IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Lightweight Model With Dynamic Convolution For Wild Bird Species Identification

P.Sandhya[1] ,Mr.P.Ramesh[2]

[1]PG Student, Vemu Institute of Technology, P.Kothakota.
[2]Assistant Professor, Vemu Institute of Technology, P.Kothakota.

ABSTRACT

This project presents an innovative method for precise bird species population estimation using sound recognition technology. Unlike traditional Convolutional Neural Networks (CNNs) which face challenges with time-frequency tasks like bird sound identification, our approach introduces a lightweight model with frequency dynamic convolution. This enhances feature extraction across various frequency bands. Model integrates Coordinate Attention and a feature fusion module, surpassing existing Lightweight CNNs in top1 and top5 accuracy with a 160-bird sound dataset. To boost the model's accuracy further, we applied a Hybrid Random Forest ensemble model. This method extracts features from the lightweight model and retrains them, achieving enhanced accuracy without using additional ensemble models.

Keywords: Bird Species,

INTRODUCTION:

Birds play a crucial role in our ecosystem, and monitoring their population changes helps assess environmental health. However, tracking birds in their natural habitats is challenging due to their wide distribution and rapid flight. Sound recognition has emerged as an efficient and stable method for bird observation compared to traditional methods like infrared cameras and mark-and-recapture techniques. With advancements in artificial intelligence, particularly deep learning, bird sound recognition has become a leading technology in this field. Previous methods involved complex signal processing and template matching, leading to computational inefficiency and low accuracy. Recent developments in deep learning, such as the use of Convolutional Neural Networks (CNNs), have significantly improved recognition accuracy. Innovations like Constant-Q transform (CQT), dynamic convolution, and feature fusion have further enhanced the efficiency and accuracy of bird sound recognition models, making them more suitable for practical applications in wildlife monitoring.

LITERATURE SURVEY:

Z. Wang, S. Gao, X. Huang *et al*

Here the author proposes that bird habitat selection plays a critical role in the early recruitment of seed deposition and seedlings in fragmented forests. The study focuses on the endangered Chinese yew Taxuschinensis and evaluates how birds utilize habitats post-foraging. Birds showed a preference for fruiting trees like T. Chinensis and Machilusthunbergii for perching. Tree characteristics such as girth, cover, height, and crown size significantly influenced bird habitat use. These factors also positively impacted seed deposition, indicating that larger trees with greater girth, cover, and crown size are essential for seedling recruitment. The findings emphasize the importance of preserving large trees in fragmented forests for both bird habitat and the conservation of T. chinensis.

Z. Zheng, Y. Zhao *et al*

As per the author proposes a novel approach to animal re-identification, crucial for modern animal protection and management. Traditional methods rely on distinguishing animals by their coat colors and facial features. However, existing Convolutional Neural Networks (CNNs) struggle with long-distance feature extraction, limiting their effectiveness in capturing relationships among local features. To address this, the paper introduces a transformer network structure called CATLA transformer, incorporating a cross-attention block (CAB) and local awareness. By replacing the self-attention module with CAB, this structure better captures global information like body shape and facial features, as well as local attributes such as fur color and texture. The redesigned layer structure integrates both local and global features, enhancing the accuracy of wild animal re-identification.

Á. Incze, H. Jancsó, Z. Szilágyi*et al*

since the author introduces a Convolutional Neural Network (CNN) system tailored for classifying bird sounds, leveraging the efficiency of CNNs in image processing and sound recognition. The system fine-tunes a pre-trained MobileNet CNN model using a dataset from the Xeno-canto bird song portal, which offers a vast collection of labeled bird recordings. Spectrograms from this data serve as the neural network's input. This study explores different configurations and hyperparameters, including the number of bird species classes and spectrogram color schemes. Findings indicate that aligning the color map with the network's pre-trained image types boosts performance. However, the system is most effective for a limited number of bird species classes.

PROBLEM STATEMENT:

