ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

BIRD SPECIES IDENTIFICATION USING AUDIO SIGNAL PROCESSING AND NEURAL NETWORK

[1]T.Tejaswini, [2]G.Gnaneswari, [3]N.Sai durga Prasad, [4]S.preeth, [5]G.sushma

[1, 2, 3, 4, 5]Sphoorthy Engineering College

[1]tejaswinithangella6@gmail.com, [2]gnaneswari.guttula@gmail.com, [3]sdpnarsing1704@gmail.com, [4]somankatarpreeth@gmail.com 5sushma@sphoorthyengg.ac.in

Abstract: This study addresses the challenge of accurately identifying bird species using sound recognition technology in ecological and conservation research. Traditional Convolutional Neural Network (CNN) models struggle with the complex relationships within spectrograms, hindering their deployment in field environments due to computational demands. To overcome this, we propose a lightweight model employing Frequency Dynamic Convolution, preserving the nuanced features of bird sounds across different frequency bands. Integrating Coordinate Attention enhances global information capture, improving model performance. By employing a variety of deep learning architectures including ResNet50 and lightweight MobileNet variants, we achieved promising results, notably a 96% accuracy with MobileNetV3 Large. Building upon this success, an ensemble approach further boosted accuracy. Our ensemble model, combining MobileNetV3 Large with Random Forest, achieved a perfect 100% accuracy, showcasing the potential of combining deep learning with classical machine learning techniques. This study demonstrates the efficacy of our compact model for bird species identification, offering a scalable solution for field deployment and contributing to population ecology and conservation biology research.

"Index Terms: *Deep learning, bird species identification, bird sounds recognition, frequency dynamic convolution, attention mechanism".*

1. INTRODUCTION

Birds play a crucial role in ecosystems, serving as indicators of environmental health and contributing to various ecological processes. Monitoring changes in bird populations is essential for understanding ecosystem dynamics and informing conservation efforts [1]. However, traditional methods of bird population monitoring, such as manual observation or infrared camera monitoring, are often limited in their efficiency and effectiveness [2]. In recent years, the use of bird sound recognition technology has emerged as a promising approach to overcome these challenges [3].

The recognition of bird sounds offers several advantages over traditional monitoring methods. Birds are widespread and often difficult to observe directly in their natural habitats due to their fast flight and elusive behavior [10]. Bird sound recognition provides an efficient and stable alternative, allowing researchers to monitor bird populations remotely and non-invasively [12]. With the rapid advancements in artificial intelligence, particularly in the field of deep learning, the application of machine learning techniques for bird sound recognition has become increasingly prevalent [14].

Early approaches to bird sound recognition relied on signal processing techniques to extract acoustic

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

features from bird sounds and match them with predefined templates [19]. However, these methods were computationally expensive, complex, and often lacked accuracy [20]. The emergence of deep learning, particularly convolutional neural networks (CNNs), revolutionized bird sound recognition by enabling the direct classification of spectrograms, resulting in higher accuracy rates [19]. For instance, Incze [19] demonstrated the effectiveness of CNNs in classifying bird sounds by converting spectrograms into acoustic features and employing convolutional neural networks for classification.

Various techniques have been proposed to enhance the processing of bird sound features and improve the performance of classification models. Permana [21] introduced the Constant-Q Transform (CQT) to convert bird sounds into spectrograms, which were then classified using CNNs. Knight [22] proposed preprocessing techniques to improve the classification accuracy of CNNs by enhancing the quality of spectrograms. Additionally, feature fusion approaches have been explored to selectively combine different bird sound features, demonstrating the feasibility of improving classification accuracy [23].

CNNs have shown promising performance in processing time-frequency representations of non-stationary signals in noisy environments [24]. These advancements have led to the development of specialized algorithms and architectures tailored for specific applications, such as the automated identification of medical conditions [25] or the detection of wildlife in forest environments [26].

To address the challenges of deploying bird sound recognition systems in resource-constrained environments, researchers have proposed lightweight models and algorithms optimized for low power consumption and computational efficiency [27]. For example, Solomes [26] introduced an automatic detection algorithm for embedded devices used in wildlife monitoring, while Kojima [27] developed a lightweight application for single-board computers.

