Local Bird Species Classification Using Advanced Deep Learning Architecture

BY

Md. Abugalib Sourov
ID: 201-15-3600

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

Supervised By

Dewan Mamun Raza
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Co-Supervised By

Amir Sohel
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"Local Bird Species Classification Using Advanced Deep Learning Architecture"** submitted by Md. Abugalib Sourov, ID: 201-15-3600 to the Department of Computer Science and Engineering, Daffodil International University, has been acknowledged as satisfactory for its style and substance and accepted as being sufficient for the accomplishment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering.

BOARD OF EXAMINERS

Chairman

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

Internal Examiner

Nazmun Nessa Moon (NNM)
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dewan Mamun Raza (DMR)
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Md. Zulfiker Mahmud (ZM)
Associate Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

I, therefore, declare that this undertaking has been finished by us under the supervision of **Dewan Mamun Raza, Lecturer (Senior Scale)**, Department of CSE, Daffodil International University. I further declare that neither an application or an educational grant has been made anywhere for this project or any part of it.

Supervised by:

Dewan Mamun Raza
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Co-Supervised by:

Amir Sohel
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Submitted by:

-Sourov

Md. Abugalib Sourov
ID: 201-15-3600
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

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

I am really grateful and wish my profound indebtedness to **Dewan Mamun Raza**, Lecturer (Senior Scale), Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "Deep Learning" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

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

I would like to thank all of my classmates at Daffodil International University who participated in the discussion while also attending the course.

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

ABSTRACT

The goal of this research is to create a reliable system for classifying local bird species using modern deep learning architectures. The dataset forms the basis for this study and consists of 1,179 photos covering seven different species of birds. A combination of deep learning and transfer learning algorithms, such as "InceptionV3," "VGG19," "VGG16," "MobileNetV2," "CNN01," and "CNN02," are used in the search for precise and effective classification. The process is a painstaking procedure that begins with web crawling to add to the dataset after data collection from web scraping. Accurate annotations for model training are ensured via manual data labeling. The construction and training of convolutional neural network (CNN) models is the study's central focus. Important designs are used, including 'InceptionV3,' which is well-known for its deep and effective design, 'VGG19' and 'VGG16' with uniform architectures, and 'MobileNetV2,' which is especially made for devices with limited resources. Furthermore, flexibility is offered by the generic CNN designs ('CNN01' and 'CNN02'). The experimental findings show that the 'MobileNetV2' design produced the maximum accuracy of 99.58%. This shows how well the model can classify local bird species and make generalizations. The study highlights the importance of transfer learning via using model training to improve productivity and faster convergence. The accuracy obtained is proof of the effectiveness of the selected deep learning architectures for classifying bird species .This work offers a dependable and automated method for identifying local bird species and provides useful details about the use of modern deep learning methods in bird study. The results have effects on protecting nature, environmental monitoring, and the larger field of computer vision in nature research.

Keywords: *Local Bird Species, Deep Learning, Transfer Learning, InceptionV3, VGG19, VGG16, MobileNetV2, CNN01, and CNN02,*

TABLE OF CONTENTS

CONTENTS **PAGE**

CHAPTER

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the field of ecological studies, aerial diversity conservation and an understanding of ecological dynamics depend heavily on the identification and classification of local bird species. The identification of bird species using traditional methods mostly depends on manual observation, which is a labor-intensive and time-consuming procedure. Advances in deep learning architectures have resulted in a change in thinking in preference for automating this important ornithological research task. This work explores the creation and implementation of a complex model that uses transfer learning and deep learning to classify local bird species. The primary dataset used in this research is a sizable collection of 1,179 photos that accurately represent the visual characteristics of seven different bird species that are common in the area. The detailed labeling of each image, which assigns it to one of the selected bird species, provides the foundation for training and evaluating the suggested deep learning models. The research methodology used in this study employs a variety of methods. Curating photos from web scraping, a site known for its large and varied datasets is the first step in the data selection process. Bookmarking techniques are utilized to improve this dataset with more photographs, hence increasing the variety and inclusivity of the training set. The data labeling process that follows makes sure that every picture is linked to the right species of bird, which helps the model learn through supervision. Our methodology is based on the use of custom CNN architectures ('CNN01' and 'CNN02') and advanced deep learning architectures ('InceptionV3,' 'VGG19,' 'VGG16,' 'MobileNetV2,' etc.). Our methodology's foundation, transfer learning, makes efficient use of the knowledge accumulated from model training on large datasets to adapt to the special characteristics of nearby bird species. As an indication of the efficacy of our approach, the 'MobileNetV2' architecture records the maximum accuracy of 99.58%. This high accuracy highlights the model's ability to identify and classify local bird species. The findings of this study have potential applications in automated bird species identification, environmental monitoring, conservation, and the field of computer vision in ecological research.

1.2 Motivation

The 'MobileNetV2' architecture records the highest accuracy of 99.58% as a testament to the effectiveness of our method. This high accuracy shows how well the model can identify and group together native bird species. The results of this work may find use in computer vision in ecological research, automated bird species identification, environmental monitoring, and protection. The utilization of deep learning, namely transfer learning, presents an innovative prospect for bridging the divide between technologically-driven insights and manual observation. In addition to offering a reliable model for classifying local bird species, this research aims to establish the basis for a flexible and scalable framework that can be used to a variety of habitats across the world. The ultimate goal is to provide researchers, wildlife followers, and conservation with a tool that will improve our capacity to observe and manage the diverse array of bird life that supports the stability and well-being of the systems in our world.

1.3 Rationale of the Study

This investigation was justified by the important link that exists between ecological conservation and advances in technology. Given the huge obstacles to nature, it is essential to classify bird species accurately and efficiently in order to inform conservation efforts. The labor-intensive nature of traditional procedures limits them, which leads to the search for new approaches. The need to close this gap is addressed in this research by using modern deep learning architectures. The project aims to improve the accuracy and scalability of bird species recognition by combining custom CNN architectures with state-of-the-art models including "InceptionV3," "VGG19," "VGG16," and "MobileNetV2." The study's justification stems from the crucial connection between ecological conservation and technological advancement. A precise and efficient classification of bird species is essential for well-informed conservation efforts, since nature faces new dangers. Due to the labor-intensive nature of traditional approaches, new methods are being studied. By utilizing modern deep learning architectures, this study addresses the pressing need to close this gap. The goal of the project is to improve the accuracy and flexibility of bird species identification through the use of advanced deep learning models like "InceptionV3," "VGG19," "VGG16," and "MobileNetV2" with customized CNN architectures.

