International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Bird Sound Classification using Deep Neural Networks: A Comparative Analysis of State-of-the-Art Models and Custom Architectures

Shubham Revadekar[1*], Vidhish Panchal[2], Pratik Kanani[3], Kamal Shah[4], Anil Vasoya[5], Ravikumar Pandey[6]

Submitted: 07/05/2023 **Revised**: 18/07/2023 **Accepted**: 07/08/2023

Abstract: Bird sound classification plays a vital role in ecological monitoring and biodiversity conservation efforts. In this research paper, we explore the efficacy of Deep Neural Networks (DNNs) for this task, conducting a comparative analysis of five well-established methods: Xception, InceptionV3, ResNet50, EfficientNet, and VGG16. The BirdCLEF 2022 dataset, sourced from Xeno-Canto on Kaggle, serves as the foundation for our investigation. To extract essential acoustic features from the dataset, we employ Mel Frequency Cepstral Coefficients (MFCC). By converting the audio files into spectrograms, we enable the utilization of image-based classification techniques on this audio data. In addition to the state-of-the-art models, we design and implement two custom-made Convolutional Neural Network (CNN) architectures. These models surpass several existing approaches, achieving accuracy rates of 80.11% and 76.94%, respectively. Our research offers valuable insights into the performance and suitability of various DNN models for bird sound classification. Furthermore, the success of our custom architectures highlights the potential for tailored solutions in this domain. The outcomes of this study have implications for bird species identification, ecological monitoring, and wildlife conservation efforts, paving the way for further advancements in avian soundscape analysis.

Keywords: *Bird sound classification, Deep Neural Networks, Mel Frequency Cepstral Coefficients, Spectrogram, CNN, Xception, InceptionV3, ResNet50, EfficientNet, VGG16, BirdCLEF 2022, Xeno-Canto.*

1. Introduction

Bird sound classification is a prominent research area that holds significant importance in various fields, such as ecology, biodiversity monitoring, and environmental conservation. Birds communicate through a wide range of vocalizations, each carrying unique information about their behavior, species, and habitat. Manual analysis of bird sounds has been traditionally employed, but it is time-consuming, labor-intensive, and subject to human biases. In recent years, advances in machine learning, particularly deep learning, have led to the development of automated bird sound classification methods, revolutionizing the field.

Transfer learning, a subfield of machine learning, has played a crucial role in enhancing the accuracy and efficiency of bird sound classification. Transfer learning involves using pre-trained models from unrelated tasks and adapting them to the specific task of classifying bird sounds. This approach leverages the knowledge and features learned from large-scale datasets, such as ImageNet, to generalize well to bird sound analysis, even with limited labeled bird sound data. Furthermore, researchers have explored the use of mel spectrograms to represent bird vocalizations. Mel spectrograms are a specialized form of audio spectrograms that convert sound signals into visual representations, capturing the distribution of frequencies over time. By incorporating mel spectrograms into the classification process, transfer learning models can extract pertinent features from bird sound data and improve the performance of automated classification systems.

In this research paper, we explore the application of various transfer learning models along with mel spectrograms for bird sound classification. The primary objective is to evaluate the effectiveness of these models in accurately identifying different bird species and vocalizations across diverse environments. By leveraging transfer learning and Mel spectrograms, researchers can overcome data scarcity challenges and achieve robust and scalable bird sound classification models. Such advancements in automated bird sound analysis have wide-ranging implications for ecological research, conservation efforts, and understanding the dynamics of avian populations in changing ecosystems.

2. Related Work

K. W. Gunawan et al. employ an approach for transfer learning in owl audio recognition that allows them to use a

[1] *Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.*
ORCID ID : 0009-0006-4474-1806
[2] *Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.*
ORCID ID : 0009-0004-8997-3585
[3] *Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.*
ORCID ID : 0000-0002-6848-2507
[4] *Thakur College of Engineering and Technology, Mumbai, India.*
ORCID ID : 0000-0002-6369-6960
[5] *Thakur College of Engineering and Technology, Mumbai, India.*
ORCID ID : 0009-0009-6495-9142
[6] *Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.*
ORCID ID : 0009-0007-6435-2848
* *Corresponding Author Email: shubham0revadekar@gmail.com*

comparatively low size of pre-trained image classification framework that is commonly accessible. The scops owl audio dataset was obtained from the xeno-canto database. The models used are Mel's spectrogram along with MFCC, they used a NN architecture with an EfficientNet model pre-trained on the large ImageNet database and also used transfer learning. Their future aim is to analyze the trade-off between the classification performance and computational performance of each model to determine the best model because the lightweight model is also required to be deployed for a fast and automatic owl sound classifier for owl conservation purposes [1].

