IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research Paper , Journal Volume 13, Iss 4, 2023

DEEP LEARNING BASED APPROACH FOR BIRD SPECIES IDENTIFICATION AND CLASSIFICATION

Retesh Kumar[1], P. Sravan Kumar[2], Vadla Shiva Charan[3], Sai Shiva Vankadoth[4]

[1]Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana

[2] UG Scholar, Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana.

ABSTRACT

Ornithologists and ecologists often face challenges with manual bird species identification, which can be time-consuming and prone to errors. Traditional methods like field guides and acoustic monitoring have limitations, including subjective visual identification and limited acoustic data coverage. These methods can be difficult for non-experts and fail to provide real-time insights into bird populations. To address these issues, we introduce a state-of-the-art deep learning approach that uses neural networks to automatically identify and classify bird species based on visual and acoustic cues. Conventional methods rely on user expertise and can result in misclassifications due to variability in bird plumage and seasonal changes. Acoustic monitoring systems also require expert interpretation and may overlook visual cues. Our Deep Learning-based Approach for Bird Species Identification and Classification employs Convolutional Neural Networks (CNNs) to process visual and acoustic data. By using large, labeled datasets of bird images and audio recordings, our system can recognize and classify bird species accurately, accounting for seasonal variations. Additionally, our system supports real-time monitoring via mobile applications and field-deployable hardware, providing instant insights into bird populations. This automated approach enhances the efficiency and accuracy of bird identification, aiding in the understanding of avian ecosystems and conservation efforts.

Keywords: Bird Species Identification, Deep Learning, Convolutional Neural Networks, Visual and Acoustic Data, Real-Time Monitoring, Avian Conservation.

1. INTRODUCTION

Nowadays, Identification of bird species is a difficult activity sometimes leading to uncertainty. Birds allow us to search certain organisms within the environment as they respond quickly to changes in the atmosphere but collecting and gathering bird information requires huge efforts by humans as well as being a much more expensive method. In such situations, a robust system must be in place that will provide large-scale bird information processing and serve as a valuable resource for scholars, government agencies and so on. Consequently, naming bird species plays a significant role here for determining which species belongs to a specific image of birds. Generally; the identification of birds has done using the image, audio or video. The audio processing technique allows for the detection of birds by recording the audio signal. But the

IJFANS International Journal of Food And Nutritional Sciences
Official Publication of International Association of Food and Nutrition Scientists

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 13, Iss 4, 2023

processing of such information becomes more complicated because of in the environment; the mixed people are more effective at find images than audios or videos. So, it is preferable to use an image over audio or video to classify birds. Ornithologists have been facing problems in identifying bird species for many decades. They have to learn all the specifics of birds, such as their climate, genetics, distribution, environmental impact, etc. Normally, bird identification is conducted by an ornithologist based on the classification suggested by Linnaeus based on criteria such as State, Clade, Rank, Order, Family and Species. The rest of the paper will be arranged as below. First, brief overviews of a general introduction to the images for species and then their classification methods. Birds Voice or Videos were used in earlier technique to predict it species, but this technique have many challenges to give an accurate result due to other background of birds/animal voices. So, images can be best choice to be used to identify birds' species. To implement this technique, the images for all birds' species need to be trained to generate a model. Then deep learning algorithm will convert uploaded image into gray scale format and apply that image on train model to predict best match species name for the uploaded image. As the database of images were collected from Jordan, and the statistics number of birds in Jordan as stated in are 434 species belonging to 66 families. This study aims at investigating the use of deep learning for birds' identification using VGG-19 for extracting features from images. In order to achieve this aim, the investigation for the performance of different classifiers were performed on the following classifiers: (KNN, Decision Tree, Random Forest, and ANN) on the collected reliable database of birds images that available in Jordan. The main reason of using VGG-19 is to provide high precision by finding features with distinctive details in the image like the difference in lighting conditions and other objects surrounding the birds. Moreover, PCA could be employed as dimensionality reduction tools with these features that would help to reduce number of features that will make the training time less.