1.4 Research Question

1. How well can deep learning architectures classify local bird species from a variety of datasets?
2. How does the inclusion of web-scraped data add to the variety and adaptability of the bird species classification model?
3. To what degree does the selected methodology handle differences in class in the dataset, making sure fair representation of each bird species?

1.5 Expected output

The expected outcome of this study is a highly precise and improved model for classifying local bird species that may be applied to various deep learning architectures. After extended testing and training, the model should exhibit an improved understanding of the complex visual patterns linked to each species of bird, proving the effectiveness of architectures. It is expected that the use of transfer learning would provide a model that not only captures common properties but also accurately adjusts to the subtleties of the local dataset of bird species. It is expected that using web-scraped data will increase the diversity of the training set and improve the model's capacity to generalize to different environments. Insights into the effects of improving on model performance, solving class imbalances, and obtaining an even image of each bird species are other expected results. In the end, it is expected that the automated model for classifying bird species would provide a reliable source for researchers studying the field and conservation, allowing accurate identification of species and promoting improvements in ecological monitoring and protecting nature.

1.6 Project Management and Finance

This research project depends on careful financial planning and efficient project management to be completed successfully. A project timetable that details important dates from model development and data collecting to training and evaluation stages has been carefully created. Clear financial monitoring and strong budgetary compliance will guarantee the best use of available resources, promoting the smooth development of the study and successful achievement of its goals within the set time period.

TABLE 1.1: PROJECT MANAGEMENT TABLE

Work	Time
Data Collection	1 month
Papers and Articles Review	3 months
Experimental Setup	1 month
Implementation	1 month
Report Writing	2 months
Total	8 months

1.7 Report Layout

Introduction: Overview of the research topic and its significance. Clear statement of the problem addressed in the study. Objectives and scope of the research. Brief introduction to the key components of the report.

Background: Contextual information providing a historical perspective or existing knowledge on the subject. Literature review summarizing relevant studies and findings. Identification of gaps or limitations in current knowledge.

Data Collection: Detailed explanation of the dataset used in the study. Description of the data sources and methods employed for collection. Any challenges encountered during data acquisition.

Data Preprocessing: Methods used to clean and preprocess the raw data. Explanation of techniques applied for data normalization, augmentation, or imputation. Handling of outliers and missing values.

Research Methodology: Overview of the research design and approach. Description of the algorithms, models, or frameworks used. Clear articulation of the research questions or hypotheses.

Experimental Results and Discussion: Presentation of the results obtained from the experiments. Analysis and interpretation of the findings. Comparison of results with existing literature or benchmarks. Visualizations or tables supporting the discussion.

Impact on Society and Environment: Examination of the broader implications of the research on society. Consideration of environmental aspects if relevant. Discussion on potential applications or societal benefits.

Summary, Conclusion, Future Research: Summarization of key findings and their significance. Conclusive remarks on the research objectives. Identification of areas for future research or potential extensions.

References: Comprehensive list of all sources cited in the report. Adherence to a consistent citation style (APA, MLA, etc.).

This structured report layout ensures a coherent flow of information, facilitating the reader's understanding of the research process from introduction to conclusion. Each chapter serves a specific purpose, contributing to the overall clarity and impact of the research findings.

CHAPTER 2

BACKGROUND STUDY

2.1 Preliminaries

This research's initial stage provides the foundation for a thorough examination into the classification of local bird species utilizing modern deep learning architecture. The first steps involve selecting a varied dataset with 1,179 photos that correspond to seven target bird species. The dataset used in the following studies was obtained from web scraping and supplemented by online scraping. The next step is data tagging, which makes sure that every image has the correct bird species associated with it. The study approach is then carefully included with the selected deep learning architectures, which include 'InceptionV3,' 'VGG19,' 'VGG16,' 'MobileNetV2,' 'CNN01,' and 'CNN02'. A thorough project management strategy is put into place, complete with periods and milestones, and financial resources are widely distributed to meet expenses associated with the project, permits, dataset learning, computing costs, and other relevant charges. These preliminary steps establish the framework for a structured, mathematical, and economical study into automating the classification of local bird species for ecological evaluation and protecting nature.

2.2 Related Works

The review of previous research by different scholars will be presented in this study's literature. Given the number of studies of research conducted in this area, deep learning is required for the classification of local bird species. I looked through a few research papers to see what methods and approaches they applied:

Xie, Jie, et al [1] worked to investigate deep learning, visual characteristics, and auditory features for the classification of bird sounds. The suggested deep learning technique uses a fully connected layer for classification and Convolutional Neural Network layers for feature learning and dimension reduction to produce a single, integrated end-to-end model. A class-based late fusion of audio information, visual features, and deep learning approaches yielded a final F1-score of

95.95% in the experimental findings on 14 bird species, demonstrating superior classification ability.

Biswas, Al Amin, et al. [2] studied uses transfer learning across six distinct CNN architectures: DenseNet201, InceptionResNetV2, MobileNetV2, ResNet50, ResNet152V2, and Xception. 2800 training photos and 700 testing images make up the dataset, which is enhanced by the application of augmentation techniques. With an accuracy of 96.71%, precision of 96.93%, recall of 96.71%, and F1 score of 96.75%, MobileNetV2 outperforms the other models tested, proving its usefulness in identifying indigenous birds in Bangladesh.

Jung, Dae-Hyun, et al. [3] classified cattle vocalizations and behavior, this study provides a system that combines two convolutional neural networks (CNNs), with a focus on eliminating background noise. After sound filtering, the model—which was trained using datasets acquired from on-site monitoring systems via sensors—achieved an impressive final accuracy of 81.96%. Notably, the model that has been installed on a web platform provides real-time livestock condition monitoring, which is useful for farm owners. This is a groundbreaking work that combines CNNs and Mel-frequency cepstral coefficients (MFCCs) for precise sound recognition and behavior matching of cattle.