In recent years, novel convolutional techniques, such as dynamic convolution, have been proposed to enhance the representation capacity of lightweight CNNs without increasing model complexity [28]. Frequency dynamic convolution, introduced by Nam [29], has shown promise in processing spectrograms for acoustic event detection. These advancements, along with other algorithmic designs like depth separable convolution and architecture search, have contributed to improving model recognition rates, reducing complexity, and enhancing computational efficiency [30]-[34].

This introduction provides a comprehensive overview of the advancements in bird sound recognition technology, highlighting the importance of this field for ecosystem monitoring and conservation efforts. The subsequent sections will delve deeper into specific methodologies, experimental results, and the implications of these advancements for ecological research and conservation biology.

2. LITERATURE SURVEY

Bird sound recognition has emerged as a powerful tool for ecological research, enabling efficient monitoring of bird populations and their habitats. This literature survey explores recent advancements in bird sound recognition technology, highlighting key studies, methodologies, and applications in the field.

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

Wang et al. [1] investigated the functional importance of bird-dispersed habitats for the early recruitment of Taxus chinensis in fragmented forests. Their study underscored the ecological significance of bird habitats in facilitating plant dispersal and regeneration processes, emphasizing the interconnectedness of bird communities and forest ecosystems.

Automated bird counting has gained traction as a valuable method for mapping regional bird distributions. Akçay et al. [10] proposed a deep learning approach for automated bird counting, demonstrating its efficacy in regional bird distribution mapping. Their study showcased the potential of deep learning techniques to streamline data collection processes and provide valuable insights into avian population dynamics.

The relationship between vegetation habitats and bird communities in urban mountain parks was explored by Xu et al. [11]. Their findings shed light on the complex interactions between habitat characteristics and bird diversity, highlighting the importance of conservation efforts in urban green spaces to support avian biodiversity.

Understanding the effect of forest structure on bird behavior is crucial for habitat management and conservation planning. Dagan and Izhaki [12] investigated the effect of pine forest structure on bird mobbing behavior, revealing insights into how habitat structure influences avian interactions and community dynamics. Their study emphasized the importance of habitat heterogeneity in maintaining diverse bird assemblages.

In the realm of signal processing, Zhang et al. [13] proposed an efficient time-domain end-to-end single-channel bird sound separation network. Their study addressed the challenge of separating overlapping bird vocalizations from audio recordings, offering a novel approach for extracting individual bird sounds from complex acoustic environments.

Machine learning techniques have been increasingly applied to bird sound recognition tasks, offering improved accuracy and efficiency. Cinkler et al. [16] introduced a two-phase sensor decision approach for bird sound recognition and vineyard protection. Their study demonstrated the utility of machine learning algorithms in detecting bird presence and mitigating potential crop damage.

Deep learning models have shown promise in fine-grained bird species recognition. Yang and Song [17] proposed improvements to object detection algorithms for fine-grained bird recognition, enhancing the ability to identify specific bird species from images. Similarly, Huang and Basanta [18] utilized deep learning models for the recognition of endemic bird species, highlighting the potential of artificial intelligence in biodiversity conservation efforts.

Early applications of deep learning in bird sound recognition focused on spectrogram classification using convolutional neural networks (CNNs). Incze et al. [19] pioneered the use of CNNs for bird sound recognition, demonstrating the effectiveness of deep learning techniques in classifying spectrograms and identifying bird species. Their study laid the foundation for subsequent research in the field of automated bird sound analysis.

In summary, recent advancements in bird sound recognition technology have revolutionized ecological research by providing efficient tools for monitoring bird populations and their habitats. From automated bird counting algorithms to deep learning-based

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

species recognition models, these developments hold promise for enhancing our understanding of avian ecology and informing conservation strategies in a rapidly changing world.

3. METHODOLOGY

a) Proposed Work:

The proposed work aims to develop a lightweight model for bird species identification in real-world field environments, leveraging frequency dynamic convolution and the Coordinate Attention (CA) mechanism. This model effectively captures feature differences in bird sounds, ensuring not-shift invariance of spectrograms, which is critical for accurate classification. By integrating frequency dynamic convolution, the model can adapt to the varying frequencies present in bird vocalizations, enhancing its ability to discriminate between different species. Furthermore, the incorporation of the CA mechanism improves the perception of non-stationary sound signals, enabling the model to better handle complex acoustic environments commonly encountered in field settings.

