IMAGE BASED BIRD DETECTION AND CLASSIFICATION THROUGH DEEP LEARNING

BY

Sumaya Islam
ID: 201-15-13675

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Abdus Sattar
Assistant Professor & coordinator MSc.
Department of CSE
Daffodil International University

Co-Supervised By

Raja Tariqul Hasan Tusher
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2023

APPROVAL

This Project titled **"Image Based Bird Detection and Classification through Deep Learning"**, submitted by Sumaya Islam, ID No: 201-15-13675 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *24th January 2024*.

BOARD OF EXAMINERS

Chairman

Dr. S.M Aminul Haque (SMAH)
Professor & Associate Head
Department of Computer Science and Engineering
Daffodil International University

Internal Examiner

Md. Abbas Ali Khan
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Internal Examiner

Mohammad Monirul Islam
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

External Examiner

Dr. Ahmed Wasif Reza
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that, this thesis has been done by us under the supervision of **Abdus Sattar, Assistant Professor, Department of CSE** Daffodil International University. I also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Raja Tariqul Hasan Tusher
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Sumaya

Sumaya Islam
ID: 201-15-13675
Department of CSE
Daffodil International University

©Daffodil International University ii

ACKNOWLEDGEMENT

First I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible to complete the final thesis successfully.

I really grateful and wish my profound my indebtedness to **Abdus Sattar**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of “Human Computer Interaction in Education” to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Professor, and Head,** Department of CSE, for his kind help to finish my thesis and also to other faculty members and the staff of CSE department of Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and passion of my parents.

ABSTRACT

Birds, as a diverse and integral component of our natural world, play a pivotal role in maintaining the balance and health of ecosystems. Their presence influences not only the environment but also various aspects of human life, making the accurate identification and understanding of birds crucial. Bird detection and classification from images is a challenging task with diverse applications, ranging from wildlife conservation and ecological studies to urban planning and agriculture. This research paper aims to explore the use of deep learning techniques for accurately detecting and classifying birds in images. Deep learning is used in many fields [1], like speech, image recognition, drug discovery and toxicology, customer management, recommendation systems, bioinformatics, NLP etc. In our paper delve into various state-of-the-art deep learning architectures, including EfficientNetB7, MobileNetV3, ResNet101. The research involves the preparation of a comprehensive dataset of bird images and evaluates the performance of different models based on various metrics, such as precision, recall, and F1-score. By enhancing our capacity to identify and comprehend birds, we strengthen our ability to protect and conserve the intricate web of life on our planet. Furthermore, we discussed the challenges encountered during the process and proposed potential avenues for future research to enhance bird detection capabilities in real-world scenarios.

Keyword: Bird detection, deep learning, EfficientNetB7, MobileNetV3, ResNet101.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Birds, with their captivating diversity of species and behaviors, have long held a fascination for humanity. Birds play an important role in the natural world and are an integral part of many ecosystems. They help to control insect populations, disperse seeds, and pollinate flowers, contributing to the health and balance of many plant and animal communities. Their diverse physical adaptations, behaviors, and vocalizations make them a rich subject of scientific research and inquiry. In this context, accurate bird identification takes on paramount importance, as it is a linchpin for conservation efforts and ecological research. Advancements in deep learning have opened new avenues for precise bird detection and classification, providing a technological solution to a longstanding challenge in ornithology. Urbanization and technology have led to a loss of knowledge and the ability to recognize various bird species and wildlife [6]. Many have worked on papers to identify birds with images, audio, or video. However, there are several complications in processing these audio or video datasets, such as the presence of mixed noise, insects, and other real-world objects. People are generally more comfortable finding something with images than with audio or video. By using automatic image recognition [2], we can recognize the image efficiently and get the corresponding information. By detecting birds, we will be able to understand population dynamics, save endangered birds, assess habitat quality, classify them, and inform conservation strategies. As we confront pressing environmental issues and biodiversity loss, the ability to precisely identify and comprehend birds becomes a pivotal tool for maintaining the health and balance of ecosystems. Deep learning is increasingly utilized for image recognition, as it allows the image to be directly input to the recognition network. This study underscores the urgent need for technology-driven bird identification, emphasizing its profound impact on scientific inquiry and the conservation of the natural world.

1.2 Motivation

The motivation behind the research paper "Image-Based Bird Detection and Classification through Deep Learning" lies in the convergence of technological advancements and the imperative to address critical challenges in ornithology and wildlife conservation. Rapid progress in deep learning presents an unprecedented opportunity to revolutionize traditional methods of bird identification, offering the potential for more accurate and efficient classification. With growing concerns about biodiversity loss and environmental changes, there is an increasing need for advanced tools to monitor and conserve wildlife. The labor-intensive and time-consuming nature of manual observation further underscores the motivation to automate these processes through deep learning techniques. The diverse array of bird species presents a complex identification task, making deep learning an attractive solution due to its capacity to learn intricate patterns from extensive datasets. By fostering interdisciplinary collaboration between ornithology and technology, the research aims to not only contribute to the scientific understanding of bird populations but also offer practical applications in wildlife monitoring, ecological research, and conservation management.

1.3 Problem Definition

The paper is looking at a problem with how we currently identify birds. Right now, people have to do it manually, which takes a lot of time, can have mistaken, and is hard when there are many different kinds of birds. It's also tough to keep up with big bird populations. The paper says we need better tools, and it's interested in using new technology called deep learning to help. Deep learning can be good at spotting details in bird pictures that are hard for people to see. The paper also mentions that birds have patterns that are tricky to understand, and our current methods struggle with that. The goal is to find a way to identify birds that is faster and uses fewer resources. The paper is also bringing together bird experts and technology experts to work on this problem together.

1.4 Research Questions

The primary inquiries that this thesis focuses on are listed below:

- What is the recreational and educational value of recognizing birds?
- How does recognizing birds contribute to citizen science participation??
- What is the economic impact of bird recognition?

