iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71743

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Bird Species Prediction Based on Voice Using CRNN

Ashish Kumar Pandey[1], Sanyam Kudale[2], Rohit Jadhav[3], Swaraj Giri[4], S. K. Wagh[5]
Department of Computer Engineering,Modern Education Society's Wadia College of Engineering, V. K. Joag Path, Pune, 411001,Maharashtra, India.

Abstract: Bird species recognition through sound is a crucial tool for biodiversity monitoring, enabling non-invasive, scalable insights into avian populations and their habitats. This project aims to develop a machine learning-based bird sound recognition system that identifies bird species from audio. The system processes audio inputs to extract relevant acoustic features, such as spectrograms and Mel-Frequency Cepstral Coefficients (MFCCs), leveraging tools like LibROSA. A Convolutional Recurrent Neural Network (CRNN) hybrid model is trained on these features, utilizing its ability to capture spatial patterns within audio data, for accurate bird species prediction. The proposed system also integrates a web interface, allowing users to upload recordings and view prediction results in real-time. This interface facilitates broader accessibility, making it useful for both research and citizen science initiatives.
Keywords: Mel-Frequency Cepstral Coefficients, CRNN, acoustic monitoring

I. INTRODUCTION

This document is Bird species identification through acoustic monitoring has gained significant attention in ecological research, conservation, and citizen science due to its non-invasive nature and potential for large-scale data collection. Traditional methods of bird identification, such as visual observation and manual sound analysis, are labor-intensive and often limited in scope, especially in dense or remote environments. Automatic bird sound recognition systems present an efficient alternative, capable of processing large volumes of audio data to detect and identify bird species with high accuracy.

This project develops an AI-based bird sound recognition system designed to identify bird species based on their vocalizations. The system captures and analyzes unique acoustic features like Mel-Frequency Cepstral Coefficients (MFCCs) and spectrograms to train a Convolutional Recurrent Neural Network (CRNN) hybrid model. This model combines convolutional layers, which detect spatial patterns in frequency, with recurrent layers, which capture sequential dependencies over time. The CRNN architecture is particularly effective in recognizing complex, time-varying bird calls, offering a significant improvement in accuracy over traditional models. Additionally, a user-friendly web interface enables users to upload audio recordings and receive real-time predictions with confidence scores, making it accessible to researchers, ecologists, and citizen scientists.

The project architecture emphasizes scalability and robust ness, using storage for audio files and a relational database for metadata, ensuring efficient data management and retrieval. By automating the bird species identification process, this system supports ongoing biodiversity monitoring and conservation efforts, contributing valuable data for ecological studies. Through this application, we aim to bridge the gap between complex acoustic analysis and practical ecological needs, high lighting the potential of machine learning to address pressing environmental challenges.

II. LITERATURE SURVEY

[1] The proposed SIAlex methodology employs an optimized AlexNet architecture for bird sound recognition, integrating convolutional layers with batch normalization to enhance efficiency and reduce inference time. A cascaded activation function improves nonlinearity, and a modified classifier preserves spatial integrity. By decoupling the model structure into training and inference stages through structural re parameterization, the SIAlex model balances speed and accuracy, facilitating low-cost computing and device deployment. However, evaluation metrics may be skewed when addressing rare bird species due to the uneven dataset distribution, which could bias model performance.

[2] The MV-MLC framework develops multi-view representations from augmented bird song audio, utilizing dedicated networks for time and spectrogram views. It incorporates hierarchical contrastive learning for feature extraction at various scales and temporal-spectrogram contrastive learning for cross-modality alignment, with a total loss function that optimizes model performance during pretraining.

The MV MLC framework outperformed traditional supervised methods and state-of-the-art self-supervised learning techniques in bird song recognition, demonstrating robust performance across different labeled data ratios. However, it does not account for the complexities of multiple bird species vocalizing simultaneously, and its effectiveness may depend on the quality and diversity of the training data, potentially impacting real-world performance.

[3] Proposed System outlines a method for identifying bird species through neural networks and audio processing, including spectrogram generation. The approach achieved 97% accuracy using a 4-species dataset with an 80:20 data split over 35 epochs. However, the study's use of a small dataset (400 samples) and the computational demands of AlexNet pose challenges for real-time processing on low-end devices.

