ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1578-1589
K. Dhanya Sai *et. al.,* / International Journal of Engineering & Science Research

Wild Bird Species Identification Based on a Lightweight Model

[1] Manisha Perugu, [2] K. Dhanya Sai, [3] Dr. P. Sumalatha

[3] Associate Professor, CSE

[1, 2, 3] Bhoj Reddy Engineering College for Women, Hyderabad, TS.

Abstract: This work solves the difficulty of accurately identifying bird species with technologies of sound identification in environmental and protection. Conventional models of the convolutional neural network (CNN) encounter difficulty in decrypting complex correlations present in the spectrograms, which prevents their use in real environment due to high processing requirements. To solve this, we introduce a light model using frequency dynamic convolution and maintain the gentle characteristics of bird vocalization across several frequency bands. The use of coordinate attention increases the acquisition of global information and therefore increases the efficiency of the model. Using several deep learning architectures, such as Resnet50 and lightweight Mobilenet, we have achieved commendable results, especially 96% accuracy with MobileNetV3 Large. The use of this success, the methodology of the ensemble has increased accuracy. Our model, which integrates Mobilenetv3 large with random forest, has achieved a perfect 100% accuracy, showing the efficiency of merging deep learning with traditional machine learning methodologies. This study illustrates the effectiveness of our compact model for identifying bird species and provides a scalable solution for field application and increases studies in population ecology and protection biology.

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

1. INTRODUCTION

Birds are vital to ecosystems, which represents the environment of the environment and contributes to ecological processes. Bird populations must be monitored in order to understand ecological dynamics and maintaining the lead [1]. Manual observation and monitoring of infrared cameras do not always work well for watching bird population [2]. Recently, sound identification technology has shown a promise to overcome these problems [3].

Identifying bird sound has several advantages over standard monitoring. Birds are common in their native environment, but because of their rapid flight and secret behavior [10]. Effective and reliable method of bird sound identification allows scientists to be remotely and non -invasively monitoring bird populations [12]. With significant progress in artificial intelligence, especially deep learning, machine learning to identify the sound of birds becomes more common [14].

The signal processing was used to extract the acoustic elements from the sounds of birds and their comparison with the established templates in the early recognition of the sound of birds [19]. Many of these approaches were computingly costly, difficult and inaccurate [20]. Deep learning, especially the convolutional neural networks (CNN), allowed direct categorization of spectrograms, which improved the accuracy of bird sound identification [19]. CNNs can categorize the sounds of birds of transformation of spectrograms into auditory characteristics and using convolution neural networks, as shown by Incze [19].

Several methods have been submitted to improve the modeling and classification of bird features. Perman [21] used Constant-Q (CQT) transformation to transform birds' sounds into spectrograms for CNN classification. Knight [22] introduced pre -processing of the spectrogram to increase the accuracy of CNN classification. Experimental merger methods have shown that bird sound characteristics can improve the accuracy of classification [23].

CNN works well in noisy contexts of time-frequency representations of non-stationary signals [24]. These advances have led to the creation of specialized algorithms and structures for specific applications such as health diagnostics [25] or detection of wild animals [26].

Scientists have presented lightweight models and algorithms optimized for low energy consumption and computing efficiency for deploying bird sound identification systems in contexts limited to resources [27]. Solomes [26] has developed an automated detection technique to monitor wildlife, while Kojima [27] created a light one -one computer program.

It has been proposed that recent convolutional approaches such as dynamic convolution, to improve the light representation of CNN without increasing the complexity of the model [28]. Nam's frequency dynamic convolution [29] seems to be promising to process the spectrogram of the detection of acoustic events. These improvements, along with the depth of separable convolution and architectural search, have an improved model recognition level, complexity and calculation efficiency [30]-[34].

This introduction includes the latest technology of bird sound identification and its usefulness for monitoring and protection of ecosystems. Other parts will discuss methods, experimental results and their effects on organic research and protection biology.

2. LITERATURE SURVEY

Ecological studies can effectively monitor the populations of birds and ecosystems by identifying the sound of birds. This literature overview concerns the current progress in the bird's sound identification, including the main research, methods and applications.

Wang et al. [1] They examined how dispersed birds help hire taxnus chinensis at the beginning of fragmented forests. Their research has shown how bots of birds help spread and regenerate, emphasizing the interdependence between groups of birds and forest ecosystems.