1.5 Research Methodology

In this section of my research paper, I describe the model's architecture, learning rate, optimizer, and experiment data set in addition to data preprocessing and model training. This chapter will conclude with a description of the suggested model's performance.

1.6 Research Objectives

There are some benefits of using AI in bird detection. They are given below:

- Sharing basic idea about Bangladeshi birds.
- Creating a relation between AI and ornithology.
- Investigate the potential for collaboration between deep learning models and human experts in ornithology.
- Exploring and highlighting practical applications of image-based bird detection in wildlife monitoring, ecological research, and conservation.

1.7 Research Layout

Chapter 1: I'll go over the project's introduction, purpose, problem definition, research question, research methodology, and anticipated results.

Chapter 2: will go through the history of the study, relevant previous research, and the present situation from Bangladesh's point of view.

Chapter 3: will discuss the state of artificial intelligence in Bangladesh's ornithology area.

Chapter 4: will talk about how well the suggested model performs for this task and identify which performs the best.

Chapter 5: It focuses on the outcome and analysis of comparisons.

Chapter 6: It outlines the findings of this investigation.

Chapter 7: Here are all of the sources I looked up for this study.

CHAPTER 2

BACKGROUND

2.1 Terminologies

The application of deep learning techniques alongside other computer vision methods to precisely recognize and categorize various bird species based on their visual attributes. This interdisciplinary approach focuses on refining image interpretation through the development of innovative algorithmic and methodological frameworks. Deep learning models are trained and tested using a diverse dataset of bird images, encompassing a wide range of species and visual characteristics. The fraction of correctly identified samples compared to the total number of samples is used to determine the accuracy of bird species recognition. The effectiveness of trained deep learning models in accurately classifying bird species not included in the training dataset demonstrates their proficiency in distinguishing unobserved avian species. This capability is validated through the training dataset. The process involves identifying, analyzing, and implementing strategies to mitigate potential risks and uncertainties that may impede the successful completion of the bird detection project.

2.2 Related Works

Deep learning for image detection and classification is prevalent in healthcare for medical image analysis, autonomous vehicles for object recognition, and retail for product identification. In these applications, popular models include CNN, such as VGG16, ResNet, and EfficientNet, known for their effectiveness in learning hierarchical features from images. With this, AlexNet, GooLeNet, VGGNet, FCNN, U-Net, SegNet, PSPNet, and many more are performed [3].

The authors [4] used Convolutional Neural Networks (CNN) for image detection and recognition on both the MNIST and CIFAR-10 datasets. Specifically, they implemented CNN models for image recognition on the MNIST dataset and object detection on the CIFAR-10 dataset. They trained mode on a single CPU unit, and real-time data augmentation is used on the CIFAR-10 dataset. To reduce overfitting on the datasets, Dropout is used. Among the 70,000 images for image recognition from the MNIST dataset, the training and testing sets contain, respectively, 60,000 and

10,000 images with an accuracy of 99.6%. The CIFAR-10 dataset used for object detection contains 60,000 32x32-pixel color images, with 10 classes and 6,000 images per class. The accuracy achieved on the test set is 80.17%. The test batch for each class consists of 1,000 randomly selected images. It is mentioned that the accuracy of CIFAR-10 can be further improved by training with larger epochs and on a GPU unit.

The researchers [5] used the Kaggle-230-birds dataset, which consists of 20,000 images of 180 species of birds. The classification algorithms used in the study are divided into three groups. First, traditional machine learning algorithms include SVM, lLDA, AdaBoost, k-means, MLP, RF, QDA, Gaussian naive Bayes theorem, bagging classifiers, DT, k-neighbors, and gradient boosting, which got accuracy of 50–80% when there were 3 and 6–50% when there were 20 species. After that, CNN algorithms include ResNet-50, GoogLeNet, DensNet-121, and AlexNet, with some models reaching 97-98% accuracy. Finally, CNN algorithms based on transfer learning, including SVM and DT, showed even better results, with classification accuracy ranging from 90% to 100%. From the paper [6] it can be seen that the Convolutional Neural Network model was created and developed to perform bird species identification. The model consists of three convolutional layers, each followed by a pooling layer. The researchers developed a CNN model to extract information from bird images and trained the model using a dataset gathered through Microsoft's Bing Image Search API v7. The classification accuracy of the CNN model on the training set was observed to be 93.19%, while the accuracy on the testing set was 84.91%. The paper also discusses related works, such as bird species recognition using support vector machines, color features, and deep learning techniques. The authors suggest possible enhancements for the application, including increasing the number of predicted bird species, improving accuracy, using the Google Maps API for displaying bird locations, and developing a mobile application.

The researchers on this paper [7] used a CNN algorithm for bird detection in Bangladesh's wildlife. They designed the models to identify 1500 bird species using deep learning techniques. The models were trained using PyTorch and utilized pre-trained DCNNs on ImageNet. The architecture of the model included four convolutional layers followed by two fully joined layers. The model also incorporated batch normalization and utilized activation functions such as stride and padding. The researchers optimized the model using Adam's Optimizer with a learning rate of 0.01 with

95.52% validation accuracy. The researchers also used CNN and YOLOV3 framework for detecting input images.

In this paper [8] aerial photographs of both wild birds and bird decoys in different environments are captured using a UAV. Image preprocessing techniques such as cropping and augmentation are applied to obtain several hundred sub-images of the aerial photographs. These sub-images are used for training the deep-learning models. The training process involves using five different deep-learning-based object-detection methods, namely Faster R-CNN, R-FCN, SSD, Retinanet, and YOLO. Each method has its own meta-architecture and feature extractor. The hidden layers of each model learn feature representations. Finally, the performance of the bird detection models is evaluated through a testing process using average precision (AP) as the performance index. AP is calculated as the area under the precision-recall graph, with true positives, false positives, and false negatives being considered.