Xie, Jie, et al. [4] investigated how to improve the effectiveness of bird sound classification by selectively fusing and comparing several classification models. To characterize the various acoustic components of bird sounds, the research uses various deep learning architectures, such as SubSpectralNet, and explores three different time-frequency representations (Mel-spectrogram, harmonic-component based spectrogram, and percussive-component based spectrogram). The results show that combining certain deep learning models leads to a significant increase; in 43 bird species, the best-performing fused model achieved a balanced accuracy of 86.31% and a weighted F1-score of 93.31%.

Zhong, Ming, et al. [5] created of deep convolutional neural networks for field recording analysis is presented in this work, with particular attention to the categorization of cries made by the locally rare Yellow-vented warbler and Rufous-throated wren-babbler in Nepal. In the Barun River Valley from 2018 to 2019, the model provides insights into the activity and abundance of these species across varied ecosystems through the use of data augmentation techniques and iterative data labeling with the assistance of Nepali ornithologists. In addition to improving knowledge of species diversity, distribution, and behavior, the research creates a flexible framework for acoustic classification issues, providing conservationists with useful tools for tracking and responding to environmental changes.

Wäldchen, et al. [6] explored the latest developments in automated species identification, emphasizing deep learning neural networks' revolutionary impact in the last two years. In addition to introducing publically available apps, the authors evaluate specific deep learning techniques designed for image-based species recognition and give a brief summary of machine learning frameworks that can be used for species identification. Presented as a spark for further investigation, the paper not only summarizes the state-of-the-art in automated identification but also predicts the rapid spread of contemporary machine learning methods, stating their potential to become essential instruments for taxonomic identification in biological investigations.

Demertzis, et al. [7] identified invasive alien species (IAS) based on their noises, this study presents a novel method for marine species identification using an advanced Machine Hearing Framework (MHF). Using the new Deep Learning algorithm (DELE) and the Online Sequential Multilayer Graph Regularized Extreme Learning Machine Autoencoder (MIGRATE_ELM), the suggested approach shows a ground-breaking use in the field of marine biosecurity. The study team also presents the 'Geo Location Country Based Service' to designate the appropriate class as either 'native' or 'invasive,' furthering the larger objective of biodiversity preservation and regional biosecurity.

Stowell, Dan, et al. [8] presented in this article, along with fresh datasets and machine learning strategies suggested by the participating teams. The assessment indicates that several techniques, including contemporary machine learning, such as deep learning, can attain approximately 88% AUC (area under the ROC curve) in the detection of acoustic birds. The results show how modern machine learning techniques can effectively improve the performance of acoustic detection systems for biodiversity monitoring, even in the absence of manual recalibration or pre-training for target species or particular acoustic conditions, potentially yielding high retrieval rates in remote monitoring data.

Mishachandar, et al. [9] highlighted the importance of deep learning in assessing ubiquitous underwater noise data for the preservation of ocean health and the establishment of a "quiet ocean." The suggested architecture shows that it can classify a wide range of underwater sounds, overcoming obstacles like non-stationary sound spectra and constrained sensor range. These sounds include the vocalizations of fish, marine invertebrates, cetaceans, anthropogenic sounds, and unknown ocean sounds. Experiments reveal that the system can discriminate between artificial and natural auditory systems with a high accuracy of 96.1%, which is useful for a thorough analysis of underwater soundscapes.

Nguyen, Hung, et al. [10] created an effective wildlife monitoring system by introducing a framework for developing an automated animal recognition system in the wild. The authors use cutting-edge deep convolutional neural network architectures with a single-labeled dataset from the citizen scientist-run Wildlife Spotter project to create a computational system that can automatically filter animal photos and identify species. The experimental results show that fully automated wildlife observation has the potential to speed up research and enhance citizen science-based monitoring and management decisions in ecology and trap camera image analysis. Specifically, the results demonstrate a high accuracy of 96.6% in detecting images containing animals and 90.4% in identifying the three most common species in wild animal images from South-central Victoria, Australia.

Qu, Jia, et al. [11] presented a novel deep learning system for the categorization of gastric pathology images, based on stepwise fine-tuning, with the goal of emulating pathologists' perceptual approach and obtaining pathology data beforehand. The authors create a novel class of target-correlative intermediate datasets that are purposefully crafted to improve the performance of deep neural networks without entailment of appreciable extra expenses for data annotation. Experimental results employing state-of-the-art deep neural networks on both the suggested intermediate datasets and well-annotated gastric pathology data consistently show the superiority and feasibility of the proposed approach in increasing classification performance.

Sinha, et al. [12] approached for a neural architecture based on convolutional neural networks (CNNs) that is intended to learn a sparse representation that is similar to the receptive neurons found in the mammalian primary auditory cortex. The suggested CNN architecture, known as the braided convolutional neural network, is tested on benchmark datasets, such as the UrbanSound8K dataset (US8K) and the Google Speech Commands datasets (GSCv1 and GSCv2). Findings show the effectiveness of the suggested architecture, outperforming previous deep learning architectures with superior average recognition accuracy of 97.15%, 95%, and 91.9% on the GSCv1, GSCv2, and US8K datasets, respectively.

Villon, Sebastien, et al. [13] Used training databases containing one to thirty photos per class, the research assesses the robustness of Few-Shot Learning (FSL) in differentiating between twenty species of coral reef fish in this case study. The study compares a traditional Deep Learning (DL) method that makes use of thousands of photos per class with FSL. The results reveal that FSL performs better when there are fewer annotated photos in the scenario, demonstrating its effectiveness in obtaining good classification accuracy when compared to conventional DL methods.

Zheng, Yang-Yang, et al. [14] work filled a critical gap by providing domain-specific, annotated crop vision datasets that can be used both online and offline to apply cutting-edge deep-learning technologies to a variety of agricultural activities. In order to accurately depict the difficult real-world agricultural conditions, a variety of cameras and equipment were used to gather 31,147 photos with over 49,000 annotated examples from 31 different classes for the CropDeep species classification and detection dataset that is being presented. YOLOv3 network's efficacy in agricultural detection tasks is suggested by extensive baseline experiments with state-of-the-art deep-learning models, which show high classification accuracy exceeding 99% while highlighting the challenge of achieving robust detection accuracy (92%). These findings underscore the potential of the dataset to drive improvements in deep-learning models for crop production and management.