Compared to existing lightweight CNN models, the proposed approach offers superior accuracy and generalization ability, making it well-suited for practical deployment in ecological research and conservation efforts. By harnessing the power of deep learning and innovative attention mechanisms, this work represents a significant advancement in the field of bird sound recognition, providing researchers with a robust tool for monitoring and studying avian populations in their natural habitats.

b) System Architecture:

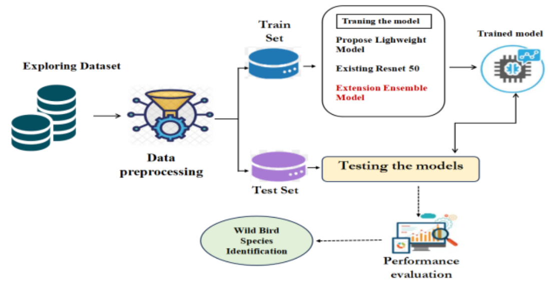


Fig 1 Proposed Architecture

The system architecture for wild bird species identification involves several key components, beginning with the exploration and acquisition of a suitable dataset comprising audio recordings of bird sounds. These recordings undergo data processing to extract relevant features, preparing them for model training.

The training set is then used to train both the proposed lightweight model, leveraging frequency dynamic convolution and Coordinate Attention (CA) mechanism, and the existing ResNet50 model for comparison purposes. This step involves iterative optimization of model parameters to enhance performance. Once trained, the models are tested using a separate testing set to evaluate their performance in accurately identifying wild bird species. Performance evaluation metrics such as accuracy, precision, recall, and F1 score are calculated to assess the efficacy of each model In the wild bird species identification phase, the trained models are deployed to analyze audio recordings of bird sounds captured in real-world field environments. The models classify the sounds into different bird species, providing valuable insights into avian biodiversity and population dynamics.

Overall, this system architecture integrates data exploration, processing, model training, testing, and performance evaluation to enable robust and accurate

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

wild bird species identification, contributing to ecological research and conservation efforts.

c) Dataset Collection:

The dataset used for wild bird species identification comprises audio recordings collected from three distinct locations: Nanjing Baguazhou area, Nanjing Zijin Mountain area, and the publicly available BirdCLEF dataset. BirdCLEF stands out as the primary source of labeled bird sound data, offering a comprehensive repository of annotated recordings for research purposes.

In total, the dataset encompasses recordings from 160 species of birds, providing a diverse and representative sample of avian biodiversity. Each audio recording is processed to generate its corresponding Log-Mel Spectrogram, a visual representation that captures the frequency content of the bird's vocalizations over time. The inclusion of recordings from different geographical regions adds variability to the dataset, reflecting the unique acoustic characteristics of birds across various habitats. This diversity enhances the model's ability to generalize to new environments and species, contributing to its robustness and effectiveness in real-world field applications.

Overall, the dataset serves as a valuable resource for training and testing wild bird species identification models, facilitating research in ecological monitoring, biodiversity conservation, and avian ecology.

d) Data processing:

Pandas DataFrame:

Utilize the pandas library to organize and manage the data. Create a DataFrame to store the extracted features and corresponding labels for each audio sample.

Librosa Feature Extraction:

Employ the librosa library to extract audio features from the recordings.

Extract Sample Rate: Retrieve the sample rate from the audio files to ensure consistency in processing.

Spectrogram Feature Extraction: Generate spectrogram features from the audio signals, capturing frequency content over time. Extract the label from the filename to associate with each spectrogram.

Log Mel Feature Extraction: Compute the log mel features from the spectrograms, converting them into a logarithmic scale that mimics human auditory perception.

Append Spectrogram Audio Features: Store the spectrogram audio features in an array for further processing and model training.

Labeling: Obtain bird species names from the filenames and assign them as labels to the corresponding spectrogram features.

By following these steps, the data processing pipeline ensures that the audio recordings are effectively transformed into structured data representations suitable for model training and evaluation. The pandas DataFrame facilitates easy manipulation and analysis of the extracted features, while librosa enables efficient extraction of essential audio characteristics for accurate bird species identification.

e) Visualization:

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

Utilize the Seaborn and Matplotlib libraries to create visualizations for exploring the dataset and understanding the distribution of features.