[4] The Methodology employs Mel Frequency Cepstral Coefficient (MFCC) and Linear Frequency Cepstral Coefficient (LFCC) features from bird sound recordings, which are classified using a Deep Convolutional Neural Network (CNN) and a Support Vector Machine (SVM). The Deep CNN outperforms traditional methods, and the hybrid CNN-SVM approach improves classification accuracy. However, model performance heavily depends on dataset quality and size, impacting generalization.

[5] The Propose System uses machine learning techniques, including Random Decision Tree, Extra Tree Regressor, and Support Vector Machines (SVM), for bird species classification. Key steps involve Short Time Fourier Transform (STFT) for spectrogram creation, noise reduction with median filtering, and morphological operations to enhance feature detection. The Random Decision Tree classifier achieved a high AUC of 0.962 in 12-fold and 14-fold cross-validation. Background noise and overlapping sounds pose challenges, and SVM demonstrated poor performance for image-based feature classification.

[6] The Method evaluates algorithms like K-Nearest Neighbors, Random Forest, Multi-Layer Perceptron, and Naive Bayes. Support Vector Machine (SVM) with an RBF kernel was selected for its superior performance, achieving 96.7% accuracy in bird sound classification. The study highlights challenges in distinguishing closely related bird species dueto similar sound features.

[7] The Proposed System investigates self-supervised learning (SSL) for scenarios with limited labeled data, benefiting bioacoustic studies that often rely on vast unlabeled audio collections. The SupCon method achieved the highest accuracy (64.55%) in a 5-way 1-shot classification task, while Barlow Twins excelled among SSL techniques with 48.90%. Challenges include potential bias from using a pre-trained audio neural network (PANN) for segment selection and limitations of data augmentation strategies in capturing fine grained acoustic variations.

[8] Proposed Method describes using an ANN in MATLAB to classify bird species based on their sounds. Raw audio from four bird species was converted to WAV format and pre-processed using Power Spectral Density (PSD) for frequency-based waveforms. The ANN effectively identified bird species, highlighting distinct acoustic patterns in the PSD graphs. A limitation is that the system can only identify one species at a time, affecting scalability in real-world scenarios with overlapping bird sounds.

[9] The Method classifies Borneo bird species using audio signal processing. Bird sounds were collected, noise-reduced, and manually segmented with energy-based methods. Thirty-five features were extracted, and Linear Discriminant Analysis (LDA) was used for dimensionality reduction, paired with a Nearest Centroid (NC) classifier, achieving 96% accuracy. Limitations include reliance on manual segmentation, impractical for large datasets or real-time use, and the model's focus on only five species, restricting broader applicability.

[10] The Proposed System presents a bird sound recognition algorithm utilizing a Time Delay Neural Network (TDNN). Audio data from the Xeno-Canto database were preprocessed into WAV format, segmented, and analyzed using Mel Frequency Cepstral Coefficients (MFCCs). The TDNN achieved high accuracy, with 100% recall for Great Egret and Mandarin Duck. Limitations include reliance on the Xeno Canto dataset, which may limit generalization to other species or variations, and sensitivity to background noise and environ mental factors.

[11] The Proposed Methods describes a system that uses spectrography and digital signal processing on a Raspberry Pi to analyze bird sounds. Spectrograms are classified using convolutional neural networks (CNNs), achieving 70%–80% accuracy. Users can record, submit, and review predictions, with links to bird information on Wikipedia. Limitations include sensitivity to environmental noise and the need for high-quality recordings, impacting accuracy across different species and conditions.

[12] The Proposed method presents a novel approach by framing audio denoising as an image segmentation task. Audio signals are transformed into images using Short-Time Fourier Transform (STFT), enabling image processing techniques. The PtDeepLab model, leveraging a pyramid transformer and DeepLabV3+, outperforms existing denoising methods. However, its advanced design is computationally intensive, and its performance on other audio types or in extremely noisyconditions remains untested.

[13] The Proposed System outlines a method that analyzes audio by segmenting it into components based on frequency and amplitude. Hidden Markov models (HMMs) track changes in these features over time, and combining deep neural networks with HMMs (DNN-HMM) enhances accuracy.

The system excels in recognizing multiple bird species, achieving 97.3% accuracy when the number of species is known. However, accuracy decreases to 96.6% when the number of species is estimated. The system performs best with prior knowledge of the exact species count, as adding constraints in the maximum likelihood method has minimal impact on high-quality models.