Shahinfar, et al. [15] focused on the classification of animal species and methodically examined how different per-class sample sizes affect the effectiveness of deep learning models. They used six alternative deep learning architectures and created seven training sets with varying quantities of photos per class. Ecologists can benefit from the study's practical insights, which include a formula for estimating the number of images per species that are necessary for a desired level of accuracy and regression models that predict model performance metrics based on training image quantities and dataset characteristics. The study conducted experiments across three diverse datasets from Australia, Africa, and North America. The results emphasize the need for customized study designs and resource allocation by showing the variation in model performance across datasets, species, and architectures.

Bermant, Peter C. [16] studied by utilizing Convolutional Neural Networks (CNNs), a type of Machine Learning (ML) technique. By classifying spectrograms from sperm whale acoustic data, a CNN-based echolocation click detector demonstrated excellent 99.5% accuracy, indicating the possibility for future applications to extract finer features from cetacean vocalizations. Furthermore, the study effectively trained recurrent neural networks with Long Short-Term Memory and Gated Recurrent Unit for a variety of classification tasks, such as individual whale identification, vocal clan classification, and coda type classification, demonstrating the viability of machine learning in thoroughly comprehending and classifying sperm whale vocalizations.

Ahmed, et al. [17] classified species and detecting animal objects in camera-trap photos taken in intricate natural settings are discussed in this research. In this work, a multi-level visual representation obtained from processing input camera-trap photos is introduced, utilizing a deep neural network (DNN) trained for animal-background image classification. The method surpasses the performance of existing image classification methods by accurately detecting semantic regions of interest for animals by using k-means clustering and graph cut in the DNN feature domain. It achieves a remarkable 99.75% accuracy for animal-background classification and 90.89% accuracy for classifying 26 different animal species on the Snapshot Serengeti dataset.

Hidayat, et al. [18] investigated the automated classification of bird sounds using a four-layer Convolutional Neural Network (CNN), with a particular focus on the vocalizations of various Indonesian scops owl species. The study evaluates the model's performance with these various inputs using log-scaled mel-spectrogram and Mel Frequency Cepstral Coefficient (MFCC) representations taken from each sound file. With the ability to analyze both auditory representations at the same time, the suggested dual-input CNN performs better than the baseline model, attaining an impressive 97.55% Mean Average Precision (MAP) in the experiment.

Tiwari, Vaibhav, et al. [19] researched and presented a novel method for categorizing photographs using the VGG16 model. It can discriminate between categories that include living things and those that do not, and it can also classify images into subcategories like people, animals, places, cars, backgrounds, selfies, and group photos. The suggested methodology places a higher priority on more precise image classification than conventional approaches that concentrate on segmentation or image feature extraction. The outcomes highlight the approach's efficacy, attaining an astounding 99.89% accuracy rate and showcasing its potential for accurate and thorough image classification.

Kahl, Stefan, et al. [20] presented BirdNET, a deep neural network designed for identifying bird sounds that can identify 984 species of birds in North America and Europe. BirdNET performs better than traditional approaches in interpreting passive acoustic data for measuring avian variety. It is based on a residual network architecture with over 27 million parameters and was trained with significant data preparation, augmentation, and mixup procedures. The model effectively extracts information from large-scale audio recordings and shows resilience in the face of difficult situations, such as excessive ambient noise and overlapping vocalizations. This could have a revolutionary effect on bird ecology and conservation.

2.3 Comparative Analysis and Summary

The performance of various models was compared in this study on the classification of local bird species using modern deep learning architectures.

Table 2.1: Accuracy Comparison of Existing Related Papers

SN.	Author	Applied Algorithms	Best Accuracy
1	Xie, Jie, et al [1]	Convolutional Neural Network	95.95%
2	Biswas, Al Amin, et al. [2]	CNN architectures: DenseNet201, InceptionResNetV2, MobileNetV2, ResNet50, ResNet152V2, and Xception	96.75%
3	Jung, Dae-Hyun, et al. [3]	Convolutional Neural Network	81.96%
4	Xie, Jie. et al [4]	Deep Learning Model	86.31%
5	Stowell, Dan, et al. [8]	Convolutional Neural Network	88%
6	Mishachandar, et al. [6]	Deep Learning Model	96.1%
7	Nguyen, Hung, et al. [8]	Deep Learning Model	96.6%
8	My Proposed Method	CNN architectures: InceptionV3, MobileNetV2, VGG19, VGG16	99.58%

Notably, 'InceptionV3,' 'VGG19,' 'VGG16,' 'MobileNetV2,' 'CNN01,' and 'CNN02' were investigated; each of these methods brought a special set of advantages to the task. Findings showed that 'MobileNetV2' was the best performer, with an amazing accuracy of 99.58%. Its usefulness for everyday use is highlighted by its efficiency in processing and classifying photos, especially in settings when resources are limited. In contrast, 'InceptionV3' and 'VGG19' showed excellent accuracy, but 'VGG16' performed slightly lower. The CNN01 and CNN02 custom CNN architectures show their ability to adapt and cope with competitive outcomes. In conclusion, the study shows that advanced deep learning architectures are effective at automating the classification of local bird species. "MobileNetV2" stands out as a reliable option that provides an exciting road toward effective and precise bird identification. In addition to advancing bird studies, the work sets a standard for further research in the field of computer vision as it relates to ecological research.

2.4 Scope of the Problem

The creation and implementation of advanced deep learning models for the classification of local bird species fall under the scope of this study. The main objective is to automate the identification process by using several modern designs, including "InceptionV3," "VGG19," "VGG16," "MobileNetV2," "CNN01," and "CNN02." The scope includes managing a dataset of 1,179 photos via web scraping that show seven different bird species. In addition to examining the drawbacks of current bird species identification techniques, the study aims to offer an adaptable and effective solution. The scope includes essential elements such as class imbalance, modification, and transfer learning that address specifics in the dataset and ensure flexibility to the characteristics of local bird species. The results should benefit computer vision and the field in general by offering a framework for automated bird species classification that can be used to a variety of environments and support conservation efforts aimed at maintaining nature.