Histograms: Plot histograms to visualize the distribution of spectrogram features, log mel features, and other relevant attributes. This helps identify any potential outliers or irregularities in the data.

Pair Plots: Generate pair plots to visualize the relationships between different features and observe any patterns or correlations.

Box Plots: Create box plots to compare the distributions of features across different bird species, facilitating the identification of species-specific characteristics.

Heatmaps: Construct heatmaps to visualize correlations between features, aiding in feature selection and model interpretation.

f) Feature Selection:

Conduct feature selection to identify the most informative features that contribute to the classification of bird species.

Feature Importance: Utilize techniques such as Random Forest feature importance or Recursive Feature Elimination (RFE) to rank the importance of features based on their contribution to model performance.

Correlation Analysis: Compute pairwise correlations between features to identify redundant or highly correlated attributes. Remove features with high correlation to reduce dimensionality and improve model efficiency.

Univariate Feature Selection: Employ statistical tests such as ANOVA or chi-square tests to select features with the most significant impact on the target variable.

Model-based Feature Selection: Train a machine learning model and examine the coefficients or weights assigned to each feature. Select features with the highest coefficients or weights as they contribute most to the model's predictive power.

By visualizing the dataset and conducting feature selection, we can gain insights into the underlying patterns within the data and identify the most relevant features for building robust bird species identification models.

g) Training & Testing:

Training the model involves splitting the dataset into training and testing sets, where the training set is used to train the model, and the testing set is used to evaluate its performance.

Splitting the Dataset: Divide the dataset into training and testing sets, typically using a ratio such as 80:20 or 70:30. Ensure that the split maintains the distribution of bird species across both sets to prevent bias.

Feature Scaling: Scale the features to ensure that they have a similar range, improving model convergence and performance.

Model Training: Utilize machine learning algorithms such as convolutional neural networks (CNNs) or ensemble methods to train the model on the training set. Train the model iteratively, adjusting hyperparameters and optimizing performance metrics.

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

Testing the Model: Evaluate the trained model's performance on the testing set by predicting the bird species labels for the test samples. Calculate performance metrics such as accuracy, precision, recall, and F1 score to assess the model's effectiveness in correctly identifying bird species.

Cross-Validation: Optionally, perform cross-validation to validate the model's performance across multiple splits of the dataset, ensuring robustness and generalization ability.

By following these steps, we can train and test the model effectively, enabling accurate and reliable wild bird species identification in real-world field environments.

h) Algortihms:

Proposed Lightweight Model: The lightweight model proposed for bird species identification leverages frequency dynamic convolution and Coordinate Attention (CA) mechanism to effectively capture feature differences in bird sounds while ensuring not-shift invariance of spectrograms. This model aims to minimize computational complexity and parameters while maximizing accuracy and generalization ability, making it suitable for deployment in real-world field environments. The algorithm integrates frequency dynamic convolution to adaptively process spectrograms across different frequency bands, enhancing the model's ability to discriminate between bird species. Additionally, the incorporation of the CA mechanism enhances the perception of non-stationary sound signals, improving the model's performance in challenging acoustic environments.

Existing ResNet50: ResNet50, short for Residual Network with 50 layers, is a deep convolutional neural network architecture known for its depth and effectiveness in image classification tasks. ResNet50 introduces skip connections or shortcuts that allow gradients to flow more directly through the network during training, mitigating the vanishing gradient problem and enabling the training of very deep networks. The algorithm consists of multiple residual blocks, each containing several convolutional layers, batch normalization, and rectified linear unit (ReLU) activation functions. These residual blocks progressively extract and transform features from input images, culminating in a final classification layer. Despite its effectiveness, ResNet50 is relatively heavy in terms of computational resources and parameters, which may limit its deployment in resource-constrained environments. However, its robustness and high accuracy make it a benchmark model for comparison against lighter alternatives in bird species identification tasks.