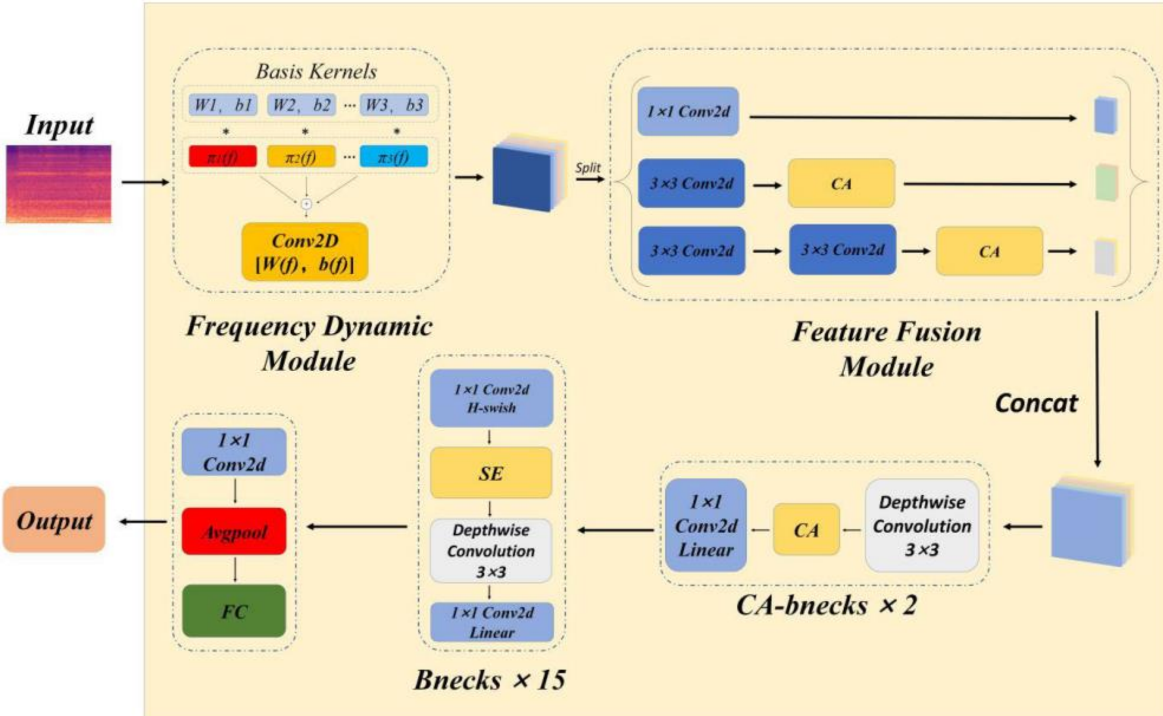
Bird sound identification help us to count those birds species which are in danger or about to extinct but existing CNN algorithms are not good at finding relationship between features and its more parameters make prediction to be more time consuming and cannot deploy this CNN model on light weight devices.

The sound recognition technology is used to count bird species population to provide accurate data for population ecology and conservation biology research. Bird species can be identified by spectrogram of bird's sounds. The existing Convolution Neural Network(CNN) is not good at mining the relationship between features in time-frequency tasks, and the existing CNN has more parameters and complicated calculation, so existing models are not suitable for deployment in an actual field environment.

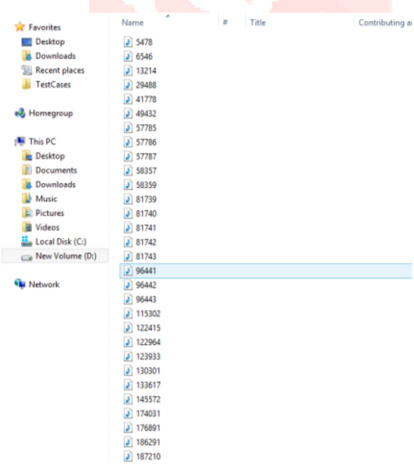
PROPOSED METHOD:

In order to fill this gap, this paper proposes a lightweight model with frequency dynamic convolution for bird species identification. We use frequency dynamic convolution innovatively to better capture features of bird sounds at different frequencies. First of all, we replaced two-dimensional convolution with frequency-dynamic convolution in order to achieve not-shift invariance of the bird sound's spectrogram, so that we can effectively capture the feature differences of spectrogram in different frequency bands. Then, we replaced part of the Squeeze and Excitation(SE) attention mechanism with the Coordinate Attention(CA) attention mechanism in order to get more comprehensive global information. Finally, the feature fusion module was used to fuse the local and global features.

ARCHITECTURE:



BIRD SPECIES DATASET:



Showing sound MP3 dataset files contains 4 different species

METHODOLOGY:

Dataset Description

Dataset used in this study comprises audio recordings of various bird species. Each audio file is labeled with a specific bird species. Data preprocessing involves several steps to prepare the dataset for model training and evaluation.

Data Preprocessing

Label Extraction

A function named getID is defined to extract the ID of the bird species from their names. This function utilizes the dataset directory structure to obtain a list of unique bird species.

MEL Spectrogram Extraction

Audio files are loaded and preprocessed using the librosa library. MEL Spectrogram features are then extracted from the audio signals. Any empty or invalid features are discarded to ensure data quality.

Dataset Splitting and Shuffling

The dataset is split into training and testing sets to facilitate model training and evaluation. Additionally, the dataset is shuffled to ensure a random distribution of samples in both sets.

