10 hours of audio from two recorders. A subset of approximately 200 of these initial EBB call examples were used as the first training dataset. This data included six different call types loosely identified as being aurally distinct. An iterative training process was followed (see (Eichinski et al., 2022). Training consisted of a two staged approach, 1) several epochs with a small batch size (size 4) and learning rate of 0.001, and 2) a fine-tuning phase for several epochs with large batch size (size 32) and a small learning rate 0.00002. During training data augmentation was implemented, consisting of sparse uniform noise, image cropping and image contrast adjustment. The classifier made predictions on all recordings detailed in table 1.

2.3 Visualization, Clustering and Variation of EBB Calls

All detections were bandpassed at 3kHz and 8kHz (the frequency range of the species) to minimise the impacts of background sounds, particularly Bell Miners (*Manorina melanophrys*). Global birdsong model embeddings (BirdNET and Perch 2.0) were then generated for all detections, alongside embeddings from the trained model. UMAP dimensionality reduction and HDBSCAN clustering were applied to the embeddings (n neighbours = 15, min dist. 0.01, min cluster size = 150, min samples = 10). An iteration with PCA applied beforehand was also tested (n components = 30).

3 PRELIMINARY RESULTS

The model performs strongly with an ROC-AUC of 0.88 and a weighted average F1 score of 0.91 at the optimal threshold. This approach was able to generalise well to field data and located song types noticeably different to those contained in the initial training dataset. This testing was not conducted on a subset of the training data, but rather on a complex held-out dataset of 5000 audio segments containing a wide range of unseen vocalisations, other species, recorder artifacts and weather noise. This was designed to test the ability of the model to generalise and be more representative of field conditions. Approximately 45000 potential Eastern Bristlebird vocalisations were detected across the 15 recorders. The model has since been used to locate the species in areas where it had not been detected in several years. The clustering results suggest that the different populations of Eastern Bristlebird appear to have distinct differences between their vocalisations, with a particular distinction between the captive vocalisations and the wild type (see figure 1). The captive birds demonstrate a vocal ‘trill’ frequently within their song which is not present in the northern population wild birds (see figure 2). As captive birds are released into the wild it provides an interesting study opportunity to observe how song patterns change into the future. Population A and B also appear to have distinct song structures. A large utility of the clustering process is that it allows for investigation into which song variations are used most frequently by a population or individual. These results are preliminary and will be investigated in more detail in future study.

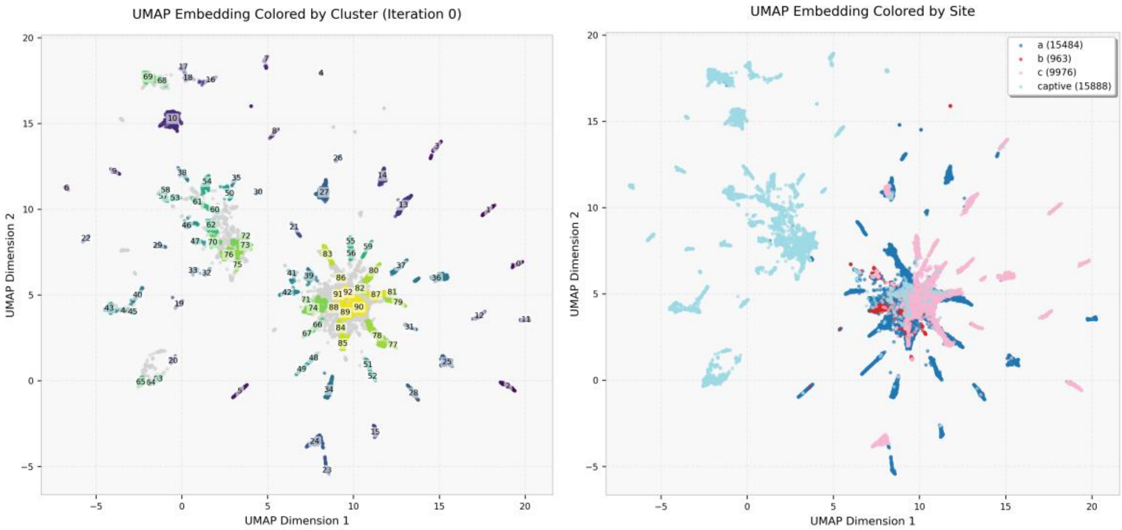

Figure 1: UMAP and HDBSCAN clustering of detection embeddings. Each plot point represents a 2.5 second vocalisation. Clusters are coloured by HDBSCAN cluster (left) or by location (right).