M. Ramashini et al. employ an SVM model for the Xeno-Canto data, extracting 3 types of cepstral features from training and testing data: GTCC, LPCC, and MFCC utilizing a 5-fold CV to train the model. Their future scope is combining GTCC with other signal features, and implementing the technique in real-time on portable multimedia devices [2]. The large Xeno- canto database has been the basis of the challenges and gives the first general insights in automated feature extraction and classification to species level for general vocalizations [3]. Next is the LifeCLEF Bird Identification Task 2016. H. Goeau et al. report the methodology of a conducted evaluation. The incorporation of soundscape sounds besides the conventional xeno-canto sounds that focus on an individual foreground species was the key novelty. The key result was that after two years of resistance from engineering-based bird song detection systems, CNN eventually outperformed them by a wide margin. It is noticeable in their results that the best-performing CNN did not use any fine-tuning so it did not benefit from the transfer learning capacities of those techniques [4].

Y. Chang et al. used distinct bird call and bird song types to increase diversity from the xeno-canto dataset and used the birdsongs for bird classification using various ML algorithms such as random forest, SVM and k-nearest neighbors, and CNN. 5 cross-validation, ReLU and sigmoid activation functions and the Adam optimizer. They would like to make possible multi-label classification problems when trying to identify multiple bird species singing at the same time. They also believe that two clips generated from the same audio with very similar patterns may fall into the training and test data sets and spuriously increase the accuracy [5]. J. Wimmer et al. proved that focused sampling approaches can offer a reliable way of analyzing huge quantities of acoustic sensor information fast and correctly by focusing on estimating a depth of a bird species by sampling ambient auditory data. To aid in the analysis of enormous amounts of acoustic sensor data, a growth of automated and semi-automatic approaches is necessary. Ultimately, analysis of large volumes of acoustic sensor data is a trade-off between analysis cost and detection accuracy [6].

C.-Y. Koh et al. focused majorly on CNN as the method for classifying bird species. They used the ResNet and the Inception Model for their CNN implementation. They draw a conclusion that it might be beneficial to learn additional information from the phase spectrogram, especially when multiple recording channels are available. Other future plans include increasing readily apparent bird voices on spectrograms of soundscape data to make them substantially comparable to the training data. Also, employing attention mechanisms can also be recommended for better performance in the future [7]. H. A. Jasim et al. use a dataset with about 8000 recordings and also use CNN as the method for bird sound classification. The DL-based method CNN using Fully Convolutional training yields superior outcomes as it removes potential future errors resulting from a lack of bird species understanding and works smoothly when programming in cohesiveness alongside the analysis of spectral kernel. The combined models were constructed by combining ML classifiers with the CNN functionality. Results show that manual feature retrieval and machine learning methods outperform baseline [8]. J. Xie et al. present a comparative analysis that uses a public dataset (CLO-43DS). The flight cries of 43 distinct North American wood warblers are included in this collection. They utilize the same DL framework with diverse inputs and they also build the fusion model using two distinct architectures. Models used include a network similar to VGG and they use a SubSpectralNet for the classification of bird recordings. Fusion strategies such as Mel-CNN, Harmonic-CNN, Percussive-CNN, and Subnet-CNN. Some of their future plans include adding more birds, another is fusing both audio and image data for classifying bird species, and also to create an efficient classification framework for recognizing bird species by intelligently fusing CNN models [9]. Another study that uses CNN to complete the assignment summarises an approach for substantial bird audio categorization in the overall setting for the LifeCLEF 2017 bird recognition test. S. Kahl et al. generated features taken from illustrations of field recordings using a range of CNNs. They tried different implementations of SOTA convolutional networks and simple CNN architectures outperformed them in certain aspects. Some of the methods suggested for future implementations for improved performance of the model are reducing dataset distortion, 3D-Convolutions, and snapshot ensembles [10].

K. Qian et al. take a different approach in the study, proposing two active learning (AL) algorithms, sparse-instance-based AL (SI-AL) and least-confidence-score-based AL (LCS-AL), both of which successfully reduce the requirement for expert human annotation. To both of these AL paradigms, a kernel-based extreme learning machine (KELM) is then integrated, and a comparison is made to the conventional support vector machine (SVM). Future work will include the comparison of more advanced AL methods via kernel-based extreme learning machines in the classification of bird sound, focusing on methods to handle such large amounts of unlabelled bird sound data [11]. Another approach used by M. Ramashini et al. is Linear Discriminant Analysis. They have demonstrated that the proposed method outperforms other complex methods such as support vector machine and K-nearest neighbor [12]. X. Ji et al. use SVM and KNN as their models and propose an improved feature selection method that reduces the feature selection time and improves the performance of the model. The dataset used by them was the CLO-43SD dataset which includes 43 bird species of North American wood-warblers [13].