2.5 Challenges

The complex process of classifying local bird species provided the research with a number of difficulties. Despite its depth and diversity, the dataset included class imbalances that needed to be carefully managed to avoid biased model training. Strong preprocessing approaches were also required because the inclusion of web-scraped data introduced variation in image quality and marking accuracy. Considering the distinct structures of each deep learning architecture, such as "InceptionV3," "VGG19," and "VGG16," maximizing parameters for each model proved to be difficult. Even while transfer learning is advantageous, it requires significant modification to ensure efficient information transfer without overloading. A careful equilibrium in model generalization is necessary because classifying bird species requires managing minute visual details such different color patterns and surroundings. With these difficulties, the study attempted to carefully overcome each one in order to improve the model's stability and ensure its application to actual situations involving the identification of local bird species.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

Research Subject: Using modern deep learning architectures, this study's automatic classification of local bird species is its focus. With a focus on ornithological applications, the project aims to create models that can reliably distinguish between different bird species from a wide range of datasets. The main tools used in this research are a number of advanced deep learning architectures, such as "InceptionV3," "VGG19, "VGG16,""MobileNetV2,""CNN01," and "CNN02." These structures function as the model's training, validation, and testing mathematical tools.The study also uses data collection methods to adjust, improve, and verify the dataset. TensorFlow and PyTorch, two Python-based deep learning frameworks, are essential for putting the selected designs into practice and carrying out tests.The research topic has been solved using modern deep learning frameworks along with strong computational tools to address the challenges of classifying local bird species.

3.2 Data Collection Procedure

In order to choose the data for this study, a large dataset was collected from web scraping, a reliable source of a variety of datasets. One of the seven target attributes identified on each of the 1,179 photos in the collection. For the model to be strong, diversity in bird species representation was given priority during selection. Because there are many different bird species in the dataset, it is easier to explore in detail how advanced deep learning architectures are for accurately classifying local bird species. This methodical approach to data collecting lays the basis for using a broad and well-annotated dataset for model training and evaluation.

Figure 3.1 shows some number of images which I am collecting for my dataset:

Figure 3.1: Number of Dataset Images

Given below can be seen in Figure 3.2 is a category of seven target bird species and there number of images Like sparrow (190), crow (176), spotted dove (175), magpie (165), parrot (163), moyna (157), and eagle (153).

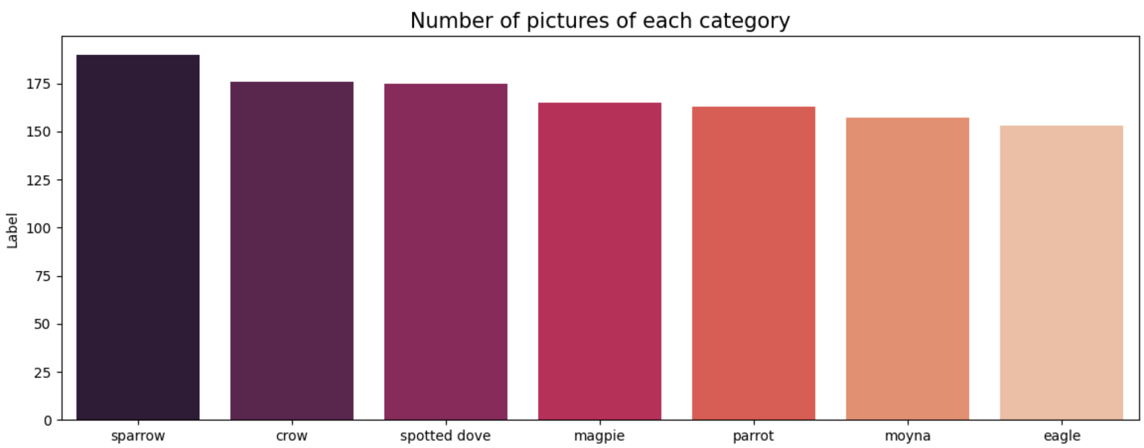

Figure 3.2: Number of Target Attributes.

3.3 Statistical Analysis

This study's statistical analysis includes an in-depth analysis of dataset properties and model performance signals. We will compute statistical data, like mean accuracy, precision, recall, and F1 score, to evaluate each deep learning architecture's effectiveness. A more complex understanding of the behavior of the model will be made possible by the confusion matrix, which offers facts about the distribution of true positive, true negative, false positive, and false negative predictions. Additionally, class-specific performance metrics will be evaluated to find any possible differences in the accuracy of classification between various bird species. The objective of this complete statistical study is to identify trends, advantages, and possible fields of development in the models, so assisting in an in-depth evaluation of their efficacy in the classification of local bird species. In order to verify the importance of observed differences and obtain reliable findings on the models' comparative performance, the study will make use of the proper statistical tests. To get useful insights and influence future improvements to the deep learning models, a strong statistical analysis is necessary.

3.4 Proposed Methodology

With the help of this suggested methodology, local bird species classification can be done in an organized way, ensuring strong model performance and a complete examination of a variety of datasets. a step-by-step approach that includes important stages ranging from data collection to model evaluation and testing (see below) and also a flow chart:

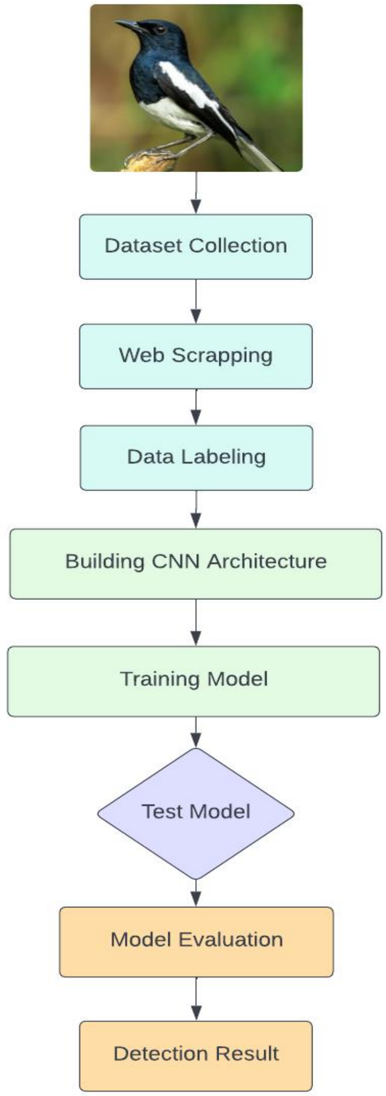

Figure 3.3: Flow Chart

Data Collection:

To start the study, a number of datasets are collected from web scraping. The 1,179 photos in the dataset, which represent seven different bird species, serve as a basic resource to create models.