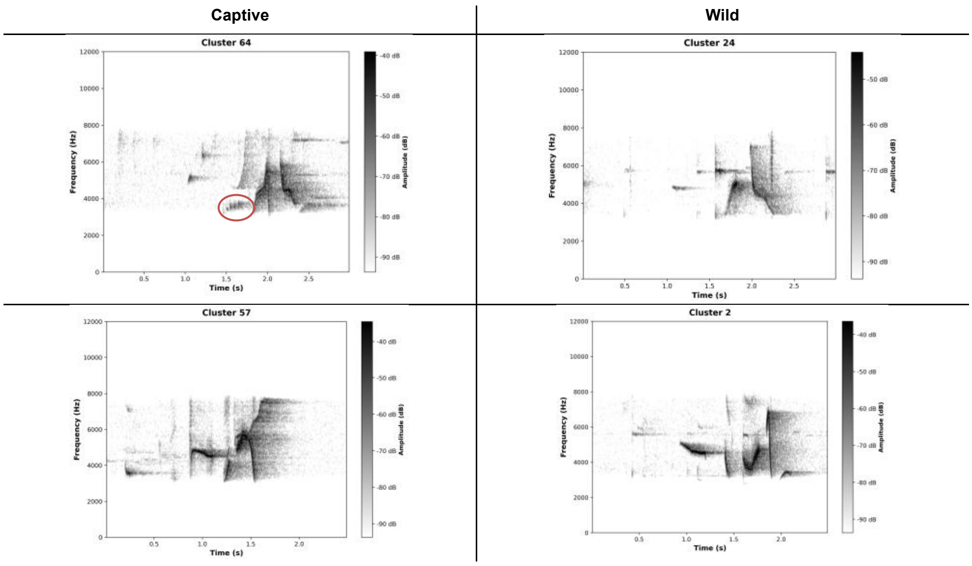

Figure 2: Example spectrograms of distinctive song types. The red circle represents a trill only observed in the captive population.

4. ACKNOWLEDGEMENTS

We wish to acknowledge the funding provided by a Queensland Government Community Sustainability Action Grant, as well as support provided by Currumbin Wildlife Sanctuary, particularly Allison Beutel and Anthony Molyneux. Thanks also the Northern Eastern Bristlebird Recovery Group, David Charley, Healthy Land and Water and BirdLife Southern Queensland for the use of their acoustic recorders and private landholders who provided land access.

5. REFERENCES

Alexander, C., Clemens, R., Roe, P., & Fuller, S. (2025). Automated note annotation after bioacoustic classification: Unsupervised clustering of extracted acoustic features improves detection of a cryptic owl. *Ecological Informatics*, 103222.

Baker, J. (1998). *Ecotones and fire and the conservation of the endangered eastern bristlebird*. University of Wollongong Wollongong. https://ro.uow.edu.au/ndownloader/files/50355888/1

Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., & White, A. E. (2022). Deep learning as a tool for ecology and evolution. *Methods in Ecology and Evolution*, *13*(8), 1640–1660. https://doi.org/10.1111/2041-210X.13901

Bravo Sanchez, F. J., English, N. B., Hossain, M. R., & Moore, S. T. (2024). Improved analysis of deep bioacoustic embeddings through dimensionality reduction and interactive visualisation. *Ecological Informatics*, *81*, 102593. https://doi.org/10.1016/j.ecoinf.2024.102593

Charley, D., Stewart, D., Stone, Z. L., Roche, K., Tasker, L., Molyneux, A., & Gillman, S. (2021). *Northern Eastern Bristlebird Dasyornis brachypterus monoides. In 'Action Plan for Australian Birds 2020'(Eds ST Garnett and GB Baker) pp. 588–591*. CSIRO Publishing: Melbourne.

Eichinski, P., Alexander, C., Roe, P., Parsons, S., & Fuller, S. (2022). A Convolutional Neural Network Bird Species Recognizer Built From Little Data by Iteratively Training, Detecting, and Labeling. *Frontiers in Ecology and Evolution*, 133.

Ghani, B., Denton, T., Kahl, S., & Klinck, H. (2023). Global birdsong embeddings enable superior transfer learning for bioacoustic classification. *Scientific Reports*, *13*(1), 22876. https://doi.org/10.1038/s41598-023-49989-z

Hill, A. P., Prince, P., Piña Covarrubias, E., Doncaster, C. P., Snaddon, J. L., & Rogers, A. (2018). AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, *9*(5), 1199–1211.

Holmes, G. (1989). Eastern Bristlebird: Species management plan for northern populations. *Draft Report to Queensland National Parks and Wildlife Service and NSW National Parks and Wildlife Service*.

Huus, J., Kelly, K. G., Bayne, E. M., & Knight, E. C. (2025). HawkEars: A regional, high-performance avian acoustic classifier. *Ecological Informatics*, *87*, 103122.

Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics*, *61*, 101236.

McGinn, K., Kahl, S., Peery, M. Z., Klinck, H., & Wood, C. M. (2023). Feature embeddings from the BirdNET algorithm provide insights into avian ecology. *Ecological Informatics*, *74*, 101995. https://doi.org/10.1016/j.ecoinf.2023.101995

Stone, Z. L., Maron, M., & Tasker, E. (2022). Reduced fire frequency over three decades hastens loss of the grassy forest habitat of an endangered songbird. *Biological Conservation*, *270*, 109570.

Stone, Z. L., Tasker, E., & Maron, M. (2018). Grassy patch size and structure are important for northern Eastern Bristlebird persistence in a dynamic ecosystem. *Emu - Austral Ornithology*, *118*(3), 269–280. https://doi.org/10.1080/01584197.2018.1425628