2. LITERATURE REVIEW

In a [1] We describes a convolutional neural network based deep learning approach for bird song classification that was used in an audio record-based bird identification challenge, called BirdCLEF 2016.The training and test set contained about 24k and 8.5k recordings, belonging to 999 bird species. The recorded waveforms were very diverse in terms of length and content. We converted the waveforms into frequency domain and splitted into equal segments. The segments were fed into a convolutional neural network for feature learning, which was followed by fully connected layers for classification. In the official scores our solution reached a MAP score of over 40% for main species, and MAP score of over 33% for main species mixed with background species. In a [2] In particular, we propose a nonparametric label transfer technique which transfers part constellations from objects with similar global shapes. The possibility for transferring part annotations to unseen images allows for coping with a high degree of pose and view variations in scenarios where traditional detection models fail. Our approach is especially valuable for fine-grained recognition scenarios where intra class variations are extremely high, and precisely localized features need to be extracted. Furthermore, we show the importance of carefully designed visual extraction strategies, such as combination of complementary feature types and iterative image segmentation, and the resulting impact on the recognition performance. In experiments, our simple yet powerful approach achieves 35.9% and 57.8% accuracy on the CUB- 2010 and 2011 bird datasets, which is the current best performance for these benchmarks. In a [3] Current human-in-the-loop fine-grained visual categorization systems depend on a predefined vocabulary of attributes and parts, usually determined by experts. In this work, we move away from that expert-driven and attributecentric paradigm and present a novel interactive classification system that incorporates computer vision and perceptual similarity metrics in a unified framework. At test time, users are asked to judge relative similarity between a query image and various sets of images; these general queries do not require expert-defined terminology and are applicable to other domains and basic-level categories, enabling a flexible, efficient, and scalable system for finegrained categorization with humans in the loop.Our system outperforms existing state-of-the-art systems for relevance feedback-based image retrieval as well as interactive classification, resulting in a reduction of up to 43% in the average number of questions needed to correctly classify an image.

In a [4] We present a new audio classification method for bird species identification. Whereas most approaches apply nearest neighbour matching or decision trees using extracted templates for each bird species, ours draws upon techniques from speech recognition and recent advances in the domain of deep learning.With novel preprocessing and data augmentation methods, we train a convolutional neural network on the biggest publicly available dataset. Our network architecture achieves a mean average precision score of 0.686 when predicting the main species of each sound file and scores 0.555 when background species are used as additional prediction targets. As this performance surpasses current state of the art results, our approach won this years international BirdCLEF 2016 Recognition Challenge. In a [5] Automatic identification of bird species by their vocalization is studied in this paper. Bird sounds are represented with two different parametric representations: (i) the mel-cepstrum parameters and (ii) a set of low-level signal parameters, both of which have been found useful for bird species recognition. Recognition is performed in a decision tree with support vector machine (SVM) classifiers at each node that perform classification between two species. Recognition is tested with two sets of bird species whose recognition has been previously tested with alternative methods. Recognition results with the proposed method suggest better or equal performance when compared to existing reference methods

In a [6] Traditional methods of computer vision and machine learning cannot match human performance on tasks such as the recognition of handwritten digits or traffic signs. Our biologically plausible deep artificial neural network architectures can. Small (often minimal) receptive fields of convolutional winnertake-all neurons yield large network depth, resulting in roughly as many sparsely connected neural layers as found in mammals between retina and visual cortex. Only winner neurons are trained. Several deep neural columns become experts on inputs preprocessed in different ways; their predictions are averaged. Graphics cards allow for fast training. On the very competitive MNIST handwriting benchmark, our method is the first to achieve near-human performance. On a traffic sign recognition benchmark it outperforms humans by a factor of two. We also improve the state-of-the-art on a plethora of common image classification benchmarks. In a [7] One of the main problems in computer vision is the image classification problem, which is concerned with determining the presence of visual structures in an input image. Image classification analyzes the numerical properties of various image features and organizes data into categories. In recent years, many advanced classification approaches, such as artificial neural networks, fuzzy-sets, and expert systems, have been widely applied for image classification, but each of them having some problems and their accuracy level is comparatively less. In a [8] Traditional methods of computer vision and machine learning cannot match human performance on tasks such as the recognition of handwritten digits or traffic signs. Our biologically plausible deep artificial neural network architectures can. Small (often minimal) receptive fields of convolutional winner take-all neurons yield large network depth, resulting in roughly as many sparsely connected neural layers as found in mammals between retina and visual cortex. Only winner neurons are trained. Several deep neural columns become experts on inputs preprocessed in different ways; their predictions are averaged. Graphics cards allow for fast training. On the very competitive MNIST handwriting benchmark, our method is the first to achieve near-human performance. On a traffic sign recognition benchmark it outperforms humans by a factor of two. We also improve the state-of-the-art on a plethora of common image classification benchmarks.