[14] The Proposed System employs spatial and spectral features to enhance the separation of bird vocalizations from mixed audio using the BACPPNet model, which incorporates a polarized self-attention mechanism. This approach outperforms others, achieving a 5.38% reduction in classification accuracy compared to ground truth, with notable effectiveness for the Great Reed-Warbler. However, the model was tested on artificial mixed data, which may not fully reflect real world conditions, and some errors in sound separation remain, potentially impacting audio quality.

[15] The Method presents a lightweight model using MobileNetV3 with adjusted depth-wise separable convolutions and multi-scale feature fusion through the Pyramid Split Attention module. This approach achieves high efficiency, reaching 95.12% Top-1 and 100% Top-5 accuracy on the dataset, although it is slightly less accurate than ResNet50. While the model performs well, its ability to handle all real-world bird sound variations may be limited, particularly in noisy or diverse environments.

[16] The Method develops a low-power wireless sensor for tracking zebra finches, capable of collecting and transmitting audio and temperature data via Bluetooth Low-Energy. The sensor can operate for up to 24 hours on a single battery and effectively reduces audio interference. It is attached to zebra f inches using tape, allowing free flight during recording. While lab tests confirm strong connectivity and low power usage, the sensor's limited battery life and potential audio interference from multiple nodes pose challenges. Additionally, sound quality may be affected by compression, and size constraints may limit use with very small birds.

[17] The Proposed method presents the Anti-Harmful Birds Repelling (AHBR) method, which uses a model-free reinforcement learning approach with a LaS policy to learn how harmful birds react to threat sounds. The method aims to play these sounds in a way that minimizes adaptability in birds, achieving an average 43.5% delay in invasions compared to traditional sound repelling methods. The LaS policy assesses harmful birds' adaptability based on detection status without needing to analyze their behavior. However, the method's effectiveness depends on the variety of threat sounds used, as birds may adapt to all sounds, and reliance on sound can create noise issues.

[18] The System describes a methodology using twowireless sensing nodes (WSNs) equipped with Node MCU and various sensors to monitor environmental parameters, including temperature, humidity, NH3 and CO2 levels, and light intensity in a poultry farm. The Node MCU collects and transmits data to a Raspberry Pi 4. The model detected 473 anomalies (data points with reconstruction losses above a threshold) from a dataset of 2,000 samples, demonstrating its effectiveness for detecting abnormal data in the farm. However, it does not identify the type of anomalies (point, collective, or contextual), and its training relies on previous data.

[19] The Proposed System presents a two-tier system for bird detection in agriculture. It utilizes sensor modules with microphones and environmental sensors for local data processing via microcontrollers, while more complex tasks are handled in the cloud through 5G, where machine learning classifies bird sounds. Data management, analytics, and deter rent triggers like sound cannons are facilitated by Microsoft Azure, with user monitoring through a dashboard. This system effectively detects animals, reduces energy consumption and network load, making it suitable for agriculture. However, it relies on cloud connectivity, which may limit functionality in remote areas, and may require customization for other animals; environmental factors could also affect performance.

[20] The Proposed Method presents a system that uses frequency dynamic convolution with frequency-adaptive kernels to process spectrograms, addressing time-frequency in consistencies. It incorporates the Coordinate Attention (CA) mechanism to enhance feature extraction by retaining spatial and positional information. Built on MobileNetV3Large, the architecture adds frequency dynamic convolution modules and utilizes Hard-Swish (HS) activation for better gradient handling and computation efficiency. The proposed model demonstrates good recognition accuracy compared to other lightweight CNNs and has significantly lower parameters and computational demands than ResNet50, making it suitable for embedded devices with high energy efficiency. However, the model's inference time on low-performance embedded devices, such as Jetson Nano, is longer, indicating room for performance improvement.

III.CONCEPTUAL WORK

The conceptual framework of this bird sound recognition project builds on the principles of audio signal processing, feature extraction, and machine learning to create an auto mated system for species classification. Bird vocalizations vary widely across species, carrying unique acoustic patterns in pitch, frequency, and modulation that can be reliably detected and analyzed. By leveraging these acoustic cues, the system is structured to identify bird species in a way that mimics human auditory processing but is far more scalable and consistent.

A. Audio Preprocessing

Audio preprocessing enhances the clarity and quality of the input sound data. Raw audio signals often contain environ mental noise or overlapping sounds from other animals, which can lead to incorrect predictions if not adequately addressed. Techniques like noise reduction, normalization, and filtering are applied to refine the input, isolating the primary bird sound frequencies and improving the signal-to-noise ratio. This preprocessing step serves to optimize the audio data for further analysis, making it more amenable to the following stages of processing and classification.