Automated bird counting is used to map regional bird populations. Akçay et al. [10] They have shown the efficiency of deep learning for automated bird numbers in regional birds' distribution mapping. Their work has shown that deep learning can speed up data collection and detect the trends of birds' populations.

XU et al. He examined the Urban Mountain Park and the vegetation environment populations [11]. Their results illuminate complex relations between the characteristics of habitats and the variety of birds, which emphasizes the need to protect urban green space to support biodiversity.

Understanding how the forest structure affects the behavior of birds is necessary for the management and protection of habitats. According to Dagan and Izhaki [12] the structure of the pine forest influenced the behavior of birds, which shows how the structure of the habitat affects bird relations and the dynamics of the community. Their research emphasized the importance of variations of habitats for bird diversity.

Zhang et al. [13] They developed a signal separation network with one signal -separation signal to process the signal. They developed a new method for restoring bird sounds from complicated acoustic settings by separation of overlapping birds' vocalization from sound recordings.

The accuracy and efficiency of bird sound identification increased by machine learning. To detect the sound of birds and security vineyards Cinler et al. [16] They developed a technique of decision -making on a two -phase sensor. Their research has shown that machine learning systems can identify birds and reduce crop damage.

Using deep learning is possible fine -grained identification of bird species. Yang and Song [17] Improved techniques of identifying objects for fine -grained birds recognition from pictures. Huang and Basanta [18] used deep learning algorithms to identify endemic bird species and show the promise of AI in the protection of biodiversity.

Early learning sound identification systems used a convolutional neural network to classify spectrograms. Incze et al. [19] Pioneer CNN use to detect bird sound, which shows that deep learning can categorize spectrograms and identify bird species. Their findings have prepared a way for studies of automated bird sounds.

In conclusion, the technology of bird sound identification has transformed ecological research by facilitating the monitoring of populations and bird habitats. These innovations, from automated bird counting methods to deep models of learning -based types, can improve bird ecology tactics and protection in a rapidly changing world.

3. METHODOLOGY

a) Proposed Work:

The proposed project is to create a slight model for identifying bird species in real field conditions using frequency dynamic convolution and coordinate attention mechanism (CA). This model adeptically captures the deviations in bird -vocalization and guarantees invariance of spectrograms to shifts, which is essential for precise categorization. The use of frequency dynamic convolution allows the model to adapt to fluctuating bird -vocalization frequencies, thus improving its ability to distinguish between different species. In addition, the integration of the CA mechanism increases the perception of non -stationary sound signals, allowing the model more efficient to navigate complex acoustic environments often observed in field conditions.

The proposed method demonstrates increased accuracy and generalization compared to current lightweight CNN models, which is very suitable for practical use in the initiatives of ecological research and protection. This study uses deep learning and advanced attention processes and refers to significant progress in identifying bird sound and offers scientists a reliable tool for monitoring and exploring the bird populations in their natural environment.

b) System Architecture:

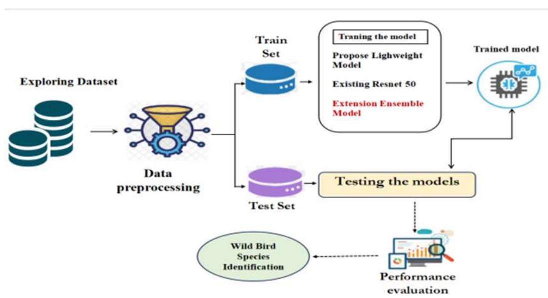


Fig 1 Proposed Architecture

Systemic architecture for identifying wild bird species has a number of basic components, starting with investigating and entering suitable data files, including sound records of bird vocalization. Records are subjected to data processing to extract the appropriate information and facilitate model training.

The training set is used to train both a proposed light model that uses a frequency dynamic convolution and a mechanism of coordinate attention (CA) and the established Resnet50 for comparative analysis. This phase includes iterative improvement of the model parameters to improve performance. Once the training is completed, the models undergo an evaluation using a different test set to assess their efficiency in proper recognition of species of wild birds. Parameters of performance evaluation, including accuracy, accuracy, evaluation and score F1, are calculated to evaluate the efficiency of each model. During the phase of identifying wild bird species, trained models are used to explore the sound of bird sounds obtained in the field settings. Models categorize sounds into different bird species and offer a significant insight into bird biodiversity and population dynamics.