The authors [9] used a flower image dataset consisting of 10,880 images and have used a tiny darknet model and an improved model based on the tiny darknet structure. There are 17 kinds of flowers datasets, which were initially comprised of 1,360 images, which were then augmented to generate additional images. The authors developed a flower classification model based on the modified tiny darknet, trained it using image data sets, and used a softmax classifier for prediction. The experimental results showed that the authors' model achieved a recognition rate of 92%.

In this paper [10], for Mnist dataset the author use two different models for image recognition, SVM and a deep learning network with two convolution layers and two full-connection layers. The dataset consists of a total of 70,000 images. The authors use SVM to create a recognition model for linearly inseparable image data, achieving accuracy 93.92%. They then construct a deep learning network with two convolution layers and two full-connection layers for handwritten number categorization, achieving 98.85% accuracy.

In this document [11], a set of 1160 panoramic dental images from patients aged 42.8 years were obtained from the hospital and school of medicine using radiographic machines from Dentsply Sirona. The performance of nnU-Net and DenseNet121 models for caries lesion segmentation and classification. For caries detection, dentists and models found similar results, but models performed better for certain lesion depths (D1 and D2).

The document [12] discusses the use of deep learning models, specifically CNN, for plant disease detection through hyperspectral images. They collect datasets like public, real-field, from farmers and IoT sensors. It mentions that the accuracy of the models varied for different classes of diseases, with 99.9% accuracy for identifying northern corn leaf blight and 87% accuracy for other disease classes. The document emphasizes the need for minimal variation in detection accuracies across different classes of crop diseases and highlights the challenges and gaps in current CNN implementations.

2.3 Bangladesh Perspective

In the context of Bangladesh, image-based bird detection holds significant importance across various domains. Bangladesh is renowned for its rich biodiversity, encompassing a diverse range of bird species that play crucial roles in maintaining ecological balance and biodiversity. The application of image-based bird detection technology in Bangladesh provides valuable insights for conservation efforts, aiding in the monitoring and protection of indigenous bird species. With the country facing environmental challenges, such as habitat loss and climate change, accurate bird detection becomes a pivotal tool for assessing the impact of these changes on avian populations. Furthermore, the agricultural landscape in Bangladesh can benefit from bird detection technology, as certain bird species contribute to natural pest control, influencing agricultural practices. Leveraging image-based bird detection in Bangladesh can enhance ecological research, contribute to wildlife conservation initiatives, and support sustainable agricultural practices, fostering a harmonious coexistence between the vibrant avian biodiversity and the local communities.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

This work obtained 99.57% accuracy by using 11 different classes of photos with Deep Learning models.

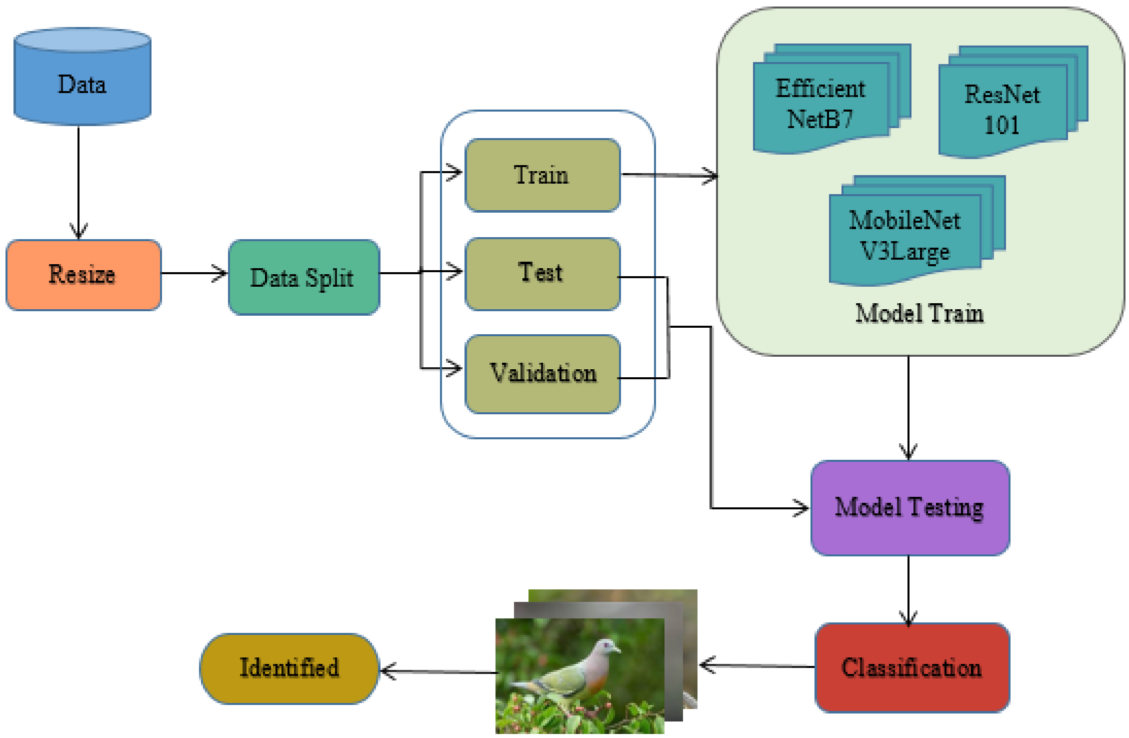

FIGURE 3.1 STEPS OF DATA COLLECTING AND PROCESSING

3.2 Experiment Data Set

There are eleven categories and two thousand five hundred sixty photos in this collection. While some of the photos in this collection were taken in natural settings, the majority came from the internet. The terrible part was that I had a hard time gathering all the data manually from the field due to a lack of equipment and incorrect place identifications. The purpose of this collection was to use images to identify birds. There are about 2560 photos in this dataset; about 2065 of them are used to train the model, 236 are used for validation, and 234 are used for testing.