L. Muller et al. used Bidirectional LSTM along with RNN for the bird sound classification. Unfortunately, their approach did not perform as well as they hoped for [14]. Y. Qiao et al. describe a seq2seq DL strategy for obtaining higher representations from bird noises without the assistance of a human. They transformed the birds' sound audio into spectrograms and then higher representations were learned by an autoencoder-based encoder-decoder scenario combined with the deep RNN. SVM, and MLP are the approaches used for the bird classification [15].

S. D. H. Permana et al. use bird sound data that was collected from the local birds in Indonesia. The CNN method is used to classify bird sounds in normal and panic conditions. The greater the epoch in training will have a positive effect on the accuracy value curve, while the greater epoch when doing the training will have a negative impact on the loss value curve [16].

F. Yang et al. provide a lightweight bird sound recognition model for using MobileNetV3 to create a feature extraction and identification network. The future work of this paper includes applying the model to embedded devices to realize real-time bird monitoring in nature reserves, collecting more bird sound data and constructing large bird datasets, simplifying bird sound feature extraction, reducing the steps and processes of feature extraction [17].

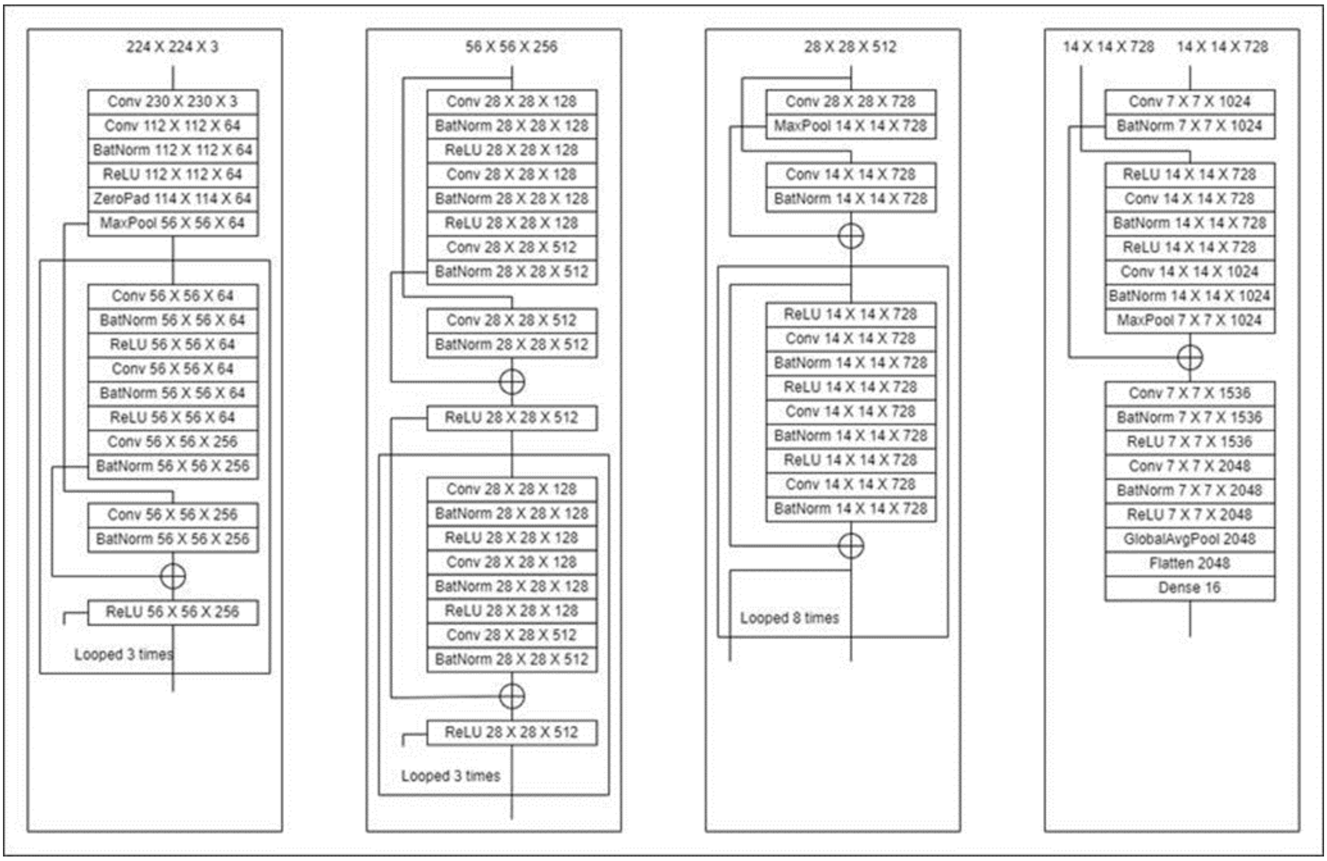


Fig. 1. Architecture Diagram for Custom Model 1

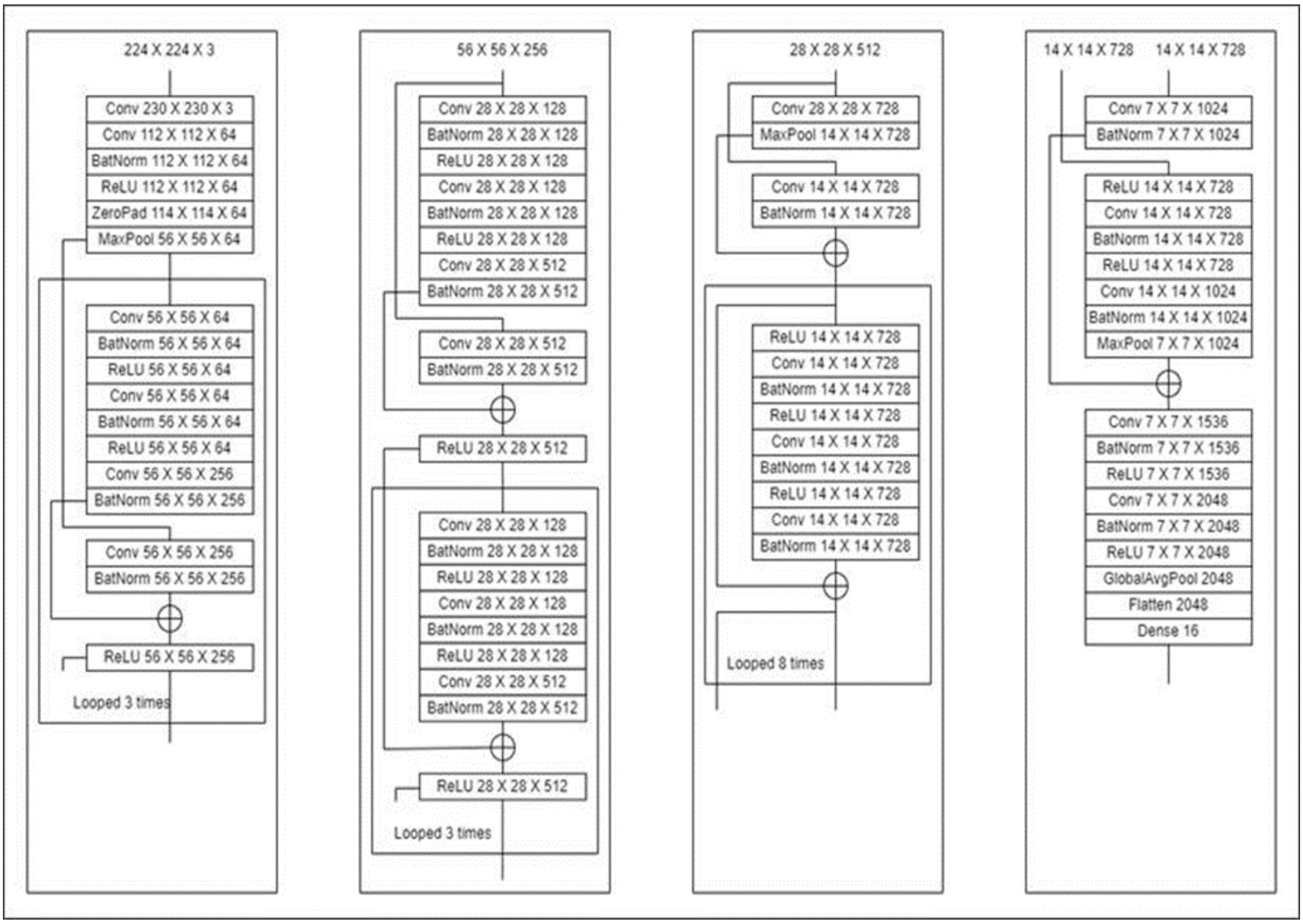


Fig. 2. Architecture Diagram for Custom Model 2

3. Proposed Methodology

We propose a solution that uses deep learning for the classification of bird sounds. We used the Transfer Learning technique for the training of deep learning models. Transfer learning is the process of employing a model that has already been trained to solve a new problem. It is presently particularly well-liked in deep learning because of its capacity to train deep neural networks with relatively minimal data. This is incredibly beneficial in tackling problems that are similar in nature because the majority of real-world scenarios frequently do not have millions of labeled data points to train such complex models. Deep learning models will be trained for the classification of bird sounds. Later, we will be analyzing the performance of these models. The proposed design is described in Figure 3.