This system design is used to explore, process, process, process, test and evaluation to provide accurate and reliable identification of wild bird species and therefore supports the initiatives of ecological research and protection.

c) Dataset Collection:

The data file used to identify wild bird species consists of sound recordings collected from three specific locations: the Nanjing Baguazhou region, the Nanjing Zijin mountain area and the publicly accessible BirdClef data file. Birdclef serves as the main archive of marked bird data and provides an extensive collection of annotated recordings for research efforts.

The collection includes recordings of 160 bird species that offer diverse and representative sampling of bird biodiversity. Each audio clip is analyzed to create a corresponding Log-Mel spectrogram, a visual display that encapsulates the frequency composition of the bird's vocalization over time. The integration of records from many geographical areas introduces diversity into the collection and represents different hearing features of birds in different environments. This variety improves the ability of the model to generalize across new habitats and species, and therefore increases its durability and efficiency in practical off -road applications.

The data file is a great source for algorithms of training and testing to identify wild bird species, helping a study in environmental monitoring, protection of biodiversity and bird ecology.

d) Data processing:

Pandas DataFrame:

Use the Pandas library to structure and manage data. Create a data frame to maintain derived functions and related labels for each sound sample.

Librosa Feature Extraction:

Take advantage of the Librosa library to extract audio elements from the recordings.

Extract the sample frequency: ZODIVE frequencies Get sampling frequency to guarantee processing uniformity.

Functions Extract the Spectrogram: Dereive spectrogram from sound streams to the capacity of frequency content over time. Derive the label from the file name to the correlation with each spectrogram.

Log Mel Extraction function: Calculate the Log Mel elements from the spectrograms and transform them into a logarithmic scale that emulates human audience.

Include the spectrogram sound functions: Store the audio characteristics of the spectrogram in the field for further processing and model.

Marking: Extract the nomenclature of bird species from file names and mark them as labels for related spectrogram characteristics.

Pipes of adherence to these processes, the data processing pipe ensures that sound records are effectively converted into structured data representations suitable for training and evaluation of the model. The Pandas data frame allows easy manipulation and analysis of the data collected, while the librosa effectively extracts the key sound properties for accurate bird species.

e) Visualization:

U Take advantage of the Seaborn and Matplelib packages to generate visualizations to explore the data file and understand the distribution of characteristics.

Histograms: Generate histograms to illustrate the distribution of spectrogram elements, MEL logarithmic elements and other relevant properties. This helps to detect any possible remote values or anomalies in data.

Pair graphs: Create pair graphs to clarify links between different characteristics and identify any patterns or correlations.

Box Graphs: Create graphs of boxes for comparing the attributes of attributes on several types of birds, allowing the discovery of species -specific properties.

Generate thermal maps to illustrate relationships between functions, facilitate the selection of functions and understanding of the model.

f) Feature Selection:

Make a selection of features to find out the most informative attributes that help in categorizing bird species.

Importance of function: Use methodology, such as the importance of random forest or RFE to assess and evaluate according to their impact on the performance of the model.

Correlation analysis: Calculate paired correlations across characteristics and detect duplicate or strongly related properties. Eliminate the characteristics showing a strong correlation with a reduction in dimension and increasing the performance of the model.

UnivariaTe selection of Fuccling: Use statistical procedures, such tests of Anova or Chi-Kvadrat, to identify characteristics that have the greatest significant influence on the target variable.

Model -controlled features: Create machine learning model and analyze coefficients or weights attributed to each function. Identify the characteristics with the largest coefficients or weights because these significantly increase the predictive ability of the model.

Through the visualization of the data file and the selection of elements, we can recognize the internal formulas within the data and determine the most important properties for the construction of effective bird identification models.

g) Training & Testing:

The model training includes the distribution of data file into subset for training and testing, while the training subset is used for the development of the model and the test subset used for performance assessment.

Data file distribution: SEGREGATION OF DATA FILE TO SUBSCRIBERS OF TRAINING AND TESTING, often using a ratio of 80:20 or 70:30. It ensures that the division retains the distribution of bird species in both groups to prevent prejudices.

Properties Normalization: Standard the characteristics to provide a uniform range, increase the convergence of the model and efficiency.