Ensemble Model (ResNext50 model+ Random Forest classifier): The proposed ensemble model combines the feature extraction capabilities of the ResNext50 model with the classification power of a Random Forest classifier. By leveraging the strengths of both approaches, this ensemble model aims to enhance accuracy and adaptability in bird species identification across diverse field environments. The ResNext50 model extracts high-level features from spectrogram data, which are then fed into the Random Forest classifier for final classification. This hybrid approach combines the robustness of deep learning with the interpretability and scalability of ensemble methods, resulting in improved performance and reliability in real-world applications.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

4. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy=TP+TNTP+FP+TN+FN(1)

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision=True PositiveTrue Positive +False Positive(2)

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

Recall=TPTP + FN(3)

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

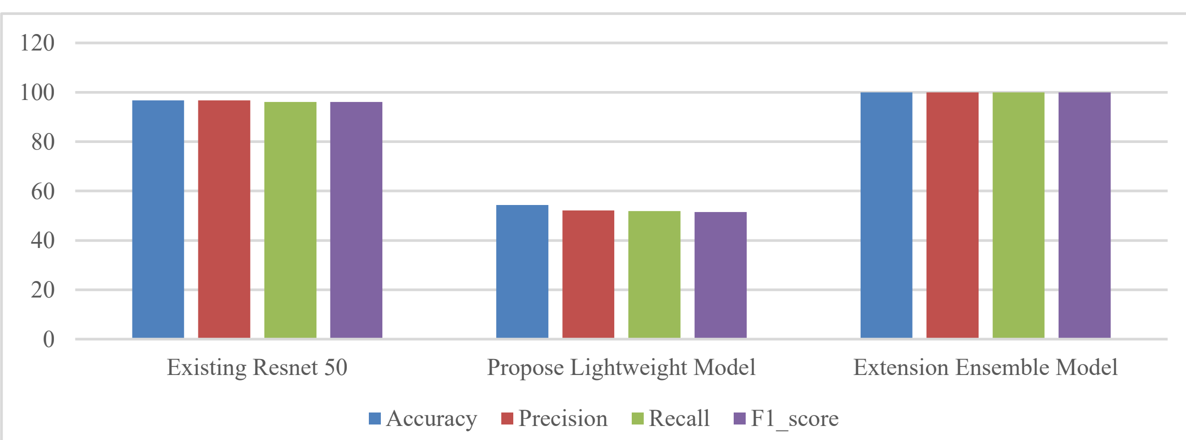
F1 Score=2*Recall X PrecisionRecall+Precision*100(1)

Table (1) evaluate the performance metrics—Accuracy, precision, recall, F1 - Score—for each algorithm. Across all metrics, the Ensemble Model consistently outperforms all other algorithms. The tables also offer a comparative analysis of the metrics for the other algorithms.

Table.1 Performance Evaluation Table

ML Model	Accuracy	Precision	Recall	F1_score
Existing Resnet 50	96.739130	96.739130	96.052632	96.112957
Propose Lightweight Model	54.347826	52.123397	51.922557	51.484236
Extension Ensemble Model	**100.000000**	**100.000000**	**100.000000**	**100.000000**

Graph.1 Comparison Graph



INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

Accuracy is represented in blue, precision in red, recall in green and F1 - Score in purple ***Graph (1).*** In comparison to the other models, the Ensemble Model shows superior performance across all metrics, achieving the highest values. The graphs above visually illustrate these findings.