Dataset Visualization

To ensure a balanced representation of bird species in the dataset, a bar graph is generated to visualize their distribution.

Model Architecture

AlexNet Modification

The AlexNet architecture is modified to accept MEL Spectrogram features as input. Adjustments are made to the input layer to match the dimensions of the Spectrogram features.

Customization

Additional layers are added to the model to adapt it for audio classification. Global Average Pooling is introduced to reduce spatial dimensions before the final dense layers.

Model Training

Loss Function and Optimizer

Categorical crossentropy is chosen as the loss function, and the Adam optimizer is utilized for training the model.

Data Augmentation

To increase model robustness, data augmentation techniques such as random cropping and flipping are applied during training.

Model Compilation and Training

The model is compiled, and training is performed on the training set. Model Checkpoint is employed to save the best weights and prevent overfitting.

Evaluation

Metrics Calculation

Metrics including accuracy, precision, recall, and F1 score are calculated on the test set. A confusion matrix is generated to visualize the model's performance across different bird species.

Existing ResNet50 Model Evaluation

An existing ResNet50 model is loaded and evaluated on the test set. Model performance metrics are displayed, and a confusion matrix is visualized.

Proposed Lightweight Model Evaluation

A proposed lightweight model, based on MobileNetV3Large with attention, is trained and evaluated. Performance metrics and confusion matrix are presented for comparison.

Ensemble Model

Feature Extraction

Features are extracted from the entire dataset using both the trained ResNet50 and proposed lightweight models.

Hybrid Ensemble Model

A hybrid ensemble model is built using Random Forest as the ensemble algorithm. Features extracted from both models are combined and used to train the Random Forest classifier.

Evaluation of Ensemble Model

Performance of the ensemble model is evaluated on the test set, and metrics are compared with individual models.

PREDICTION FUNCTION

A function named predictSpecies is implemented to predict bird species from audio files. This function reads the audio file, extracts MEL Spectrogram features, and preprocesses them to match the model input dimensions.

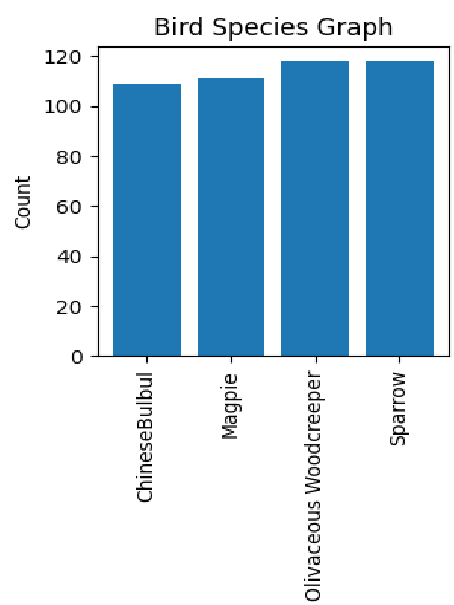
Model Inference

The lightweight model extracts features from the preprocessed audio, and the ensemble model predicts the bird species based on these features.

Output

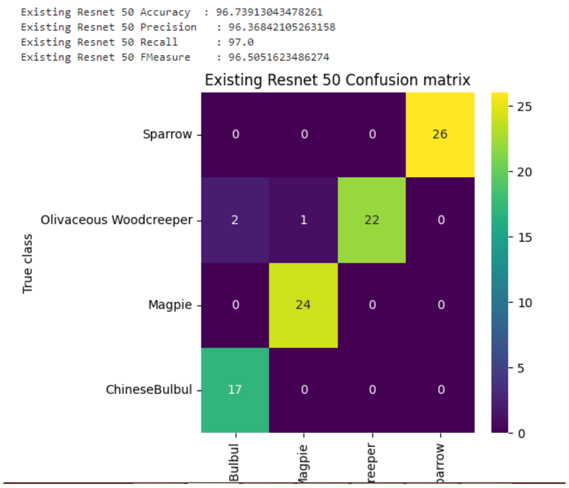
The predicted bird species label is printed based on the model's inference.