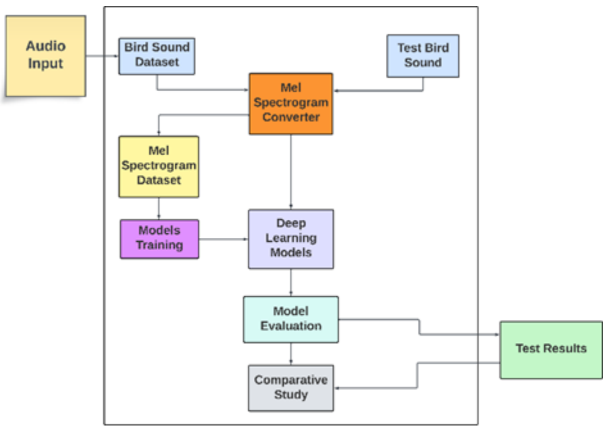


Fig. 3. Proposed Design

4. Implementation

4.1. Dataset Generation

We used the dataset from BirdCLEF 2022, which contains audio files of birds. The bulk of the training data consists of recordings of individual bird calls submitted by the xenocanto.org community. These files have been downsampled to 32 kHz and converted to the ogg format. The dataset also provided a metadata file that contains a code for the bird species, the associated audio file, and some other data. Since we will be using deep learning models, we will be converting the audio files in ogg format into spectrograms on mel scale.

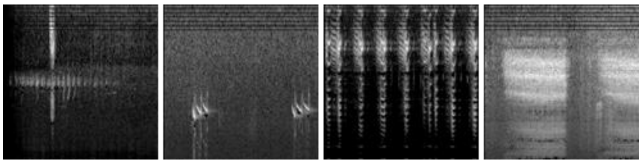


Fig. 4. Sample Mel Spectrograms of a Bird Sound Audio clips

4.2. Dataset Preprocessing

The dataset consists of various audio files of bird sounds and some metadata. This dataset contains 14,853 labeled audio files in ogg format. There are 16 different labels used in this dataset that categorize sounds into each class. The data in these audio files was converted to Mel Spectrograms. A mel spectrogram is a spectrogram where the frequencies are converted to the mel scale. For the purpose of processing the audio files and converting them into Mel Spectrograms, we used a Python library Librosa, which is used for music and audio analysis. The converted spectrograms were stored and labeled accordingly, as they will be used by deep learning networks in their training. A sample mel spectrogram of a bird sound is shown in Figure 4.

4.3. Models Used

Various models exists that are trained on ImageNet datasets, and their weights can be used, to solve different yet similar problem. The primary models we will be using are

1) Xception
2) ResNet50
3) EfficientNet
4) VGG16
5) InceptionV3

We also introduce two custom Convolutional Neural Networks that can be compared with the state-of-the-art models. In the design of its deep convolutional neural network, Xception uses Depthwise Separable Convolutions. It was created by Google researchers. Inception modules in convolutional neural networks have been interpreted by Google as a transitional process between ordinary convolution and the depthwise separable convolution operation (a depthwise convolution followed by a pointwise convolution). In this sense, a depthwise separable convolution can be thought of as an Inception module with the most towers possible. On the basis of this outcome, they propose a novel deep convolutional neural network design, with Inception modules replaced with depthwise separable convolutions [18]. An artificial intelligence model, Inception-v3 was developed to categorize and recognize items in photographs. It is a convolutional neural network that has been pre-trained and contains 48 layers. It was trained using data from the ImageNet database, where there are more than one million pictures. The network has the ability to classify photos into 1,000 distinct item types. Label Smoothing, Factorized 7 x 7 convolutions, and the use of an auxiliary classifier to convey label information lower down the network are just a few of the enhancements made by Inception-v3 [19].

Residual Network (ResNet50), another deep learning model we used, is frequently utilized in computer vision applications. It is a CNN architecture capable of supporting a high amount of convolutional layers, which could be thousands, Performance was negatively impacted by the limited number of layers that earlier CNN architectures could support. However, as more layers were added, researchers ran into the "vanishing gradient" problem [20]. Another sort of neural network architecture called EfficientNet makes use of compound scaling to improve performance. By lowering the number of parameters and FLOPs (Floating point Operations per Second), EfficientNet seeks to enhance performance while maintaining computational efficiency [21]. We have also used a VGG model, commonly known as VGGNet, which is referred to as VGG16, a CNN model of 16 layers. K. Simonyan and A. Zisserman from Oxford University came up with this model [22].