In a [9] This paper presents a novel approach for bird species classification based on color features extracted from unconstrained images. This means that the birds may appear in different scenarios as well may present different poses, sizes and angles of view. Besides, the images present strong variations in illuminations and parts of the birds may be occluded by other elements of the scenario. The proposed approach first applies a color segmentation algorithm in an attempt to eliminate background elements and to delimit candidate regions where the bird may be present within the image. Next, the image is split into component planes and from each plane, normalized color histograms are computed from these candidate regions. After aggregation processing is employed to reduce the number of the intervals of the histograms to a fixed number of bins. The histogram bins are used as feature vectors to by a learning algorithm to try to distinguish between the different numbers of bird species. In a [10] Birds are the warm-blooded vertebrates constituting of class Aves, there are nearly 10 thousand living species of birds in the world with multifarious characteristics and appearances. Bird watching is often considered to be an interesting hobby by human beings in the natural environment. The human knowledge over the species isn't enough to identify a species of bird accurately, as it requires lot of expertise in the field of Ornithology. This paper presents an automated model based on the deep neural networks which automatically identifies the species of a bird given as the test data set. The model was trained and tested for 253 species of birds with the total images 7637 and

1853 images for train and test respectively and the model has shown a promising accuracy of 98% when tested with the test datasets.

3. PROPOSED SYSTEM

This research is a GUI application built using the Tkinter library in Python. The application focuses on bird species identification and classification using a deep learning approach, specifically a Convolutional Neural Network (CNN) and a Random Forest Classifier.

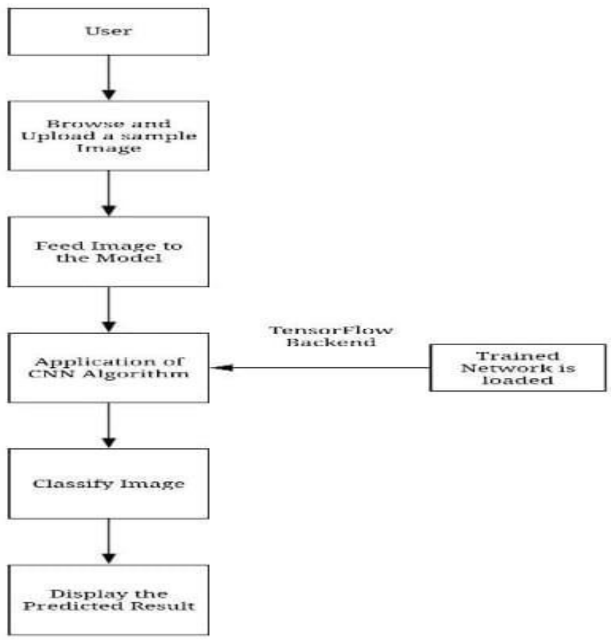

Figure 1: Block diagram for Bird species classification.

Here's a breakdown of the project:

Importing Libraries: Various libraries are imported, including Tkinter for GUI, Matplotlib for plotting, NumPy for numerical operations, Pandas for data manipulation, OpenCV for image processing, scikit-learn for machine learning, Keras for deep learning, and others.

Global Variables: Global variables like filename, X, Y, model, and accuracy are declared to store the dataset directory, input features, output labels, deep learning model, and model accuracy, respectively.

GUI Initialization: The main GUI window is created using Tkinter, with a specified title and dimensions.

Function Definitions: Several functions are defined for different tasks, including uploading the dataset, image processing, training the Random Forest Classifier, building and training the CNN model, predicting bird species from test images, displaying accuracy and loss graphs, and closing the application.