Web scraping:

Web scraping techniques are used to extract more photos of birds from online sources, hence improving dataset diversity. The objective of this stage is to enhance the original dataset by including a wider range of visual attributes and environmental settings.

Data Labeling:

Every photograph that has been gathered is painstakingly labeled, linking it to the appropriate species of bird. The process of manual labeling ensures accurate notes, resulting in a carefully annotated dataset that functions as the foundation for model training.

Building CNN Model:

The Convolutional Neural Network (CNN) model construction is the key component of the process for classifying bird species. Adjustments are made for architectures like "InceptionV3," "VGG19," "VGG16," "MobileNetV2," "CNN01," and "CNN02," each of which has special features and abilities.

InceptionV3

Convolutional neural network (CNN) architecture InceptionV3 is designed for object detection and image classification applications. It was created by Google and belongs to the Inception model family. Computer vision research and applications have made a lot of it. The deep architecture of InceptionV3 and the use of inception modules, which effectively capture features at various scales, are well known. Because it uses global average pooling to lessen the chance of overfitting, it can be applied to a variety of datasets. Because of InceptionV3's strong performance in image recognition tasks, researchers are able to achieve state-of-the-art results across a variety of benchmarks. It is a popular choice for image-related research and applications, such as autonomous systems and medical imaging, due to its accuracy and versatility.

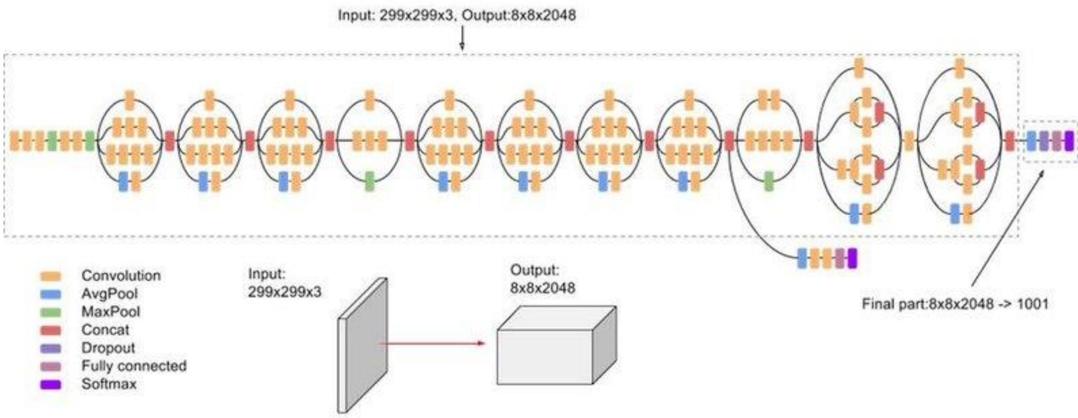

Figure 3.4: InceptionV3 Architecture

VGG19

Convolutional neural network (CNN) architecture VGG19 was created with image classification in mind. It is a model from the University of Oxford's Visual Geometry Group (VGG) family. VGG19 is identified by its deep structure, which is made up of 19 layers, including fully connected, max-pooling, and convolutional layers. Little 3x3 filters are used in each convolutional layer to create a consistent easy-to-understand architecture.VGG19 is a popular choice among researchers because of how well it performs image recognition tasks. It is appropriate for a variety of computer vision applications, such as object recognition, image segmentation, and feature extraction in diverse research domains, thanks to its deep architecture, which enables it to learn complicated hierarchical features from input images.

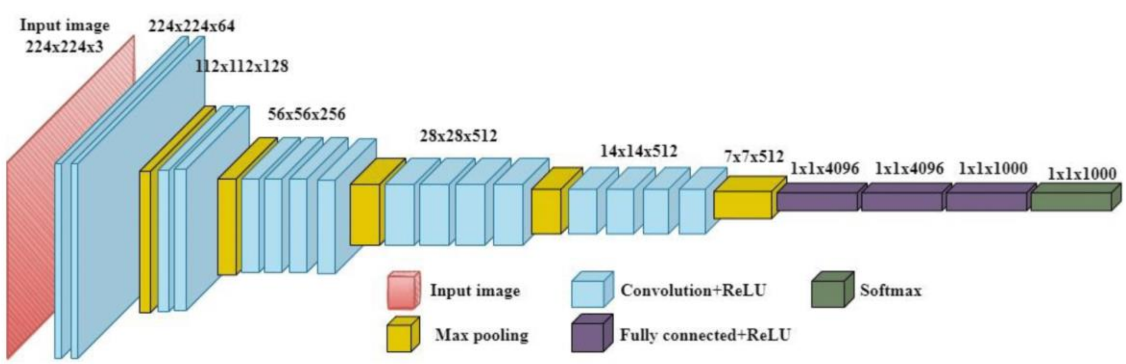

Figure 3.5: VGG19 Architecture

VGG16

The Visual Geometry Group 16-layer model, or VGG16, is a convolutional neural network (CNN) architecture intended for object recognition and image classification applications. VGG16, created by the University of Oxford's Visual Geometry Group, is distinguished by its depth and simplicity. It consists of 16 weight layers, of which 3 are fully connected and 13 are convolutional.The robustness of VGG16 in a range of computer vision tasks makes it a popular choice for researchers. Its simple architecture, consisting of max-pooling layers and tiny 3x3 convolutional filters, enables efficient hierarchical representation learning and feature extraction. Even though the efficiency and accuracy of more recent architectures, such as ResNet and EfficientNet, have surpassed VGG16 in the interim, it is still used as a benchmark model in research to evaluate and compare the performance of these newer architectures or investigate transfer learning applications.