The custom models were designed with the aim of improving the accuracy of the pre-trained models. These models only take in input as the image of 224 X 224. Custom 1 model starts with Convolution layers with zero padding to standardize the input to 230 X 230 X 3. With some convolution layers, batch normalization, and Rectified Linear Unit activation layers, the output of 56 X 56 x 64 is processed with a block of layers that runs 3 times and converted to 56 X 56 X 256. Again with some convolution layers, batch normalization, and Rectified Linear Unit activation layers that run for 3 times, the output is processed to 28 X 28 X 512. Here inception type modules run 8 times. The model outputs a 7 X 7 X 2048 which can be used for classification by using a dense layer for 16 labels. The Custom 2 model varies from Custom 1 at its start but the intermediary and output layers of the model are the same. Custom 2 starts with converting the input into a convolution of 111 X 111 X 32. It is then processed through complex yet similar layer connections for extracting features. The architecture diagrams of the custom models are shown in Figures 1 and 2.

4.4. Implementation

We use the preprocessed dataset which contains spectrogram images in the model training. Our dataset was split into batches for 64 images each with the image sizes 224 X 224 X 3. With a much higher batch size, we aimed to improve the Normalization process while training the model.

We used a learning rate scheduler to fine-tune the learning rate during the training period. The learning rate was scheduled such that the learning rate will be very low at the start and linearly increase up to a maximum value and they eventually drop to a preset minimum value. The reason behind starting with a low learning rate in the epochs was to reduce the significant changes in the pre-trained weights of the models. This will help in a stable increase in the performance of the models. The schedule is shown in Figure 5.

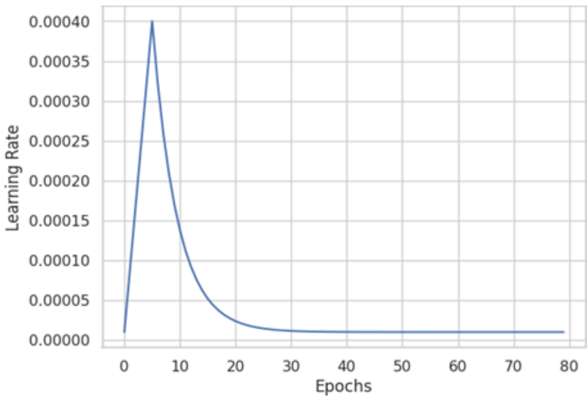


Fig. 5. Learning Rate Schedule

In the training stage, we trained five Deep Learning models: Xception, InceptionV3, EfficientNet, ResNet50, and VGG16, which were trained on the ImageNet dataset. By using these pre-trained weights, we implemented the Transfer Learning approach in classifying bird sounds. Additionally, we implemented 2 custom models (Custom1, and Custom2) to match up with these pre-trained models. The training of all these models was scheduled for 80 epochs.

5. Result and Discussion

During the training of the models, the first epoch gave a very low accuracy and did not significantly learn from the dataset. However, the following epochs showed a gradual and stable increase in accuracy. This behavior of the models was expected since our learning rate scheduler provided a very low learning rate at the start, which prevented us from breaking the pre- trained weights in the case of the primary models. Since we did not schedule the learning rates for custom models, we received fluctuating values in performance graphs. However, the custom models started showing signs of improvement later on in the training.

The training continued with the learning rate schedules that reached up to a peak, we made a curved downfall for the primary models in the study. This was done to stabilize the learning process and not induce unwanted bias. All the models were trained for 80 epochs, however, the validation accuracy stabilized around 20 epochs and an automatic early stopping was triggered. A visualization of training and testing accuracy and loss values for the pre-trained models are provided in Figure 6 and for the custom models are provided in Figure 7.

This experimentation can be used to make a comparative analysis between the performances of the models. We used validation accuracy and loss values as criteria for evaluating these models. The results of trained models are shown the Table I. At the end of the training period, we saw that among all the models we trained, Xception performed the highest with an accuracy of 80.66%. The Custom 1 performed second and was very close to outperforming the Xception with an accuracy of 80.11%. The InceptionV3 model came third with an accuracy of 79.26%. EfficientNet and ResNet50 were very close in terms of accuracy with 78.93% and 78.88% respectively. The VGG16 gave an accuracy of 77.37% and our custom 2 model came in with an accuracy of 76.94%. In terms of accuracy, the bespoke models we created nearly outperformed the state-of-the-art models.