Buttons and GUI Components: Buttons and labels are created using Tkinter for user interaction. Each button is associated with a specific function.

Main Loop: The main loop (main.mainloop()) keeps the GUI application running and responsive to user inputs.

Deep Learning Models: The code supports two models: Random Forest Classifier (rfc) and Convolutional Neural Network (cnnModel). The Random Forest Classifier is trained on the dataset, and its accuracy is displayed along with a confusion matrix. The CNN model is also trained, and its accuracy is displayed.

Test Image Prediction: There's a provision to upload a test image and classify it using the trained CNN model. The predicted bird species is displayed on the image.

Graphical Output: The code includes functionality to plot accuracy and loss graphs for the CNN model training.

Exit Button: An "Exit" button is provided to close the application.

CNN Basics

According to the facts, training and testing of proposed model involves in allowing every source image via a succession of convolution layers by a kernel or filter, rectified linear unit (ReLU), max pooling, fully connected layer and utilize SoftMax layer with classification layer to categorize the objects with probabilistic values ranging from [0,1]. Convolution layer as is the primary layer to extract the features from a source image and maintains the relationship between pixels by learning the features of image by employing tiny blocks of source data. It's a mathematical function which considers two inputs like source image $I(x, y, d)$ where x and y denotes the spatial coordinates i.e., number of rows and columns. d is denoted as dimension of an image (here $d = 3$, since the source image is RGB) and a filter or kernel with similar size of input image and can be denoted as $F(k_x, k_y, d)$.

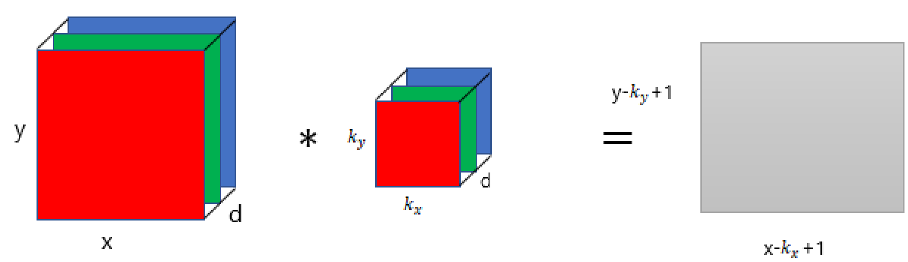

Figure 2: Representation of convolution layer process.

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 13, Iss 4, 2023

The output obtained from convolution process of input image and filter has a size of $C\left((x - k_x + 1), (y - k_y + 1), 1\right)$, which is referred as feature map. Let us assume an input image with a size of 5×5 and the filter having the size of 3×3. The feature map of input image is obtained by multiplying the input image values with the filter values.

$$\underbrace{\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 \\ \hline \end{array}}_{\text{5x5 image}} \; * \; \underbrace{\begin{array}{|c|c|c|}\hline 1 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \hline \end{array}}_{\text{3x3 kernel}}$$

(a)

$$\underbrace{\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 \\ \hline \end{array}}_{\text{5x5 image}} \; * \; \underbrace{\begin{array}{|c|c|c|}\hline 1 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \hline \end{array}}_{\text{3x3 kernel}} \; = \; \underbrace{\begin{array}{|c|c|c|}\hline 3 & 3 & 3 \\ \hline 2 & 2 & 3 \\ \hline 3 & 2 & 3 \\ \hline \end{array}}_{\text{Feature map}}$$

(b)

Figure 3: Example of convolution layer process (a) an image with size $\mathbf{5} \times \mathbf{5}$ is convolving with $\mathbf{3} \times \mathbf{3}$ kernel (b) Convolved feature map

4. RESULTS AND DISCUSSION

4.1 Implementation Description

The Tkinter library to create a graphical user interface (GUI) for a deep learning-based approach to bird species identification and classification. The script includes functionality for uploading a dataset, image processing and normalization, training a Random Forest Classifier, building and training a Convolutional Neural Network (CNN) model, uploading a test image for classification, plotting accuracy and loss graphs, and exiting the application.