FIGURE 3.2 VISUALIZATION OF MY COLLECTED DATASET

3.3 Data Pre-Processing

Some of the data was manually gathered from the field using a smartphone and camera, as the photographs were not all the same size and quality. The rest of the data was gathered from Google. The dataset was very challenging to train and test, and it displayed some inaccuracies. The dataset in this version has images with a set resolution. In compliance with the specifications of my project, I have resized every image to be square. I used a fixed resolution of 224 × 224 while creating the photos. I started by cropping the photos in order to get rid of any extraneous things. Next, I resized it to fit the 224 × 224 dimensions needed for the cropped photos. I also used the RGB color mode to train my models.

3.4 Architecture of the Model

The model is specifically built to accurately detect and classify birds from any location. The model employs many convolutional layers to improve its capacity for analyzing and extracting characteristics from the input data. In addition, the model has two completely linked layers to further enhance its predictions. To maximize the model's efficiency, it incorporates a fully linked layer with batch normalization, numerous thick layers, and dropout. With a kernel size of three, the first convolutional layer functions as the input layer and processes a 224x224 input shape in

RGB color mode. The layer has 64 as its filter size. The layer maintains the same padding characteristic and uses the Rectified Linear Unit (ReLU) activation function with a stride of 1. When the input value is less than or equal to zero, the ReLU function produces zero; otherwise, it outputs the input value if it is greater than zero**.**

$$f(x) = max(0, x) \tag{1}$$

The first average pooling layer receives the output of the first convolutional layer as input. By employing sub-regions binned average pooling, the parameter count is decreased. To do this, an average pooling layer is employed with a pool size of 2 and strides of 2. The second layer utilizes a filter size of 128, a kernel size of 3, and a stride of 1. This layer integrates batch normalization to address the problem of vanishing gradient in the convolutional network during the training process. This regularization strategy improves performance and decreases training time. In order to utilize a greater learning rate and expedite the learning process, batch normalization is employed. The subsequent pooling layer is linked to the output of the second convolution layer. The layer uses strides of two and has a pool size of two as well.

Layer 3 retains the same kernel size, strides, and batch normalization features as layer 2, but with a filter size of 256. With a pool size and stride of two, the third layer's output is linked to the third pooling layer. Layer 4's filter size of 512 indicates that its properties are the same as those of Layer 3. It is important to mention that this degree of regularization could be overly stringent.

Two dense layers correspond to a matrix-vector combination after the four preceding levels. 512 hidden units and an applied activation function make up the first dense layer. The dropout technique is utilized in this situation to mitigate the problem of overfitting. To standardize the model, the density of the second dense layer is 1024. Softmax is the activation function that is being used. This includes the model-development process.

By integrating these supplementary layers into the model, the resulting output becomes more polished and free from any interference or distortion. Moreover, the use of two fully interconnected thick layers improves the efficiency of the categorization task. These layers are enhanced with batch normalization and dropout features, which further enhance the capabilities of the model. Afterwards, by applying a softmax function to it, the data will be converted into probabilities that correspond to each class. A softmax function is a mathematical function that assigns a value between 0 and 1 to each real number.

The categorical cross-entropy is a loss function employed to quantify the loss in categorical classification problems. Afterwards, the optimizer will aid in carrying out gradient descent, while the metrics will be updated to ensure accuracy, as this is a classification problem.

Figure 3, which depicts the architecture of the model's summary:

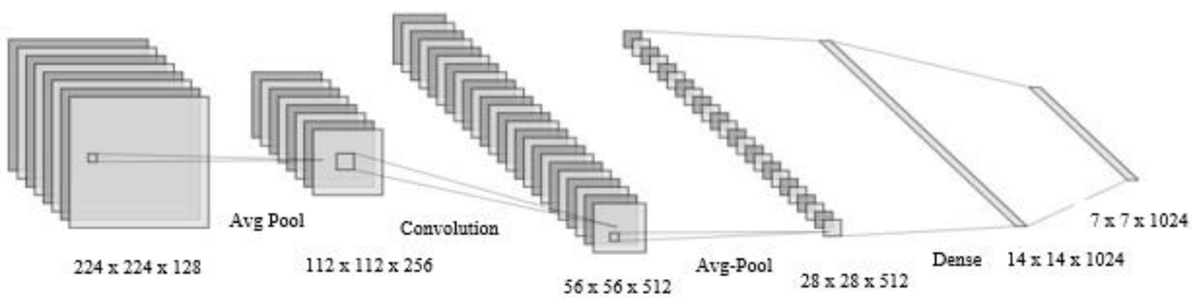

FIGURE 3.4 REPRESENTATION OF MODEL SUMMARY

3.5 Learning rate and Optimizer of the model

Update network weights based on training data using Adam Optimizer 0.001 to achieve faster results at the first learning level. Technically speaking, they did this by moving on to the following iteration, which updated the network weights. The learning rate and optimizer are critical components in the training of machine learning models. The size of the step executed in each optimization iteration is determined by the learning rate, which functions as a hyperparameter. It plays a pivotal role in balancing the convergence speed and stability of the model.