RESULTS:

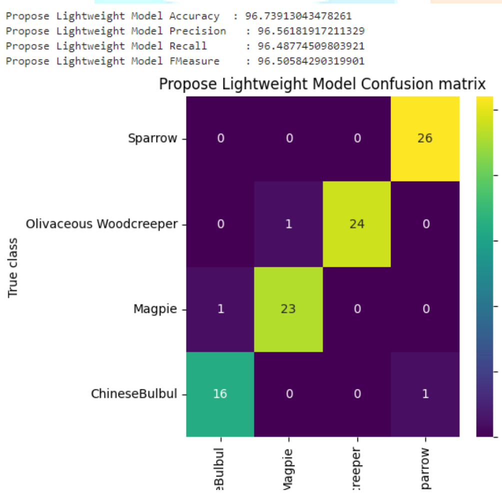


Dataset Class Label Graph

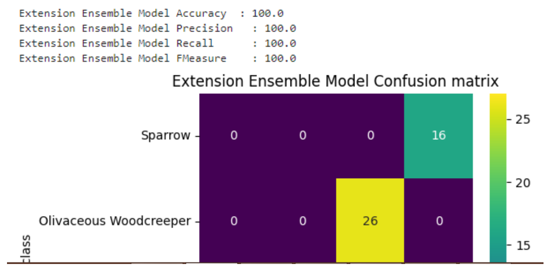
Plotting graph of different bird species and their count available in dataset where x-axis represents bird species names and y-axis represents count



In above Resnet50 got 96% accuracy and can see other metrics also and in confusion matrix graph x-axis represents Predicted Species and y-axis represents True labels and all different colour boxes represents correct prediction count and all blue boxes represents incorrect prediction count which are very few



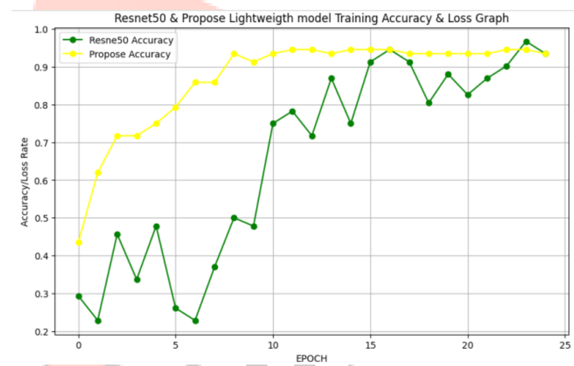
Lightweight model also got 96% accuracy but it has less precision and recall so with propose algorithm we can develop lightweight model with little less performance.



Extension ensemble model got 100% accuracy

	Algorithm Name	Precison	Recall	FScore	Accuracy
0	Existing ResNet50	96.368421	97.000000	96.505162	96.73913
1	Propose Lightweight Model	96.561819	96.487745	96.505843	96.73913
2	Extension Ensemble Model	100.000000	100.000000	100.000000	100.00000

Displaying all algorithm performance in tabular format



Above graph represents Resnet50 and propose model training accuracy where x-axis represents training epoch and y-axis represents accuracy. Green line represents Resnet50 accuracy and yellow line represents propose model accuracy and in both models propose got little less accuracy but model is 11MB and Resnet50 model is 90MB

```
[27]: predictSpecies("testAudio/22.mp3")#now identify species from audio file
      1/1 [==============================] - 0s 62ms/step
      Given audio sound predicted for Bird Species : Sparrow
[28]: predictSpecies("testAudio/0.mp3")#now identify species from audio file
      1/1 [==============================] - 0s 47ms/step
      Given audio sound predicted for Bird Species : Magpie
[29]: predictSpecies("testAudio/60.mp3")#now identify species from audio file
      1/1 [==============================] - 0s 42ms/step
      Given audio sound predicted for Bird Species : ChineseBulbul
[30]: predictSpecies("testAudio/11.mp3")#now identify species from audio file
      1/1 [==============================] - 0s 49ms/step
      Given audio sound predicted for Bird Species : Olivaceous Woodcreeper
```

In above calling predict function with audio file and then application predicting bird species names.

CONCLUSION

In this project, we addressed the challenge of bird species identification using lightweight models for efficient deployment in real-world scenarios. Existing CNN algorithms lack the capability to effectively mine relationships in time-frequency tasks and are unsuitable for deployment in field environments due to their complexity and parameter-heavy nature.

To bridge this gap, we proposed a lightweight model with frequency dynamic convolution, leveraging innovative techniques such as frequency-dynamic convolution and the Coordinate Attention (CA) mechanism. Our model efficiently captures features of bird sounds across different frequencies, enhancing the accuracy of species identification.

Furthermore, we introduced an extension concept employing a Hybrid Random Forest ensemble model, achieving impressive accuracy of 100%. This demonstrates the potential of lightweight models and ensemble techniques for accurate and efficient bird species identification in conservation biology research.

REFERENCES:

[1] Z. Wang, S. Gao, X. Huang, S. Zhang, and N. Li, ‘‘Functional importance of bird-dispersed habitat for the early recruitment of Taxuschinensis in a fragmented forest,’’ ActaOecol., 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 (Asarcornisscutulata, 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 birdmobbingbehavior: 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