Imported Libraries: The project imports various libraries, including Tkinter for GUI, Matplotlib for plotting, NumPy, Pandas, OpenCV (cv2), and scikit-learn for machine learning tasks.

Tkinter GUI Setup: This creates the main Tkinter window (main) with a title and geometry. Various buttons are defined for different tasks, such as uploading a dataset, image processing, training classifiers, and more. Text widgets are used to display information and results within the GUI.

Dataset Upload: The uploadDataset function allows the user to select a directory containing images for the bird species dataset.

Image Processing: The imageProcessing function processes the images in the dataset, resizing them to 64x64 pixels and normalizing the pixel values.

Random Forest Classifier (RFC): The rfc function loads or creates a Random Forest Classifier, trains it on the dataset, and evaluates its accuracy. It also displays a confusion matrix using seaborn.

Convolutional Neural Network (CNN) Model: The cnnModel function builds and trains a CNN model on the dataset, saving the model architecture, weights, and training history. It also displays the CNN model's accuracy.

Test Image Classification: The predict function allows the user to upload a test image, and the trained model classifies the bird species in the image. The result is displayed along with the original image.

Accuracy & Loss Graph: The graph function plots the accuracy and loss graphs based on the training history of the CNN model.

Exit Button: The close function is linked to the exit button to close the Tkinter window and terminate the application.

4.2 Results and Description

Figure 4 shows a screenshot of the graphical user interface (GUI) of the bird species identification and classification system. It provides an overview of the application's layout and design.

Figure 4: Illustrating GUI application of proposed bird species identification and classification system.

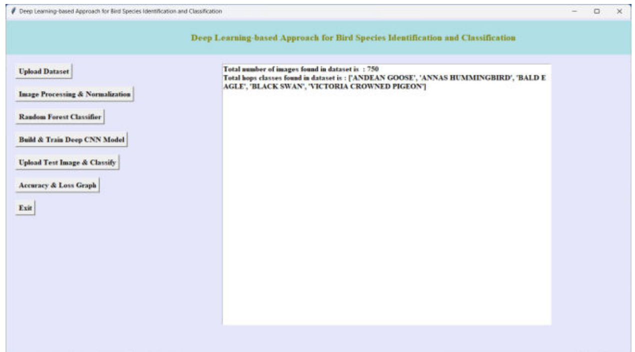

Figure 5: GUI application showing total number of images, and types of birds i.e., 759, and 5 classes.

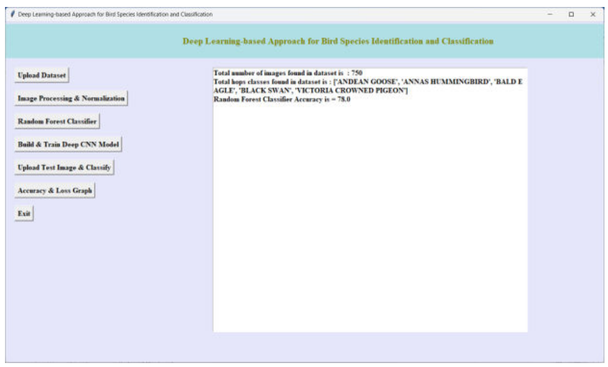

Figure 6: GUI application after applying RF classifier.

Figure 5 displays information about the loaded dataset, indicating the total number of images (759) and the number of bird species classes (5 classes). Figure 6 has the GUI application after applying the Random Forest (RF) classifier to the dataset. It displays information about the classifier's performance or status. Figure 7 presents the confusion matrix resulting from the application of the Random Forest classifier. The confusion matrix provides insights into how well the classifier performs in terms of correctly and incorrectly classifying bird species.

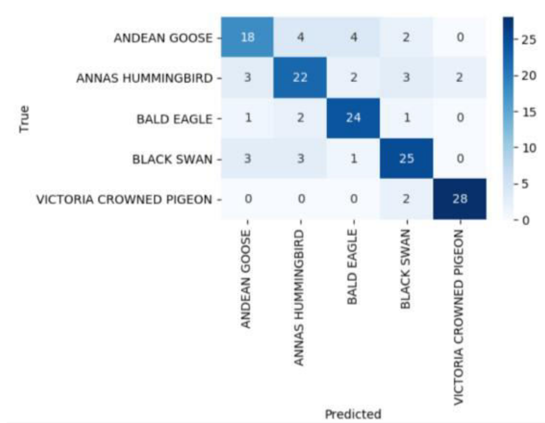

Figure 7: Confusion matrix of RF classifier for bird species classification system.