B. Feature Extraction

Extracting relevant features is a fundamental step in translating audio signals into a format that the machine learning model can process effectively. Mel-Frequency Cepstral Coefficients (MFCCs) and spectrograms are employed to capture essential characteristics of the audio signal. MFCCs break down the audio into compact representations that emphasize the perceptually important frequencies, similar to how the human ear detects pitch and tone, allowing the model to distinguish between subtle differences in calls across species. Spectrograms visually represent the frequency spectrum over time, providing a powerful way to analyze time-dependent changes in pitch and energy, which are particularly useful for classifying dynamic and varied bird calls.

C. Model Training

AConvolutional Recurrent Neural Network (CRNN) hybrid model is used in this project for classification. The CRNN combines the strengths of both convolutional layers (from CNNs) and recurrent layers (like LSTM or GRU) to analyze the audio features more effectively. Convolutional layers process spatial patterns in the frequency domain, identifying unique audio textures, while recurrent layers analyze sequential patterns, capturing time-dependent changes in bird sounds. This hybrid approach allows the model to recognize intricate patterns across both time and frequency, significantly enhancing species recognition accuracy, especially for recordings with temporal dynamics.

D. Classification Layer

The final classification layer in the CRNN model outputs the predicted bird species along with a confidence score. This layer typically uses a fully connected (dense) layer followed by a softmax activation function, which converts the model’s outputs into probability scores for each bird species, allowing the model to determine the most likely species based on the input audio features.

E. Postprocessing

After obtaining the initial predictions, postprocessing techniques may be applied to improve accuracy and robustness. Thresholding can be used to discard predictions below a certain confidence level, reducing the likelihood of misclassification. Additionally, ensemble methods, where multiple models are combined to make a final prediction, can enhance accuracy by leveraging the strengths of different models or features. This approach is particularly effective in handling varied audio conditions or addressing overlapping bird calls.

F. Prediction and Output

With the model trained and deployed, the system can process new audio recordings to predict bird species and output a confidence score. This prediction output includes the identified species and confidence levels, presented through a user-friendly web interface. Such a setup provides easy access to results for researchers, ecologists, and citizen scientists, who can interact with the system in real-time, contributing to both scientific research and conservation.

G. Additional Considerations

Besides the core pipeline, the system’s design addresses general machine learning and ecological challenges, such as handling diverse audio environments and the potential for novel bird sounds. To accommodate these, periodic retraining on newly collected data can be implemented, improving robustness over time. Additionally, the system can be expanded to recognize multiple species within a single recording, adding flexibility to support complex environments where bird species are highly interdependent.