3.6 Data Augmentation

By rotating 30 degrees and changing the width, height shift range of 0.2, and shear range of 0.15, where the zoom range is 0.15, I augmented my training, testing, and validation data. Transforming training data through data augmentation is a powerful technique aimed at improving the accuracy and robustness of a classifier. By applying various transformations to each sample x_i, such as rotations, flips, shifts, and shears, the classifier becomes more adept at generalizing to diverse scenarios. The goal is to generate an augmented dataset where each transformed instance not only represents a modified version of the original sample but also enhances the model's ability to handle real-world variations in input data. Within a neural network, the anticipated class label y_i and a probability distribution p_{ij} for every class j are the outputs for every augmented sample.

$$y = argmax_j \, max_{1 \le i \le r} \, (p_{ij}) \qquad (2)$$

$$Averaging\ all\ predictions: y = \frac{1}{r} \sum_{i=1}^{r} p_{ij} \qquad (3)$$

Equation (2) proposes an approach that involves averaging the probabilities for each class across all augmented samples. The equation (3) calculates the predicted class label (y) by averaging the probabilities (p_{ij}) for each class across all augmented samples (r instances) for a given input sample.

Data augmentation is a potent fundamental method that assists in producing more durable, generalizable, and effective machine learning models. In my paper, I make various transformations to the existing training samples, like rotation, flip, zoom, shear, and shift. By augmenting the dataset with these variations, the model learns to recognize patterns and features that are invariant to such transformations.

3.7 Training the model

With over 66 million parameters, EfficientNetB7 is a deep neural network with an intricate design that achieves efficiency through a compound scaling method, depth wise separable convolutions, and squeeze-and-excitation blocks for image classification tasks. Because of this, the model performs better for me once I train it using my own collection of bird datasets. Here, a batch size of 32 with 10 epochs was chosen. With a test accuracy of 99.57% and a validation accuracy of 96.60%, my model exhibits amazing fineness accuracy. 224 × 224 is the intended image size for both testing and training. EfficientNetB7 excels in image identification tasks, leveraging its deep neural network architecture and efficient design for high-performance image classification.

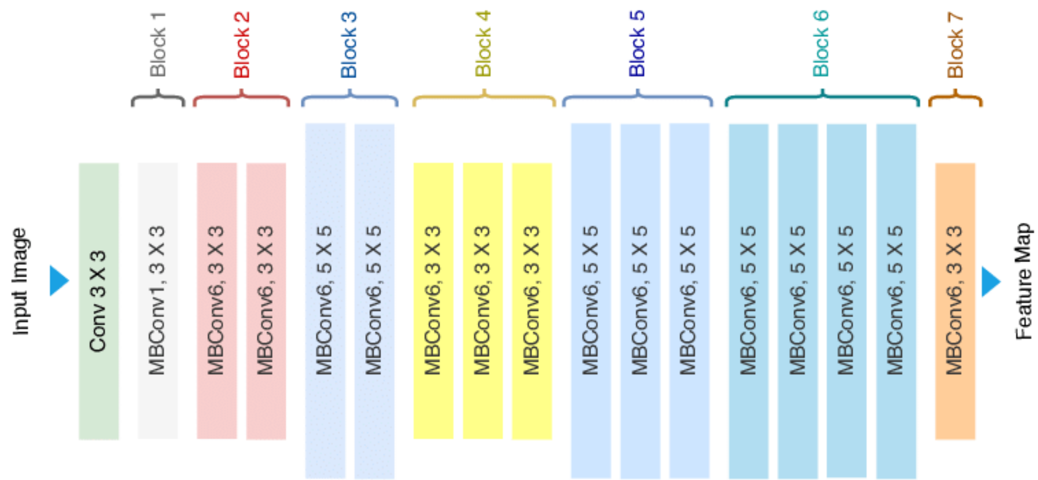

FIGURE 3.7 EFFICIENTNETB7 LAYER [source [13]]

CHAPTER 4
RESULT COMPARISON AND ANALYSIS

4.1 Training, Testing and the Validation of the model

I divide the data into training, test, and validation sets; there are roughly 2560 total photos in all. Approximately 80% of the total data (2065) is utilized for training, 10% is used for testing (234 total), and the remaining 10% (236 total) is used for model validation. I tested and trained the data sets using 32-batch processing in RGB color mode using categorical mode. I utilized three distinct models for this. Based on these, the research's proposed EfficientNetB7 model performs well. The birds in my validation data, together with my train and test data sets and other birds that the model does not know about, appear as follows:

FIGURE 4.1 VALIDATION DATASET

4.2 Model efficiency

The models utilized in this work, including EfficientNetB7, achieved 99.57% accuracy on the testing dataset and 96.60% accuracy on the validation data set after ten epochs. After completing the training and testing cycles, the work obtains a consecutive accuracy rate. It's possible to conclude that the suggested Deep Learning model is appropriate for use in bird identification tasks based on the accuracy and confusion matrix obtained from this work. This is a representation of the work's performance.