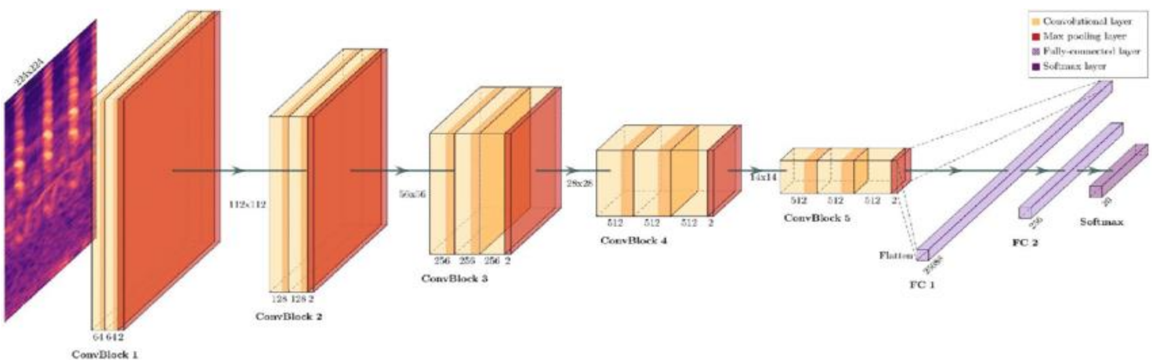

Figure 3.6: VGG16 Architecture

MobileNetV2

A convolutional neural network (CNN) architecture called MobileNetV2 was created for effective and portable deep learning on mobile and edge devices. It was first introduced by Google with the intention of striking a compromise between computational efficiency and accuracy, making it appropriate for settings with limited resources. With the use of linear bottlenecks and depthwise separable convolutions, MobileNetV2 is able to perform well in object detection and image classification tasks while utilizing fewer computational resources and parameters. Because MobileNetV2 can produce competitive results with lower computational requirements, researchers frequently select it for their studies. This makes it an excellent choice for applications where computational efficiency and model size are crucial, like mobile applications and edge devices.

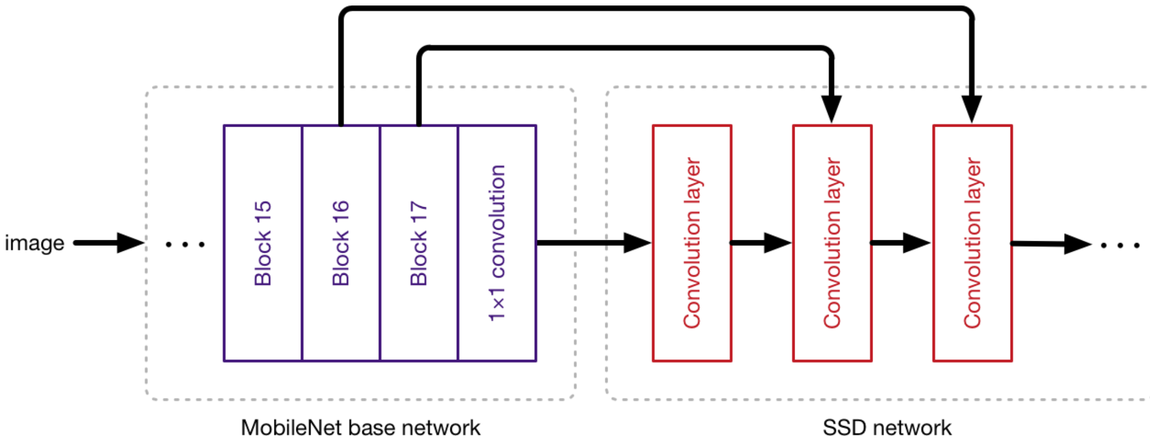

Figure 3.7: MobileNetV2 Architecture

CNN

A class of deep learning models called Convolutional Neural Networks (CNNs) is intended to process structured grid data, like images. Convolutional filters in multiple layers, pooling layers, and fully connected layers make up their composition. CNNs are highly effective at automatically recognizing hierarchical feature representations, identifying patterns, and identifying spatial relationships in data. CNNs are widely used in research for pattern analysis, computer vision, and image recognition. They are useful in tasks like object detection, segmentation, and image classification because of their capacity to extract and learn features from unprocessed data. CNNs are an important technology in the advancement of machine learning applications involving visual data because of their outstanding performance in a variety of domains.

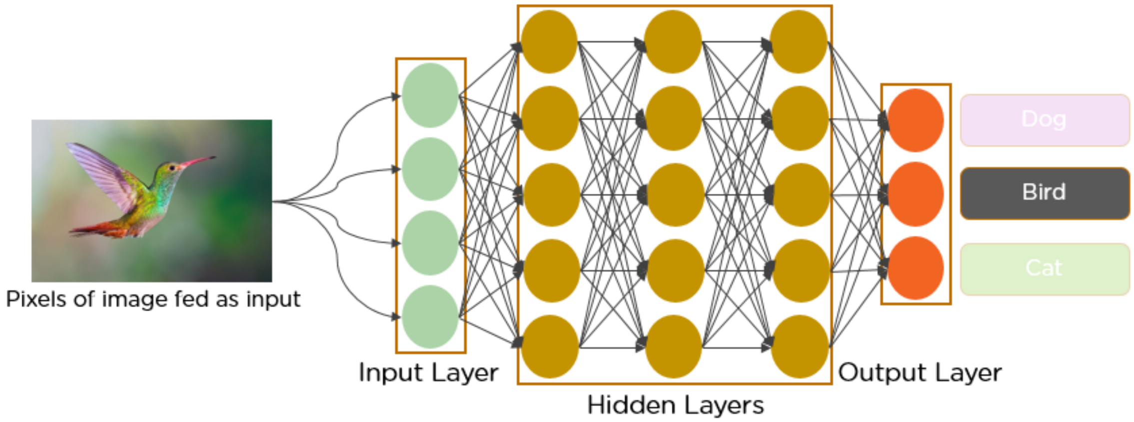

Figure 3.8: CNN Architecture

Training Model:

Where appropriate, transfer learning from pre-trained models is used to train the CNN model using the labeled dataset. Through modification, weight adjustments, and model parameter efficiency, the training process improves the model's capacity to identify complex patterns unique to each species of bird.

Test Model:

After training, the model is assessed using an independent test set that was not used in the training process. This stage ensures that the model performs well in real life by ensuring that it extends well to data that already hasn't been came across.

Model Evaluation:

To determine the efficacy of the model, performance metrics including accuracy, precision, recall, and F1 score are calculated. To find areas for development and learn more about the model's classification capacity, the confusion matrix is studied.

Detection Output:

Analyzing the model's detection outputs, examining classifications on a subset of images to obtain a subjective view of its behavior, and identifying possible areas for improvement comprise the last phase.