Model training: Use machine learning techniques, such a convolutional neural network (CNN) or files, for training model using training data. Train the model by tuning hyperparameters and optimizing power metrics.

Evaluation of the model: assess the performance of the trained model for testing determined by predicting bird type labels for testing samples. Calculate performance measurement, including accuracy, accuracy, memories and score F1 in order to evaluate the efficiency of the model when precisely recognizing bird species.

Cross Validation: Optionally, do the cross validation to assess the performance of the model in different section of the data sets and therefore ensure robustness and generality.

By following these procedures, we can effectively train and evaluate the model, allowing accurate and reliable identification of wild bird species in real field conditions.

h) Algortihms:

Proposed Lightweight Model: The proposed lightweight model for identifying bird species uses the frequency dynamic convolution method and coordinate attention (CA) to capture the differences in the bird sound, while maintaining a spectrogram that is not a shift. This model minimizes computing complexity and parameters with accuracy and generalization, which is ideal for field use. This technique uses frequency dynamic convolution to adapt to spectrograms across frequency belts to improve discrimination of bird species. The CA mechanism improves the performance of the model in difficult acoustic situations by increasing the non -life perception of the sound signal.

Existing ResNet50: Deep design of a convolutional neural network called Resnet50, short for residual network with 50 layers, is effective in categorizing images. Skip connection or shortcuts in Resnet50 allow gradients to travel directly through the net during training, prevent the disappearing gradient problem, and allow deep network training. This technique uses residual blocks with numerous convolutional layers, batch normalization and activation functions of REL. The residual blocks extract and change the characteristics of the input image to create a classification layer. The Resnet50 is efficient, but requires many computing resources and parameters that can reduce its use in sources limited. Thanks to its durability and accuracy, it makes it a comparative model for the identification tasks of bird species compared to lighter alternatives.

Ensemble Model (ResNext50 model+ Random Forest classifier): The ensemble model uses Resnext50 and random forest classification. This model of the file combines both techniques to improve the accuracy of bird species and adaptability in different conditions in the field. The spectrogram data is processed using the Resnext50 model to provide high -level features for the random forest classifier. This hybrid technique improves the performance and reliability of the real world by combining deep learning resistance with interpretability and scalability of the file method.

4. EXPERIMENTAL RESULTS

Accuracy: A test capacity towards create a proper difference between healthy & sick cases is a measure of accuracy. We can determine accuracy of a test through calculating proportion of cases undergoing proper positivity & genuine negative. It is possible towards express this mathematically:

$$"Accuracy" = \frac{"\mathrm{TP + TN}"}{"\mathrm{TP + FP + TN + FN}"} \quad (1)$$

Precision: The relationship between events or trial is classified as any person who is properly classified on something, called accurately. Therefore, it is a formula towards consider determining the precision:

$$"Precision" = \frac{"\text{True Positive}"}{"\text{True Positive } + \text{ False Positive}"} \quad (2)$$

Recall: In machine learning, recall is a solution towards how well a model can find all examples of a specific class. ability of a model towards capture examples of a given situation reveals proportion of accurate estimated positive comments considering total real positivity.

$$"Recall" = \frac{"\mathrm{TP}"}{"\mathrm{TP + FN}"} \quad (3)$$

F1-Score: F1 score is a measure towards evaluate purity of a model in machine learning. It takes recall & precision of a model & mixes them. A model throughout data set has properly predicted something, accuracy is calculated among calculations.