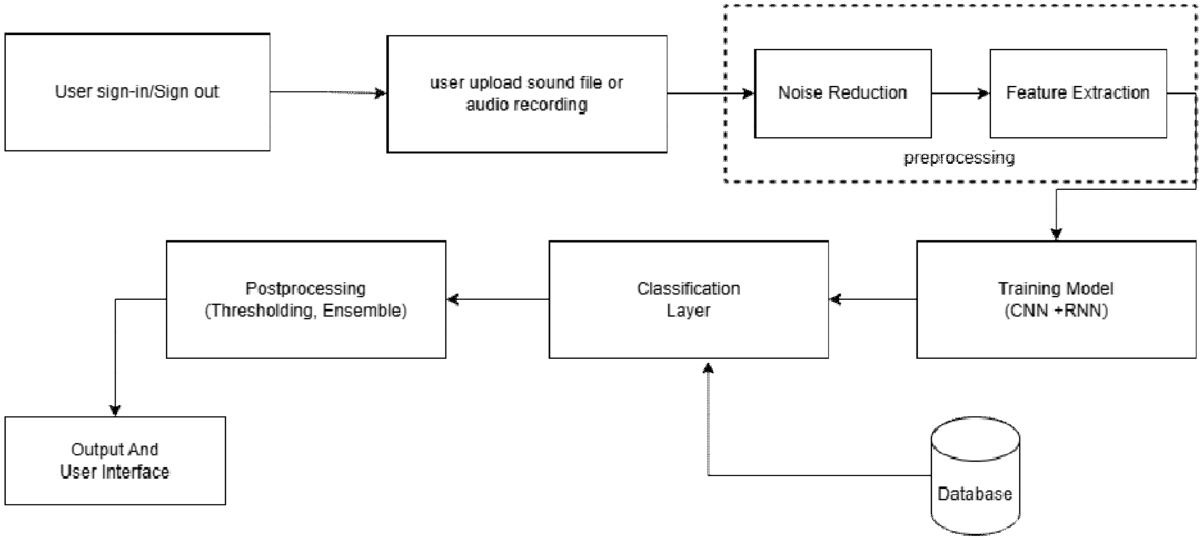

Fig. 1. System Architecture Diagram

This comprehensive approach integrates both the theoretical underpinnings of signal processing and the practical considerations of deep learning to solve the challenges of bird sound recognition. The resulting system provides a high accuracy, scalable solution that supports ecological research, conservation, and biodiversity monitoring, demonstrating the potential of AI to contribute valuable insights in environmental sciences.

IV. RESEARCH GAP

1) Limited Species Diversity in Datasets: Many studies use datasets with a limited number of bird species, often ranging from 4 to 5 species, which restricts the model's generalizability to a broader set of birds. A more extensive dataset encompassing diverse species from different geographic regions is needed to improve model robust ness and applicability in real-world scenarios.
2) Challenges in Real-Time Classification: Several studies have reported high computational costs associated with deep learning models like CNNs and AlexNet, which limit the feasibility of real-time classification on low-power devices. Addressing this limitation with lightweight architectures or optimized models suitable for real-time processing remains a critical gap.
3) Sensitivity to Background Noise and Overlapping Vocalizations: Background noise and overlapping calls from other species significantly affect model accuracy, as noted in multiple studies. Current models are often tested in controlled environments, so additional research is needed to develop noise-resistant algorithms that can perform reliably in natural, noisy environments.
4) Limited Exploration of Hybrid Models: While CNNs and SVMs have been widely used individually, few studies have explored hybrid models (e.g., CRNNs) that combine CNNsfor feature extraction and RNNs for sequential data analysis. Such hybrid architectures could enhance performance by capturing both spatial and temporal patterns, particularly useful for the dynamic characteristics of bird sounds.
5) Dependency on High-Quality Audio Data: Many models depend on clean, high-quality recordings for accurate classification. However, natural data collection often results in lower-quality recordings with various environ mental factors. Developing models that can effectively handle lower-quality data without significant performance loss is an area that needs further exploration.
6) Few-Shot Learning for Rare Species: Some studies have explored few-shot learning, but it remains underdeveloped for bird sound classification. Given the difficulty of obtaining labeled audio data for rare or endangered species, few-shot learning or self-supervised approaches could improve classification accuracy in low-data scenarios.
7) Limited Field Testing and Practical Deployment: Many studies focus on accuracy metrics in laboratory settings rather than evaluating the models in real field environments where conditions are variable. Further research is needed to validate these models in diverse field conditions and explore deployment challenges for practical applications.

V. RESULTS AND DISCUSSION

These are the implementation results based on the conceptual framework as discussed earlier:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