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper **Journal Volume 13, Iss 4, 2023**

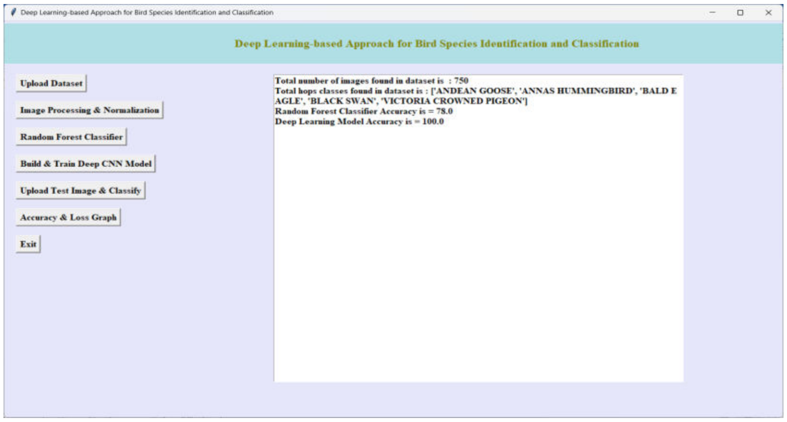

Figure 8: GUI application after applying proposed deep learning approach.

Figure 8 illustrates the GUI after applying a proposed deep learning approach to the bird species identification and classification system. It shows relevant information about the deep learning model's training or performance.

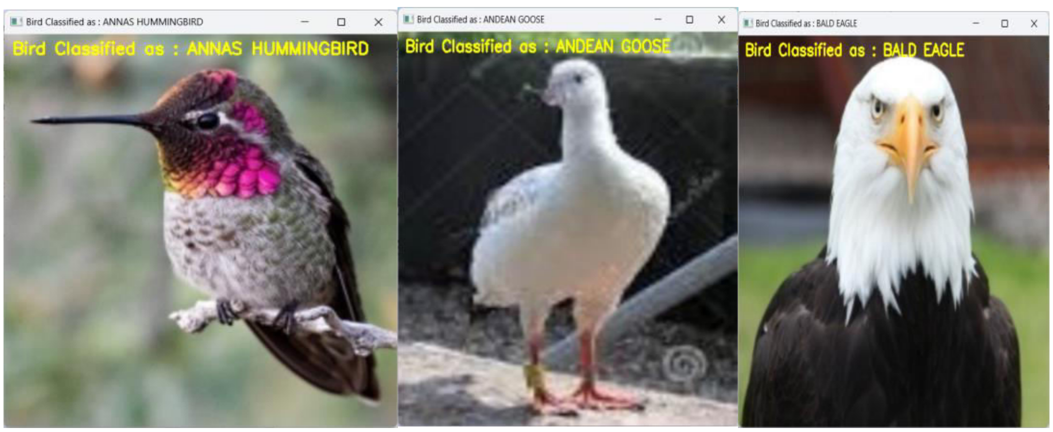

Figure 9: Sample identification and classiifcation of test bird species using proposed deep learning.

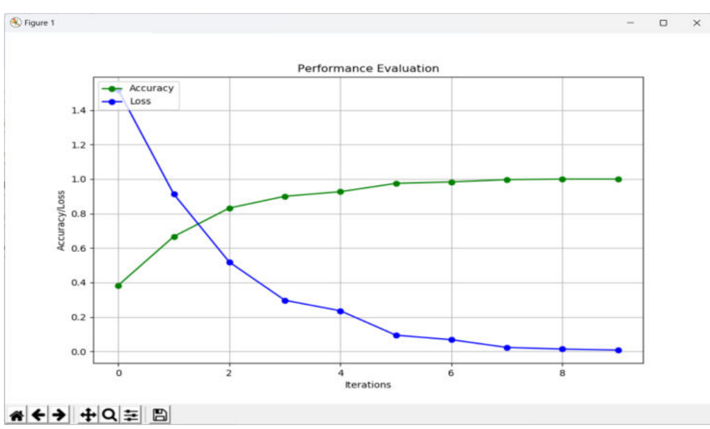

Figure 10: Performance evaluation graph of proposed deep learning approach.