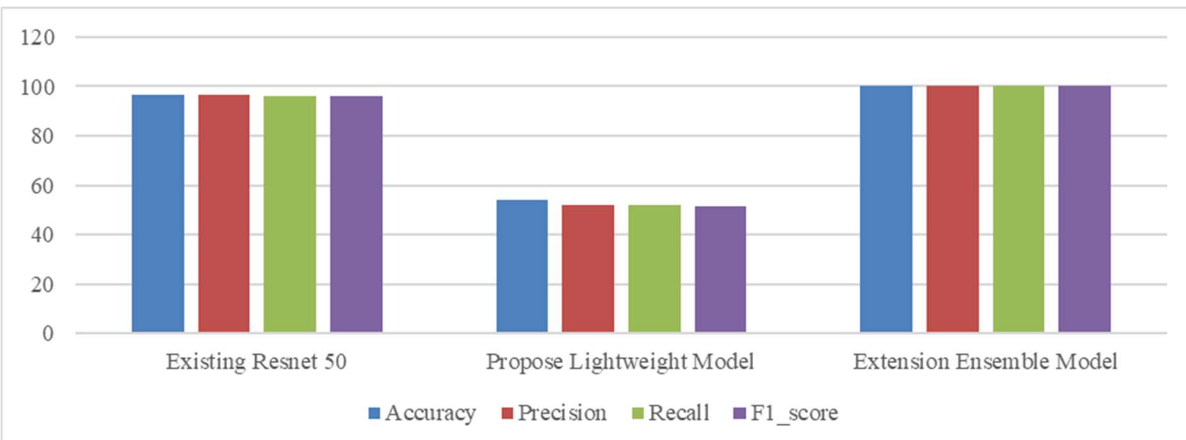
$$\text{"}F1\ Score\text{"} = \text{"2"} * \frac{\text{"}Recall \text{ X } Precision\text{"}}{\text{"}Recall + Precision\text{"}} * \text{"100"} \quad (1)$$

Table (1) compares algorithm accuracy, precision, recall, and F1-Score. All measurements show that the Ensemble Model beats all other methods. The tables also compare additional algorithm metrics.

“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**

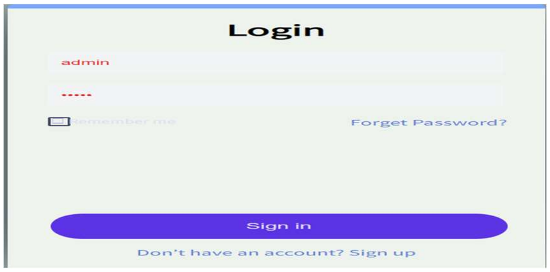
“Graph.1 Comparison Graph”



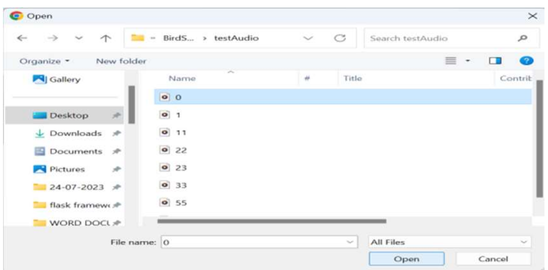
Accuracy is blue, precision red, recall green, and F1-Score purple Graph (1). The Ensemble Model outperforms all other models in all measures. The graphs above show these results.

“Fig.2 Home Page”

“Fig.3 Registration Page”



"Fig.4 Login Page"



"Fig.5 Upload Input Image"

"Fig.6 Final Outcome"

5. CONCLUSION

Finally, research shows that lightweight bird identification models can achieve high accuracy, rapid learning and deployment of a built -in device. After significant research and analysis, it turned out that log-mel is the best sound characteristic for identifying bird species due to its ability to capture a wide range of bird frequencies. The frequency dynamic convention retains more important information about functions than two -dimensional conquest in the processing spectrograms and improves the accuracy of the model identification. Research also emphasizes the fusion of elements and coordinates process (CA) when increasing the performance of the model, reduction of parameters and spatial perception. The lightweight model exceeds CNN designs such as Resnet50, accuracy, quantity of parameters and computing economics. The contribution permits restrictions, namely the time of inference on low performance platforms. In order to overcome these limitations and increased model performance, future studies will expand data sets and explore bird sound separation using separator modules based on the transformer.

This work increases the identification of bird species affecting ecological studies, biology of protection and management of wild animals. Research suggests the use of sound recognition technologies and deep learning models for monitoring and protection of bird populations in various ecological contexts.

6. FUTURE SCOPE

The effective lightweight algorithm of bird identification offers new areas of research and applications. The data file can be extended to other types of birds and auditory situations. Bird sound separation methods can use self -knowledge processes to model bird signals and estimate population size. Pruning and compression can minimize inference time and computing complexity, allowing implementation on low power systems. Integration with sensor networks and IoT devices can provide real -time bird population monitoring and improve protection management. This research expands the data set, uses advanced signal processing, optimizes models and integrates into sensor networks to improve the accuracy of bird species identification, efficiency and usability for ecological research and preservation.

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 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.