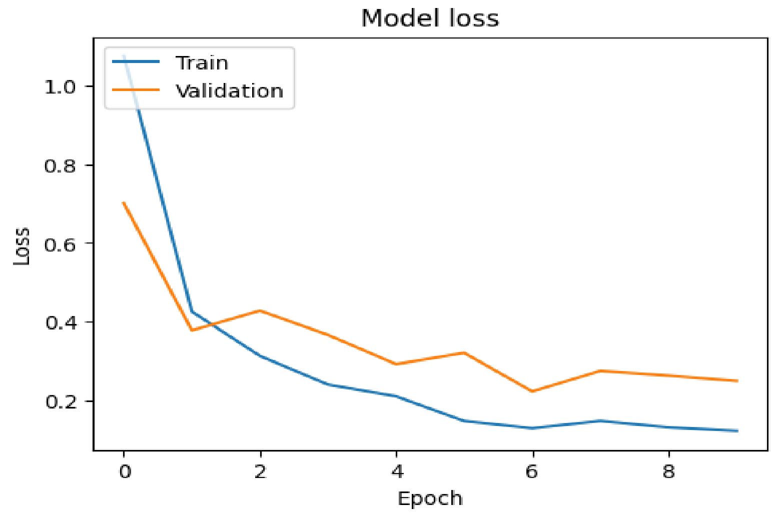

FIGURE (I) REPRESENTING TRAINING LOSS AND VALIDATION LOSS

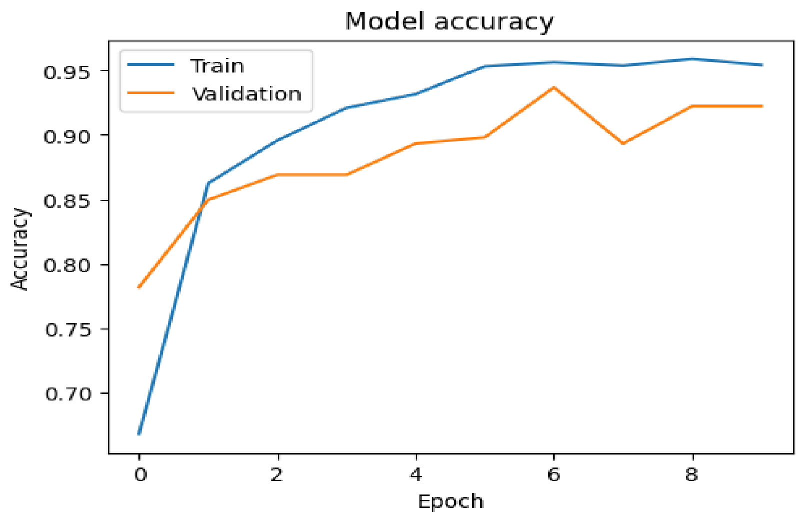

FIGURE (II) REPRESENTING TRAINING ACCURACY AND VALIDATION ACCURACY

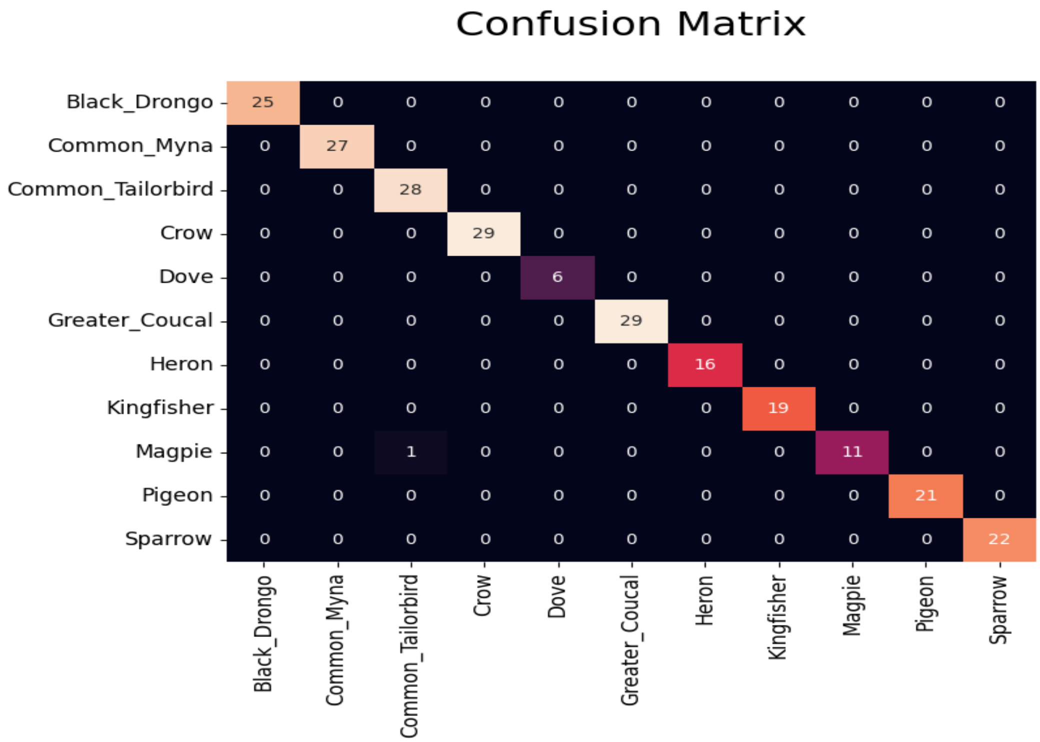

FIGURE (III) REPRESENTING CONFUSION METRIX

TABLE 4.2 BEST MODEL CLASSIFICATION REPORT

Label Name	Precision	Recall	F1-Score	Support
Black_Drongo	1.00	1.00	1.00	25
Common_Myna	1.00	1.00	1.00	27
Common_Tailorbird	0.97	1.00	0.98	28
Crow	1.00	1.00	1.00	29
Dove	1.00	1.00	1.00	6
Greater_Coucal	1.00	1.00	1.00	29
Heron	1.00	1.00	1.00	16
Kingfisher	1.00	1.00	1.00	19
Magpie	1.00	0.92	0.96	12
Pigeon	1.00	1.00	1.00	21
Sparrow	1.00	1.00	1.00	22
accuracy			1.00	234
macro avg	1.00	0.99	0.99	234
weighted avg	1.00	1.00	1.00	234

The presented table outlines the precision, recall, F1-score, and support metrics for various bird species classifications in the model evaluation. Each label, representing a specific bird species, demonstrates high precision, recall, and F1-score values, with the majority achieving perfect scores of 1.00, indicating a high level of accuracy in both positive and negative predictions. Notably, the model excels in recognizing species like Black Drongo, Common Myna, Crow, Dove, Greater Coucal, Heron, Kingfisher, Pigeon, and Sparrow, achieving perfect scores across all metrics. While the Magpie label exhibits a slightly lower precision of 1.00 and a recall of 0.92, resulting in a F1-score of 0.96, the overall performance is exceptional. The macro and weighted averages further reinforce the model's effectiveness, with macro F1-score at 0.99 and weighted averages at 1.00. These metrics collectively highlight the robust performance of the classification model across diverse bird species.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (4)$$

The performance metrics (4) are often used to characterize the classification model's performance. This is the False Positive (FP) rate, which occurs when negative values are expected to be positive even when the observation is negative; A true positive (TP) rate indicates that both the observation and the prediction are positive; For each kind of dataset, there is a true negative (TN) rate, where projected values are accurately forecasted as an actual negative, and a false negative (FN) mean, where positive values are predicted as negatively.