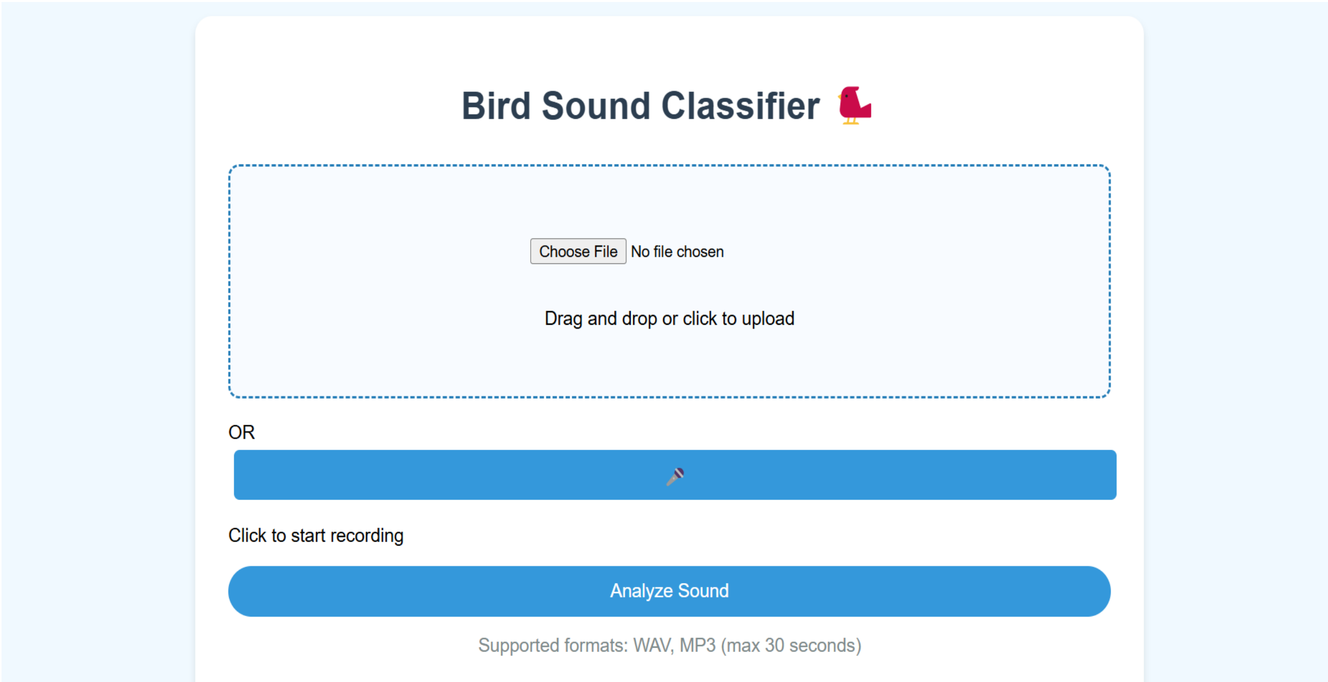

Fig. 2 User Interface

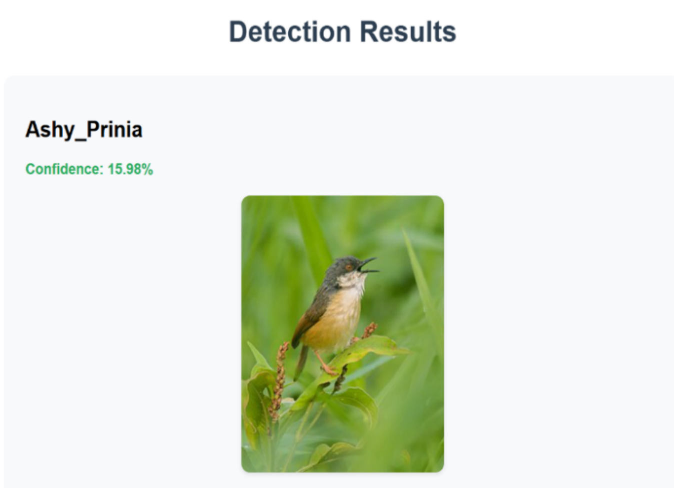

Fig.3 Result Page-i

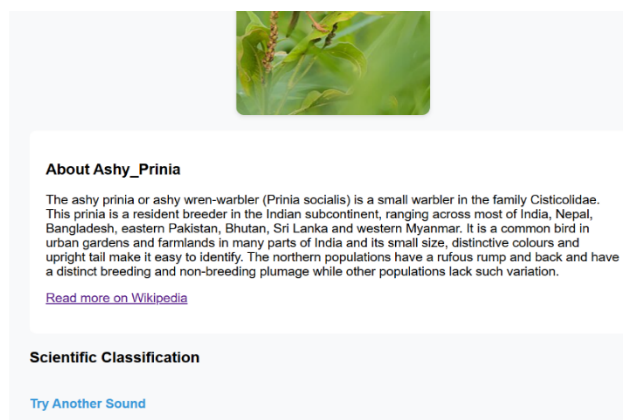

Fig.4 Result Page-ii

Following Points are based on findings in the literature andgive insights into the effectiveness of various approaches:

1) Comparative Analysis of Model Accuracy and Performance: Hybrid Models such as Convolutional Recurrent Neural Networks (CRNNs), which combine CNN layers for spatial features and recurrent layers (RNNs or LSTMs) for temporal patterns, have shown promising results in capturing both frequency and time-based features. CRNNs tend to outperform standalone CNNs and RNNs, achieving higher accuracy and robustness in noisy environments, as they handle complex bird vocalizations more effectively.
2) Feature Extraction Techniques: Spectrograms and Mel Frequency Cepstral Coefficients (MFCCs) are the most commonly used features due to their ability to represent frequency and time characteristics of bird calls. Studies indicate that MFCCs are effective in identifying general patterns but may miss finer acoustic details. Spectrograms, on the other hand, provide richer data, which CNN-based methods can leverage for improved classification accuracy.