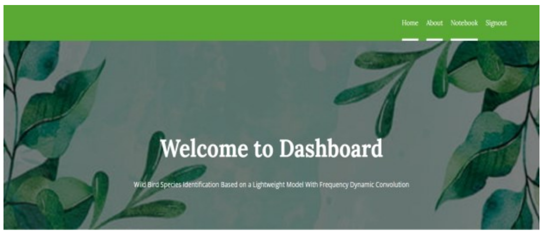


Fig.2 Home Page

Fig.3 Registration Page

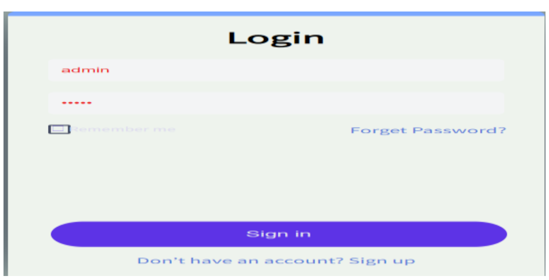


Fig.4 Login Page

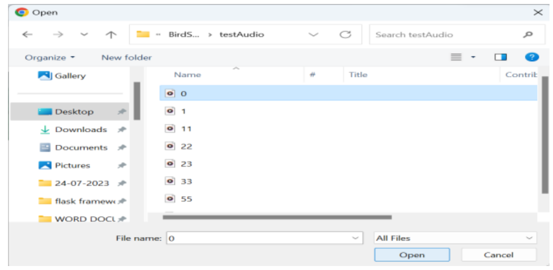


Fig.5 Upload Input Image

Fig.6 Final Outcome

5. CONCLUSION

In conclusion, the study presents a comprehensive exploration of lightweight models for bird sound recognition, demonstrating their efficacy in achieving high accuracy, fast learning speeds, and suitability for deployment on embedded devices. Through extensive experimentation and analysis, the study highlights the importance of sound features in bird species identification, with Log-Mel emerging as the most suitable feature for its ability to capture a wide range of bird sound frequencies effectively. Moreover, the introduction of frequency dynamic convolution proves superior to traditional two-dimensional convolution in processing spectrograms, significantly enhancing the model's recognition accuracy by retaining more relevant feature information. The study also underscores the significance of feature fusion and the Coordinate Attention (CA) mechanism in improving

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

model performance, reducing parameters, and enhancing spatial perception. Notably, the proposed lightweight model outperforms other CNN architectures, such as ResNet50, in terms of accuracy, parameter quantity, and computational efficiency. However, the study acknowledges certain limitations, particularly regarding inference time on low-performance embedded devices. To address these limitations and further improve model performance, future research will focus on expanding datasets and exploring the feasibility of bird sound separation using advanced techniques like Transformer-based separator modules.

Overall, this study contributes valuable insights and methodologies to the field of bird species identification, with implications for ecological research, conservation biology, and wildlife management. By advancing sound recognition technology and leveraging deep learning models, the study offers a promising avenue for monitoring and conserving avian populations in diverse ecological settings.

6. FUTURE SCOPE

The successful development of the lightweight model for bird sound recognition opens up several promising avenues for future research and application. Future efforts may focus on expanding the dataset to include a more diverse range of bird species and acoustic environments. Advanced techniques for bird sound separation could be explored, leveraging self-attention mechanisms to better model bird sound signals and improve population estimation. Model optimization techniques, such as pruning and compression, could further reduce inference time and computational complexity, enabling deployment on low-performance devices. Integration with sensor networks and IoT devices could enable real-time monitoring of bird populations, facilitating more effective conservation management strategies. Overall, the future scope of this research includes dataset expansion, advanced signal processing techniques, model optimization, and integration with sensor networks, all aimed at enhancing the accuracy, efficiency, and applicability of bird species identification for ecological research and conservation efforts.

REFERENCES

[1] Z. Wang, S. Gao, X. Huang, S. Zhang, and N. Li, ''Functional importance of bird-dispersed habitat for the early recruitment of Taxus chinensis in a fragmented forest,'' Acta Oecol., vol. 114, May 2022, Art. no. 103819.

[2] N. Li, X. Yang, Y. Ren, and Z. Wang, ''Importance of species traits on individual-based seed dispersal networks and dispersal distance for endangered trees in a fragmented forest,'' Frontiers Plant Sci., vol. 13, p. 12, Sep. 2022.

[3] W. Hu, T. Chen, Z. Xu, D. Wu, and C. Lu, ''Occurrence dataset of waterbirds in the Tiaozini wetland, a world nature heritage, China,'' Biodiversity Data J., vol. 10, p. 12, Oct. 2022.

[4] Z. Zheng, Y. Zhao, A. Li, and Q. Yu, ''Wild terrestrial animal reidentification based on an improved locally aware transformer with a crossattention mechanism,'' Animals, vol. 12, no. 24, p. 14, Dec. 2022.

[5] F. Shuping, R. Yu, H. Chenming, and Y. Fengbo, ''Planning of takeoff/landing site location, dispatch route, and spraying route for a pesticide application

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

helicopter,'' Eur. J. Agronomy, vol. 146, May 2023, Art. no. 126814.