4.3 Work comparison

When compared to the other articles, this study yields the highest accuracy for me. On this work, a few researchers have collaborated closely. The relationships between my work and some earlier research on various bird identification and classification tasks are displayed in Table 5.1.

TABLE 4.3 CONSIDERATION WITH SELECTED PRIOR WORKS

Work	Algorithm	Accuracy
Effective classification of birds' species based on transfer learning [5]	SVM	84.91%
Image based Bird Species Identification using Convolutional Neural Network [6]	CNN	84.91%
Machine Learning Approach for Bird Detection[7]	CNN	95.52%
Flower identification based on Deep Learning[9]	Darknet	92%
Image Based Bird Detection and Classification through Deep Learning	EfficientNetB7	99.57%

It's evident from the above table that numerous researchers approach different types of algorithms with varying accuracy. However, 95.52% is the highest accuracy out of all the accuracy. The researchers also used CNN and the YOLOV3 framework for detecting input images. To obtain better accuracy, they use high-resolution images, though their accuracy is less than mine. Using CNN, another study achieved 84.91% accuracy. They tried to improve the accuracy of their work by utilizing the Deep Belief Network as well, but they were unsuccessful. Following their examination of the research, they put into practice a CNN-based model that achieved 92% accuracy, the second-highest accuracy of all the studies in the above table. They train their model using the Adam gradient descent algorithm to save training time and improve convergence.

The table demonstrates that my accuracy surpasses that of every work mentioned above. All I did was resize the picture to fit my needs. I found better accuracy in EfficientNetB7 which is 99.57%. This model is capable of accurately identifying the bird, and it may be a superior method for identifying unfamiliar species in our environment.

4.4 Experimental Results & Analysis

TABLE 4.4 ACCURACY OF MODELS

Model Name	Validation Accuracy	Test Accuracy	Training Time (sec)
EfficientNetB7	96.60%	99.57%	401.63
MobileNetV3Large	94.66%	95.30%	381.36
ResNet101	92.23%	92.74%	392.28

The table presents the performance metrics of three deep learning models—EfficientNetB7, MobileNetV3Large, and ResNet101—applied to image-based bird detection and classification. EfficientNetB7 exhibits the highest validation accuracy at 96.60% and an impressive test accuracy of 99.57%, outperforming the other models. MobileNetV3Large and ResNet101, while demonstrating respectable validation and test accuracies of 94.66% and 92.23%, and 95.30% and 92.74%, respectively, showcase competitive results. The training times for EfficientNetB7, MobileNetV3Large, and ResNet101 are 401.63 seconds, 381.36 seconds, and 392.28 seconds, respectively. These metrics provide valuable insights into the trade-offs between model accuracy and computational efficiency, aiding researchers in selecting an appropriate model based on their specific requirements.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on society

The research paper carries significant implications for society, as its findings on accurate bird detection and classification have far-reaching impacts across various domains. Within the realm of wildlife conservation, the newfound ability to identify and monitor bird species emerges as a vital tool in the ongoing effort to preserve biodiversity. The influence extends into urban planning, where insights derived from bird populations shape decisions concerning green spaces and contribute to the establishment of ecological equilibrium within urban environments. Moreover, the potential applications of this research stretch into ornithological and environmental monitoring, providing invaluable datasets for scientific research in these fields. By elevating our capability to recognize and study birds through deep learning, this research underscores its potential to cultivate a harmonious coexistence between human activities and the avian world. This, in turn, holds the promise of contributing significantly to the well-being of both ecosystems and society at large.

5.2 Impact on Environment

The research's impact on the environment carries profound implications. The precision of bird detection becomes a linchpin for a nuanced comprehension of avian populations, assuming a pivotal role in sustaining overall ecosystem health. This knowledge becomes instrumental not only in assessing habitat quality but also in the identification of endangered species, paving the way for the implementation of targeted conservation strategies. By harnessing the capabilities of deep learning for efficient bird classification, the research introduces a technologically advanced tool for environmental scientists and conservationists. This facilitates more comprehensive decision-making processes, fostering an informed approach to protecting and sustaining biodiversity. In its broader scope, the research emerges as a catalyst for the positive transformation and preservation of natural habitats, acknowledging the intricate ecological relationships that birds foster within their environments.

5.3 Ethical Aspects

Exploring the ethical dimensions of image-based bird detection and classification through deep learning becomes imperative in navigating the intricate intersection of technology and environmental ethics. Privacy concerns loom prominently, particularly in the conscientious collection and utilization of data, necessitating meticulous considerations and unwavering adherence to ethical guidelines to safeguard the fundamental rights of individuals. A pivotal facet of ethical diligence lies in ensuring transparency throughout the developmental trajectory of deep learning models. Identifying and rectifying biases within training data becomes a critical mission to preclude potential discriminatory outcomes. Ethical scrutiny extends beyond privacy to encompass environmental impacts, scrutinizing potential disturbances induced by data collection processes. Striking an equilibrium between the rapid pace of technological advancements and ethical principles becomes paramount, serving as the linchpin to foster responsible practices and uphold the unassailable integrity of research in this dynamic and consequential domain.