3) Challenges with Environmental Noise and Overlapping Calls: Sensitivity to Noise: Environmental noise remains a significant limitation for many bird sound recognition systems. While preprocessing techniques like noise reduction and normalization help, studies report that most models exhibit decreased accuracy in real-world environments with background noise. Overlapping Sounds: Recognition accuracy is also affected when multiple bird species vocalize simultaneously. Although some advanced models, like CRNNs and ensemble approaches, handle overlapping calls better than basic models, achieving reliable separation and accurate classification in such cases is still challenging.
4) Generalization and Dataset Limitations: Dataset Diversity: Many studies use datasets with limited species diversity, impacting the generalizability of models. For example, models trained on data with only a few species may perform well in controlled tests but struggle in broader applications with new species or varied environmental conditions. Data Augmentation and Synthetic Data: To address data scarcity, some studies use data augmentation techniques (e.g., pitch shifting, time stretching) to artificially expand the dataset. Synthetic data helps improve model robust ness, though its effectiveness varies, especially when applied to real-world scenarios.
5) Emerging Trends and Future Directions: Studies suggest that integrating IoT devices and cloud computing could make bird sound recognition systems more accessible for remote monitoring. Real-time cloud-based processing enables complex models to be run without device limitations, though reliance on connectivity may limit functionality in remote areas.

VI. CONCLUSIONS

Bird sound recognition technology holds immense potential for advancing ecological research and conservation efforts by providing efficient, scalable tools for species identification and monitoring. This review highlights various methodologies, such as convolutional neural networks (CNNs), support vector machines (SVMs), and feature extraction techniques like Mel Frequency Cepstral Coefficients (MFCCs) and spectrogram analysis, which have shown promising results in bird species classification. However, despite the progress, several research gaps remain, including limitations in species diversity, sensitivity to background noise, and dependency on high-quality audio data.

Current models often struggle with real-time classification, generalization to noisy, natural environments, and deployment on low-power devices, posing challenges for practical field applications. Moreover, there is a need for more extensive datasets and hybrid model architectures (such as CRNNs) that can capture both spatial and temporal features in bird calls. Addressing these limitations could lead to significant advancements in model robustness, making automated bird sound recognition more accurate and accessible in real-world conditions.

In conclusion, future research should focus on creating noise-resilient, lightweight, and generalizable models, as well as exploring innovative approaches like few-shot learning for rare species. Bridging these gaps will enhance the practicality and effectiveness of bird sound recognition systems, empowering ecologists, researchers, and citizen scientists to contribute to biodiversity monitoring and conservation on a broader scale.

REFERENCES

[1] Lin Duan, Lidong Yang, Yong Guo, SIAlex: Species identifi cation and monitoring based on bird sound features, Ecological Informatics, Volume 81, 2024, 102637, ISSN 1574-9541,https://doi.org/10.1016/j.ecoinf.2024.102637.

[2] Wei Wu, Ruiyan Zhang, Xinyue Zheng, Minghui Fang, Tianyuan Ma, Qichang Hu, Xiangzeng Kong, Chen Zhao, Orchard bird song recognition based on multi-view multi-level contrastive learning, Applied Acoustics, Volume 224, 2024,110133, ISSN 0003-682X, https://doi.org/10.1016/j.apacoust.2024.110133.

[3] B. Chandu, A. Munikoti, K. S. Murthy, G. Murthy V. and C. Nagaraj, ”Automated Bird Species Identification using Audio Signal Processing and Neural Networks,” 2020 International Conference on Artificial Intelligence and Signal Processing (AISP), Amaravati, India, 2020, pp. 1-5, doi: 10.1109/AISP48273.2020.9073584.

[4] G. Rane, P. P. Rege and R. Patole, ”Bird Classification based on Bird Sounds,” 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2021, pp. 1143-1147, doi: 10.1109/SPIN52536.2021.9566071.

[5] S. Debnath, P. P. Roy, A. A. Ali and M. A. Amin, ”Identification of bird species from their singing,” 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, Bangladesh, 2016, pp. 182-186, doi: 10.1109/ICIEV.2016.7759992.