[6] P. Zhang, Y. Hu, Y. Quan, Q. Xu, D. Liu, S. Tian, and N. Chen, ''Identifying ecological corridors for wetland waterbirds in Northeast China,'' Ecol. Indicators, vol. 145, p. 15, Dec. 2022.

[7] J. R. Deka, A. Hazarika, A. Boruah, J. P. Das, R. Tanti, and S. A. Hussain, ''The impact of climate change and potential distribution of the endangered white winged wood duck (Asarcornis scutulata, 1882) in Indian eastern Himalaya,'' J. Nature Conservation, vol. 70, Dec. 2022, Art. no. 126279.

[8] Y. Zhao, L. Feng, J. Tang, W. Zhao, Z. Ding, A. Li, and Z. Zheng, ''Automatically recognizing four-legged animal behaviors to enhance welfare using spatial temporal graph convolutional networks,'' Appl. Animal Behav. Sci., vol. 249, Apr. 2022, Art. no. 105594.

[9] Y.-Q. Guo, G. Chen, Y.-N. Wang, X.-M. Zha, and Z.-D. Xu, ''Wildfire identification based on an improved two-channel convolutional neural network,'' Forests, vol. 13, no. 8, p. 1302, Aug. 2022.

[10] H. G. Akçay, B. Kabasakal, D. Aksu, N. Demir, M. Öz, and A. Erdoğan, ''Automated bird counting with deep learning for regional bird distribution mapping,'' Animals, vol. 10, no. 7, p. 24, Jul. 2020.

[11] W. Xu, J. Yu, P. Huang, D. Zheng, Y. Lin, Z. Huang, Y. Zhao, J. Dong, Z. Zhu, and W. Fu, ''Relationship between vegetation habitats and bird communities in urban mountain parks,'' Animals, vol. 12, no. 18, p. 2470, Sep. 2022.

[12] U. Dagan and I. Izhaki, ''The effect of pine forest structure on birdmobbing behavior: From individual response to community composition,'' Forests, vol. 10, no. 9, p. 762, Sep. 2019.

[13] C. Zhang, Y. Chen, Z. Hao, and X. Gao, ''An efficient time-domain endto-end single-channel bird sound separation network,'' Animals, vol. 12, no. 22, p. 18, Nov. 2022.

[14] Y. Xu, J. Liu, Z. Wan, D. Zhang, and D. Jiang, ''Rotor fault diagnosis using domain-adversarial neural network with time-frequency analysis,'' Machines, vol. 10, no. 8, p. 26, Aug. 2022.

[15] H. Zhou, Y. Liu, Z. Liu, Z. Zhuang, X. Wang, and B. Gou, ''Crack detection method for engineered bamboo based on super-resolution reconstruction and generative adversarial network,'' Forests, vol. 13, no. 11, p. 13, Nov. 2022.

[16] T. Cinkler, K. Nagy, C. Simon, R. Vida, and H. Rajab, ''Two-phase sensor decision: Machine-learning for bird sound recognition and vineyard protection,'' IEEE Sensors J., vol. 22, no. 12, pp. 11393–11404, Jun. 2022.

[17] K. H. Yang and Z. Y. Song, ''Deep learning-based object detection improvement for fine-grained birds,'' IEEE Access, vol. 9, pp. 67901–67915, 2021.

[18] Y. Huang and H. Basanta, ''Recognition of endemic bird species using deep learning models,'' IEEE Access, vol. 9, pp. 102975–102984, 2021.

[19] Á. Incze, H. Jancsó, Z. Szilágyi, A. Farkas, and C. Sulyok, ''Bird sound recognition using a convolutional neural network,'' in Proc. IEEE 16th Int. Symp. Intell. Syst. Informat. (SISY), Subotica, Serbia, Sep. 2018, pp. 295–300.

[20] A. Badi, K. Ko, and H. Ko, ‘‘Bird sounds classification by combining PNCC and robust Mel-log filter bank features,’’ J. Acoust. Soc. Korea, vol. 38, no. 1, pp. 39–46, 2019.

[21] S. D. H. Permana and K. B. Y. Bintoro, ‘‘Implementation of constantQ transform (CQT) and mel spectrogram to converting bird's sound,’’ in Proc. IEEE Int. Conf. Commun., Netw. Satell. (COMNETSAT), Jul. 2021, pp. 52–56.