5.4 Sustainability

Consideration of sustainability in the context of image-based bird detection and classification through deep learning is essential. Sustainable practices encompass various aspects, including data collection methods, model development, and the broader ecological impact. Implementing sustainable data collection involves minimizing disturbances to bird habitats and ecosystems, ensuring the long-term health of the environment. In model development, adopting energy-efficient algorithms and hardware contributes to sustainable computing practices. Moreover, a commitment to ongoing research that prioritizes ecological balance and the well-being of bird species reinforces the sustainability of the overall endeavor. Striving for sustainability in every facet of the research process is integral to ensuring the longevity and positive impact of image-based bird detection and classification efforts.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this paper, I explored the application of deep learning models, specifically EfficientNetB7, ResNet101, and MobileNetV3Large, for the task of bird detection and classification based on image data. Our analysis revealed that EfficientNetB7 outperformed the other models in terms of classification accuracy, showcasing its suitability for complex image recognition tasks. Additionally, the ResNet101 and MobileNetV3Large models displayed commendable performance, offering valuable insights into the trade-offs between model complexity and computational efficiency in the context of bird classification. The results highlight the importance of accurate bird detection in various domains, including wildlife conservation, urban planning, and their potential in ornithological and environmental monitoring applications. By addressing challenges and exploring future research avenues, we anticipate improved bird detection capabilities that will contribute to the advancement of ecological studies and environmental monitoring. However, challenges persist, warranting future investigations. Despite the success of deep learning models, we encountered some challenges during the research, such as limited training data for certain rare bird species and handling images with multiple overlapping birds. We discuss potential solutions to these challenges and suggest future research directions to improve bird detection performance further.

6.2 Future work

I analyze possible remedies to overcome obstacles and propose future avenues of investigation to enhance the accuracy of bird detection even further. Subsequent research should prioritize the implementation of sophisticated data augmentation methods, leveraging transfer learning from extensive datasets, employing IoT-based detection systems, and integrating multi-modal approaches to enhance the resilience of the model. Furthermore, the immediate implementation and involvement of the public through citizen science activities are crucial for the actual use and growth of the produced models in ecological research and conservation endeavors. Future work should focus on advanced data augmentation techniques, transfer learning from larger datasets, and multi-modal integration to improve model robustness. Additionally, real-time deployment and community engagement through citizen science initiatives are vital for the practical application and expansion of the developed models in ecological studies and conservation efforts. This research lays the foundation for the intersection of artificial intelligence and ornithology, emphasizing the need for continued exploration to address real-world challenges and promote sustainable environmental practices.

REFERENCES

[1]. NF Hordri, SS Yuhaniz, SM Shamsuddin "Deep Learning and Its Applications: A Review" Conference on Postgraduate Annual Research on Informatics Seminar, 2016. [**Google Scholar**]

[2]. G Olague, S Cagnoni, E Lutton "Introduction to the special issue on evolutionary computer vision and image understanding" Pattern Recognition Letters, 27(11), 2006. [**Google Scholar**]

[3]. L Cai, J Gao, and D Zhao "A review of the application of deep learning in medical image classification and segmentation" Annals of translational medicine, 8(11), 2020. [**Google Scholar**] [**CrossRef**]

[4]. R Chauhan, KK Ghanshala, RC Joshi "Convolutional neural network (CNN) for image detection and recognition" 2018 first international conference on secure cyber computing and communication (ICSCCC), pp. 278-282. IEEE, 2018. [**Google Scholar**] [**CrossRef**]

[5]. M Alswaitti, L Zihao, W Alomoush, A Alrosan, K Alissa "Effective classification of birds' species based on transfer learning" in IJECE, 2022. [**Google Scholar**] [**CrossRef**]

[6]. S Raj, S Garyali, S Kumar, S Shidnal "Image based Bird Species Identification using Convolutional Neural Network" in IJECE, 2020 Aug 1; 12(4), pp. 4172-4184. [**Google Scholar**]

[7]. F Mashuk, A Sattar, N Sultana "Machine Learning Approach for Bird Detection" in Third International Conference on Intelligent Communication Technologies and Virtual Mobile Network (ICICV), pp. 818-822. IEEE, 2021. [**Google Scholar**] [**CrossRef**]

[8]. SJ Hong, Y Han, SY Kim, AY Lee, G Kim "Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery" Sensors 19(7), (2019): 1651 [**Google Scholar**]

[9]. M Tian, H Chen, Q Wang "Flower identification based on Deep Learning" Journal of Physics: Conference Series, Vol. 1297, No. 2, p. 022060, IOP Publishing, 2019. [**Google Scholar**]

[10]. Y Lai "A Comparison of Traditional Machine Learning and Deep Learning in Image Recognition" Journal of Physics: Conference Series, Vol. 1314, No. 1, p. 012148. IOP Publishing, 2019. [**Google Scholar**]

[11]. L Lian, T Zhu, F Zhu, H Zhu "Deep Learning for Caries Detection and Classification" Diagnostics 11(9), p. 1672, 2021. [**Google Scholar**] [**CrossRef**]

[12]. M Nagaraju, P Chawla "Systematic review of deep learning techniques in plant disease detection" International Journal of System Assurance Engineering and Management, 2020 Jun 11, pp. 547-560 [**Google Scholar**] [**CrossRef**]

[13]. T Ahmed, NHN Sabab "Classification and understanding of cloud structures via satellite images with EfficientUNet" SN Computer Science, 3, 99. 2022 [**Google Scholar**]

IMAGE BASED BIRD DETECTION AND CLASSIFICATION

ORIGINALITY REPORT

14%	11%	6%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
2	Submitted to Daffodil International University Student Paper	2%
3	www.mdpi.com Internet Source	1%
4	Submitted to FH Campus Wien Student Paper	1%
5	www.ijert.org Internet Source	1%
6	www.researchgate.net Internet Source	1%
7	Mohammed Alswaitti, Liao Zihao, Waleed Alomoush, Ayat Alrosan, Khalid Alissa. "Effective classification of birds’ species based on transfer learning", International Journal of Electrical and Computer Engineering (IJECE), 2022 Publication	<1%