[6] Y. Jadhav, V. Patil and D. Parasar, ”Machine Learning Approach to Clas sify Birds on the Basis of Their Sound,” 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, pp. 69-73, doi: 10.1109/ICICT48043.2020.9112506. interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] I. Moummad, N. Farrugia and R. Serizel, ”Self-Supervised Learning for Few-Shot Bird Sound Classification,” 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Korea, Republic of, 2024, pp. 600-604, doi: 10.1109/ICASSPW62465.2024.10627576.

[8] M. M. M. Sukri, U. Fadlilah, S. Saon, A. K. Mahamad, M. M. Som and A. Sidek, ”Bird Sound Identification based on Artificial Neural Network,” 2020 IEEE Student Conference on Research and Development (SCOReD), Batu Pahat, Malaysia, 2020, pp. 342-345, doi: 10.1109/SCOReD50371.2020.9250746.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

[9] M. Ramashini, P. E. Abas, U. Grafe and L. C. De Silva, ”Bird Sounds Classification Using Linear Discriminant Analysis,” 2019 4th International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), Kedah, Malaysia, 2019, pp. 1 6, doi: 10.1109/ICRAIE47735.2019.9037645.

[10] M. Gou, Z. Zhao, B. Liu, Z. Huang and R. Yang, ”Bird Sounds Recognition Algorithm Using a Time Delay Neural Network,” 2023 China Automation Congress (CAC), Chongqing, China, 2023, pp. 1738 1742, doi: 10.1109/CAC59555.2023.10450312.

[11] H. B. Sale, A. P. Ghodake, R. Patil, S. Patil, S. Patil and V. E. Pawar, ”Sound-Based Bird Recognition System using Rasp berry Pi,” 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIAN CON58793.2023.10270055.

[12] J. Li, P. Wang and Y. Zhang, ”DeepLabV3+ Vision Transformer for Visual Bird Sound Denoising,” in IEEE Access, vol. 11, pp. 92540 92549, 2023, doi: 10.1109/ACCESS.2023.3294476.

[13] P. Jancovic and M. K¨ ok¨uer, ”Bird Species Recognition Using Unsupervised Modeling of Individual Vocalization Elements,” in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 5, pp. 932-947, May 2019, doi: 10.1109/TASLP.2019.2904790.

[14] J. Xie et al., ”Automatic Bird Sound Source Separation Based on Passive Acoustic Devices in Wild Environment,” in IEEE Internet of Things Journal, vol. 11, no. 9, pp. 16604-16617, 1 May1, 2024, doi: 10.1109/JIOT.2024.3354036.

[15] F. Yang, Y. Jiang and Y. Xu, ”Design of Bird Sound Recognition Model Based on Lightweight,” in IEEE Access, vol. 10, pp. 85189-85198, 2022, doi: 10.1109/ACCESS.2022.3198104.

[16] M. Magno, F. Vultier, B. Szebedy, H. Yamahachi, R. H. R. Hahnloser and L. Benini, ”A Bluetooth-Low-Energy Sensor Node for Acoustic Monitoring of Small Birds,” in IEEE Sensors Journal, vol. 20, no. 1, pp. 425-433, 1 Jan.1, 2020, doi: 10.1109/JSEN.2019.2940282.

[17] C. W. Lee et al., ”Anti-Adaptive Harmful Birds Repelling Method Based on Reinforcement Learning Approach,” in IEEE Access, vol. 9, pp. 60553-60563, 2021, doi: 10.1109/ACCESS.2021.3073205.

[18] V. Goyal, A. Yadav, S. Kumar and R. Mukherjee, ”Lightweight LAE for Anomaly Detection With Sound-Based Architecture in Smart Poultry Farm,” in IEEE Internet of Things Journal, vol. 11, no. 5, pp. 8199-8209, 1 March1, 2024, doi: 10.1109/JIOT.2023.3318298.

[19] T. Cinkler, K. Nagy, C. Simon, R. Vida and H. Rajab, ”Two-Phase Sensor Decision: Machine-Learning for Bird Sound Recognition and Vineyard Protection,” in IEEE Sensors Journal, vol. 22, no. 12, pp. 11393-11404, 15 June15, 2022, doi: 10.1109/JSEN.2021.3134817.

[20] Y.-P. Sun, Y. Jiang, Z. Wang, Y. Zhang and L.-L. Zhang, ”Wild Bird Species Identification Based on a Lightweight Model With Frequency Dynamic Convolution,” in IEEE Access, vol. 11, pp. 54352-54362, 2023, doi: 10.1109/ACCESS.2023.3281361.