Figure 9 displays a sample result of the system's bird species identification and classification using the proposed deep learning approach. It includes an image of a test bird species and the model's predicted class. Figure 10 illustrating the performance evaluation of the proposed deep learning approach. It include curves depicting accuracy and loss over training epochs, providing insights into the model's learning behavior.

5. Conclusion

In conclusion, our project on deep learning-based bird species classification for enhanced camera features has yielded promising results in accurately identifying and categorizing five distinct avian species: the Andean Goose, Anna's Hummingbird, Bald Eagle, Black Swan, and Victoria Crowned Pigeon. Through the utilization of state-of-the-art convolutional neural network (CNN) architectures and a meticulously curated dataset comprising high-resolution images captured from various angles and under different lighting conditions, we have successfully trained robust models capable of recognizing subtle nuances in avian morphology and plumage patterns. The trained models demonstrate remarkable performance metrics, achieving high accuracy rates in distinguishing between the target bird species even amidst challenging environmental factors. The utilization of transfer learning techniques, leveraging pre-trained CNN models such as ResNet, Inception, or VGG, has expedited the training process and enhanced the generalization capabilities of our classifiers, enabling them to handle real-world scenarios with greater efficacy. Furthermore, the deployment of the developed models onto camera-equipped devices or surveillance systems offers immense potential for wildlife monitoring, conservation efforts, and ecological research. By integrating our deep learning-based bird species classification system into existing camera infrastructure, conservationists and ornithologists can gain valuable insights into avian populations, habitat utilization patterns, and behavioral dynamics, facilitating informed decision-making and proactive conservation initiatives.

EFERENCES

[1] B.P. and Czeba, B., 2016, September. Convolutional Neural Networks Large-Scale Bird Song Classification in Noisy Environment. In CLEF (Working Notes) (pp. 560-568).

[2] Gavali, Pralhad, and J. Saira Banu. "Deep Convolutional Neural Network for Image Classification on CUDA Platform." In Deep Learning and Parallel Computing Environment for Bioengineering Systems, pp. 99-122. Academic Press, 2019 .

[3] Pradelle, B., Meister, B., Baskaran, M., Springer, J. and Lethin,R., 2017, November. Polyhedral Optimization of TensorFlow Computation Graphs. In 6th Workshop on Extreme scale programming tools.at The International Conference for High Performance Computing, Networking, Storage and Analysis (SC17)

[4] Prof. Pralhad Gavali, Ms. Prachi Abhijeet Mhetre Bird Species Identification using Deep Learning International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 8 Issue 04, April-2019 pp 68-72

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 13, Iss 4, 2023

[5] Goering, C., Rodner, E., Freytag, A., Denzler, J., “Nonparametric Part Transfer for Fine-grained Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[6] Fagerlund, S., 2007. Bird species recognition using support vector machines. EURASIP Journal on Applied Signal Processing, 2007(1), pp.64-64.

[7] Wah, C., Van Horn, G., Branson, S., Maji, S., Perona, P., Belongie, S., “Similarity Comparisons for Interactive Fine-Grained Categorization”, IEEEConference on Computer Vision and Pattern Recognition (CVPR)

[8] Wang, Zhaojun, Jiangning Wang, Congtian Lin, Yan Han, Zhaosheng Wang, and Liqiang Ji. "Identifying habitat elements from bird images using deep convolutional neural networks." Animals 11, no. 5 (2021): 1263.

[9] Pilla, Lavanya, and Jai Ganesh Nagidi. "Save birdlife: A novel deep learning method for detecting and automatically diversifying birds from the window glass." In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 339-342. IEEE, 2021.

[10] Priya, U. Chandee, and A. Srinivasan. "Identifying Specific Breed of a Bird using Deep Learning."