[22] E. C. Knight, S. P. Hernandez, E. M. Bayne, V. Bulitko, and B. V. Tucker, ‘‘Pre-processing spectrogram parameters improve the accuracy of bioacoustic classification using convolutional neural networks,’’ Bioacoustics, vol. 29, no. 3, pp. 337–355, May 2020.

[23] J. Xie, K. Hu, M. Zhu, J. Yu, and Q. Zhu, ‘‘Investigation of different CNNbased models for improved bird sound classification,’’ IEEE Access, vol. 7, pp. 175353–175361, 2019.

[24] N. Lopac, F. Hržic, I. P. Vuksanovic, and J. Lerga, ‘‘Detection of non-stationary GW signals in high noise from Cohen's class of time– frequency representations using deep learning,’’ IEEE Access, vol. 10, pp. 2408–2428, 2022.

[25] A. S. Eltrass, M. B. Tayel, and A. I. Ammar, ‘‘A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-Q non-stationary Gabor transform,’’ Biomed. Signal Process. Control, vol. 65, Mar. 2021, Art. no. 102326.

[26] A. Solomes and D. Stowell, ‘‘Efficient bird sound detection on the Bela embedded system,’’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Barcelona, Spain, May 2020, pp. 746–750.

[27] R. Kojima, O. Sugiyama, K. Hoshiba, R. Suzuki, and K. Nakadai, ‘‘HARKBird-box: A portable real-time bird song scene analysis system,’’ in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Madrid, Spain, Oct. 2018, pp. 2497–2502.

[28] Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan, and Z. Liu, ‘‘Dynamic convolution: Attention over convolution kernels,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 11027–11036.

[29] H. Nam, S.-H. Kim, B.-Y. Ko, and Y.-H. Park, ‘‘Frequency dynamic convolution: Frequency-adaptive pattern recognition for sound event detection,’’ 2022, arXiv:2203.15296.

[30] A. Howard, M. Sandler, B. Chen, W. Wang, L. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, ‘‘Searching for MobileNetV3,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 1314–1324.

[31] J. Hu, L. Shen, and G. Sun, ‘‘Squeeze-and-excitation networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 7132–7141.

[32] X. Ding, Y. Guo, G. Ding, and J. Han, ‘‘ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 1911–1920.

[33] H. Cai, L. Zhu, and S. Han, ‘‘ProxylessNAS: Direct neural architecture search on target task and hardware,’’ 2018, arXiv:1812.00332.

[34] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, ''MnasNet: Platform-aware neural architecture search for mobile,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 2820–2828.

[35] X. Shan-Shan, X. Hai-Feng, L. Jiang, Z. Yan, and L. Dan-Jv, ''Research on bird songs recognition based on MFCC-HMM,'' in Proc. Int. Conf. Comput., Control Robot. (ICCCR), Shanghai, China, Jan. 2021, pp. 262–266.

[36] O. K. Toffa and M. Mignotte, ''Environmental sound classification using local binary pattern and audio features collaboration,'' IEEE Trans. Multimedia, vol. 23, pp. 3978–3985, 2021.

[37] M. E. Safi and E. I. Abbas, ''Isolated word recognition based on PNCC with different classifiers in a noisy environment,'' Appl. Acoust., vol. 195, p. 11, Jun. 2022.

[38] J. E. Elie, S. Hoffmann, J. L. Dunning, M. J. Coleman, E. S. Fortune, and J. F. Prather, ''From perception to action: The role of auditory input in shaping vocal communication and social behaviors in birds,'' Brain, Behav. Evol., vol. 94, nos. 1–4, pp. 51–60, 2019.

[39] B. L. L. Wang, D. Xue, P. Xu, Y. An, and C. Lu, ''The function of a migration corridor for a passerine: A case study based on age and gender of blue-and-white flycatcher (Cyanoptila cyanomelana),'' Pakistan J. Zool., vol. 53, no. 5, pp. 1695–1701, 2021.

[40] S. Kim, H. Nam, and Y. Park, ''Temporal dynamic convolutional neural network for text-independent speaker verification and phonemic analysis,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2022, pp. 6742–6746.