6. Conclusion

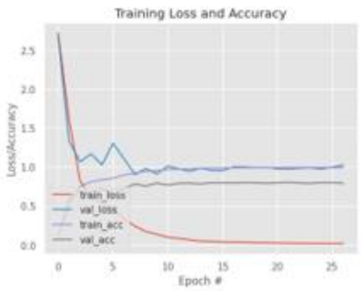
In conclusion, our research delved into bird sound classification using Deep Neural Networks (DNNs) and conducted a comparative analysis of five prominent models: Xception, InceptionV3, ResNet50, EfficientNet, and VGG16. Utilizing the BirdCLEF 2022 dataset from Kaggle, sourced from Xeno-Canto, we employed Mel Frequency Cepstral Coefficients (MFCC) for feature extraction and converted audio files into spectrograms for classification. Our findings demonstrated the commendable performance of all five DNN models in bird sound classification, with variations in accuracy and computational efficiency. Additionally, our custom-made CNN architectures surpassed the accuracy of some state-of-the-art models, highlighting the potential for domain-specific solutions.

Future research in this domain could explore various ensemble methods, advanced data augmentation, and attention mecha- nisms. Also, investigating the models' robustness to background noise or overlapping bird calls, as these challenges are common in real-world soundscapes. We would like to train a model in such a manner that it can classify multiple amounts of birds from a single audio input and also implement this technique in a real-time manner.

These endeavors will enhance the field of bird sound classification, contributing to ecological monitoring, wildlife conservation, and avian soundscape analysis.

Table 1. Performance Matrix

Model	**Accuracy**	**Loss**
Xception	80.66%	0.9776
InceptionV3	79.26%	1.0437
EfficientNet	78.93%	0.9334
ResNet50	78.88%	0.9933
VGG16	77.37%	1.1467
Custom Model 1	80.11%	0.9713
Custom Model 2	76.94%	1.0307

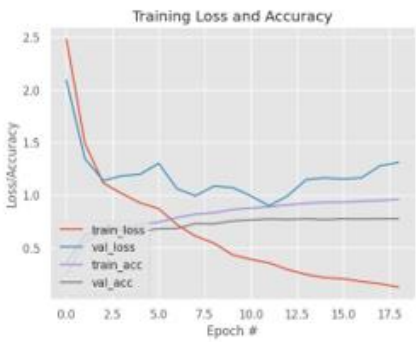


(a) Xception

(b) InceptionV3

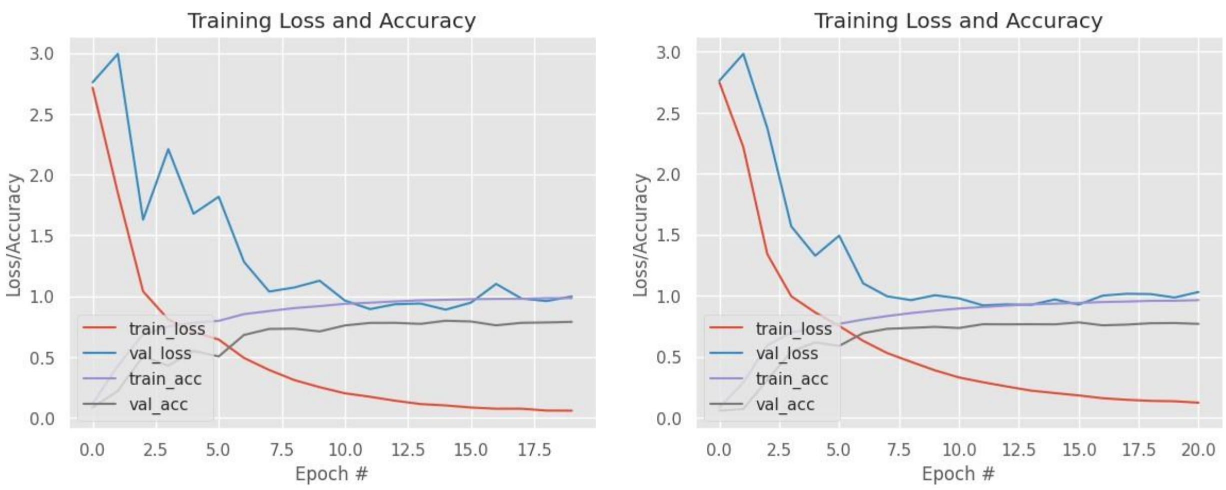
(c) EfficientNet

(d) ResNet50



(e) VGG16

Fig. 6: Accuracy and Loss Plotting for Pre-trained Models



(a) Custom Model 1

(b) Custom Model 2

Fig. 7: Accuracy and Loss Plotting for Custom

References

[1] K. W. Gunawan, A. A. Hidayat, T. W. Cenggoro, and B. Pardamean, "Repurposing transfer learning strategy of computer vision for owl sound classification," *Procedia Computer Science*, vol. 216, pp. 424–430, 2023.

[2] M. Ramashini, P. E. Abas, K. Mohanchandra, and L. C. De Silva, "Robust cepstral feature for bird sound classification," *Int. J. Electr. Comput. Eng.(2088-8708)*, vol. 12, pp. 1477–1487, 2022.

[3] W.-P. Vellinga and R. Planque´, "The xeno-canto collection and its relation to sound recognition and classification." in *CLEF (Working Notes)*, 2015.

[4] H. Goe¨au, H. Glotin, W.-P. Vellinga, R. Planque´, and A. Joly, "Lifeclef bird identification task 2016: The arrival of deep learning," in *CLEF: Conference and Labs of the Evaluation Forum*, no. 1609, 2016, pp. 440–449.

[5] Y. Chang and R. O. Sinnott, "Machine learning-based classification of birds through birdsong," *arXiv preprint arXiv:2212.04684*, 2022.

[6] J. Wimmer, M. Towsey, P. Roe, and I. Williamson, "Sampling environmental acoustic recordings to determine bird species richness," *Ecological Applications*, vol. 23, no. 6, pp. 1419–1428, 2013.

[7] C.-Y. Koh, J.-Y. Chang, C.-L. Tai, D.-Y. Huang, H.-H. Hsieh, and Y.-W. Liu, "Bird sound classification using convolutional neural networks." in *CLEF (Working Notes)*, 2019.

[8] H. A. Jasim, S. R. Ahmed, A. A. Ibrahim, and A. D. Duru, "Classify bird species audio by augment convolutional neural network," in *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. IEEE, 2022, pp. 1–6.

[9] J. Xie, K. Hu, M. Zhu, J. Yu, and Q. Zhu, "Investigation of different cnn-based models for improved bird sound classification," *IEEE Access*, vol. 7, pp. 175 353–175 361, 2019.

[10] S. Kahl, T. Wilhelm-Stein, H. Hussein, H. Klinck, D. Kowerko, M. Ritter, and M. Eibl, "Large-scale bird sound classification using convolutional neural networks." *CLEF (working notes)*, vol. 1866, 2017.

[11] K. Qian, Z. Zhang, A. Baird, and B. Schuller, "Active learning for bird sound classification via a kernel-based extreme learning machine," *The Journal of the Acoustical Society of America*, vol. 142, no. 4, pp. 1796–1804, 2017.

[12] M. Ramashini, P. E. Abas, U. Grafe, and L. C. De Silva, "Bird sounds classification using linear discriminant analysis," in *2019 4th International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*. IEEE, 2019, pp. 1–6.

[13] X. Ji, K. Jiang, and J. Xie, "Lbp-based bird sound classification using improved feature selection algorithm," *International Journal of Speech Technology*, vol. 24, pp. 1033–1045, 2021.

[14] L. Mu¨ller and M. Marti, "Bird sound classification using a bidirectional lstm." in *CLEF (Working Notes)*, 2018.

[15] Y. Qiao, K. Qian, and Z. Zhao, "Learning higher

representations from bioacoustics: A sequence-to-sequence deep learning approach for bird sound classification," in *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 18–22, 2020, Proceedings, Part V*. Springer, 2020, pp. 130–138.

[16] S. D. H. Permana, G. Saputra, B. Arifitama, W. Caesarendra, R. Rahim *et al.*, "Classification of bird sounds as an early warning method of forest fires using convolutional neural network (cnn) algorithm," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4345–4357, 2022.

[17] F. Yang, Y. Jiang, and Y. Xu, "Design of bird sound recognition model based on lightweight," *IEEE Access*, vol. 10, pp. 85 189–85 198, 2022.

[18] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[19] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[21] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[23] Kamatchi, S. B. ., Agme, V. N. ., Premkumar, S., Prasad, K. ., V, D. G. ., & Gugan, I. . (2023). Enhancing Microcomputer Edge Computing for Autonomous IoT Motion Control. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 58–67. https://doi.org/10.17762/ijritcc.v11i3.6202

[24] Rossi, G., Nowak, K., Nielsen, M., García, A., & Silva, J. Machine Learning-Based Risk Analysis in Engineering Project Management. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/114