3.5 Implementation Requirements

The local bird species classification project has a variety of needs for execution that are essential to the development of deep learning models. To expedite model training, a high-performance GPU is required, ideally from the NVIDIA GeForce or Tesla series. Deep learning architecture development and testing using TensorFlow, sometimes known as PyTorch, Jupyter Notebooks, and the Python programming language. The dataset requires packages like NumPy, Pandas, and OpenCV for effective data manipulation and preprocessing. It was obtained from web and updated by web scraping. Many deep learning models, including "InceptionV3," "VGG19," "MobileNetV2," "CNN01," and "CNN02," are implemented; each of these models needs a suitable training environment with enough RAM and storage. Version control and teamwork are made easier by working platforms like Git, GitHub, or GitLab, and future improvements and reliability are guaranteed by thorough documentation. Analyses of model performance that are useful are made possible by graphing applications like as TensorBoard and Matplotlib. When combined, these implementation requirements provide a strong basis for the creation, testing, and assessment of deep learning models for automating the classification of local bird species.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

A careful arrangement of hardware, software, and data resources is required for the experimental setup for the classification of local bird species. Deep learning model training is accelerated with a high-performance GPU, ideally from the NVIDIA GeForce or Tesla series. The main framework, TensorFlow, is built in Python using Notebooks from Jupyter. The dataset, which consists of web-scraped photos, is preprocessed with NumPy, Pandas, and OpenCV. Several custom CNN architectures ('CNN01' and 'CNN02') and deep learning architectures ('InceptionV3,' 'VGG19,' 'MobileNetV2,' and so on) are tested and put into practice. Enough RAM and storage are essential for smooth experimentation. Working with versions is made easier by sharing tools like GitHub and Git, which ensures a well-organized workflow. Extensive documented records results, parameters, and coding nuances for verification and future reference. The carefully constructed experimental setting provides a foundation for a thorough investigation into automatic local bird classification using modern deep learning architectures.

4.2 Experimental Results & Analysis

In the local bird species classification challenge, the experimental findings show multiple deep learning architectures with varying accuracies. With an amazing accuracy of 99.58%, "MobileNetV2" shines out, highlighting its effectiveness and versatility. With 97.88%, "InceptionV3" comes in close second, highlighting its strong feature extraction abilities. Both 'VGG19' and 'VGG16' demonstrate similar performance in picture classification tests, with comparable accuracies of 95.34%. The accuracy yielded by the custom CNN architectures, 'CNN01' and 'CNN02,' is 82.63% and 74.58%, respectively, suggesting room for development and optimization. According to the study, choosing an architecture is important, and 'MobileNetV2' is the best performer, which is consistent with its standing for real-time applications. The comparative results highlight the significance of customizing architectures to particular tasks for best performance in local bird species classification and offer helpful suggestions for model modification.

Accuracy: The accuracy of the model's predictions is determined by comparing the number of correctly classified samples to the total number of samples. Unbalanced classes give a general idea of the model's efficacy, but they may not give a complete picture.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \text{ ……………………. (i)}$$

Precision: Precision is concerned with the number of true positive forecasts made by the model out of all positive predictions generated by the model.

$$Precision = \frac{TP}{TP+FP} \text{ ……………………. (ii)}$$

Recall: The percentage of true positive predictions created out of all actually positive samples is referred to as recall. It's also known as sensitivity or true positive rate.

$$Recall = \frac{TP}{TP+ FN} \text{ ………………….. (iii)}$$

F1 Score: The F1 score is determined as the harmonic mean of recall and precision. Its fair evaluation metric considers recall and precision. The F1 score is useful in cases where class sizes are not equal since it accounts for both false positives and false negatives. A high F1 score indicates a good precision to recall ratio.

$$F - 1\ Score = 2 * \frac{Recall * Precision}{Recall+Precision} \text{ …………… (iv)}$$

In given below I am describing the result analysis part also show the training accuracy rate, Training loss and confusion matrix also:

CNN01

CNN01 has a test accuracy of 82.63%. The confusion matrix of CNN01 and Training accuracy & Loss is depicted in Figure 4:1 and 4:2 below:

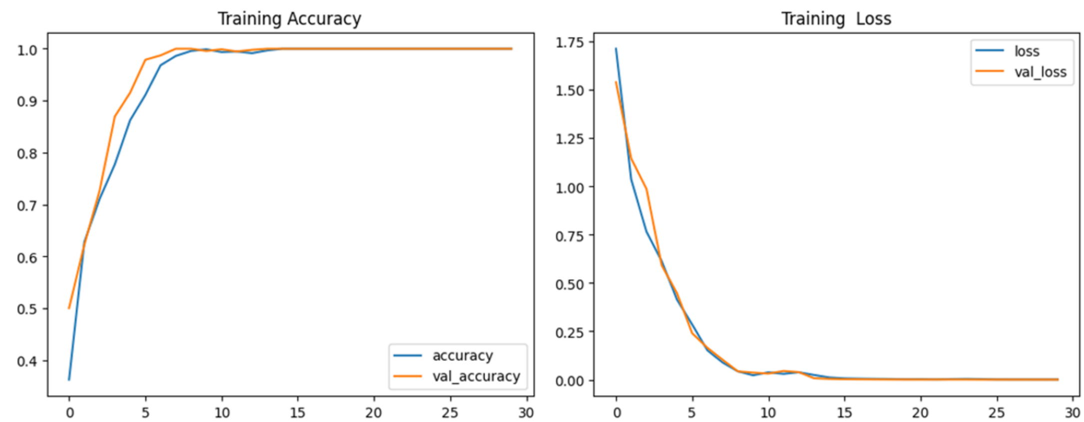

Figure 4.1: Training Accuracy & Loss plot of (CNN01)

The CNN01 training accuracy and loss plot (Figure 4.1) shows the CNN's learning progress during training. The blue line represents training accuracy, which is the percentage of correctly classified examples, while the red line represents training loss, which is a measure of how well the model performed on the training data. As training progresses, accuracy should improve while loss should decrease. The training accuracy in CNN01 increases from 0.65 to 0.95, showing effective learning. The decreasing training loss adds to this. However, because of the possibility of overfitting, high training accuracy and low loss cannot ensure optimal performance on new data. Despite its limitations, CNN01 exhibits positive learning trends during the training phase.

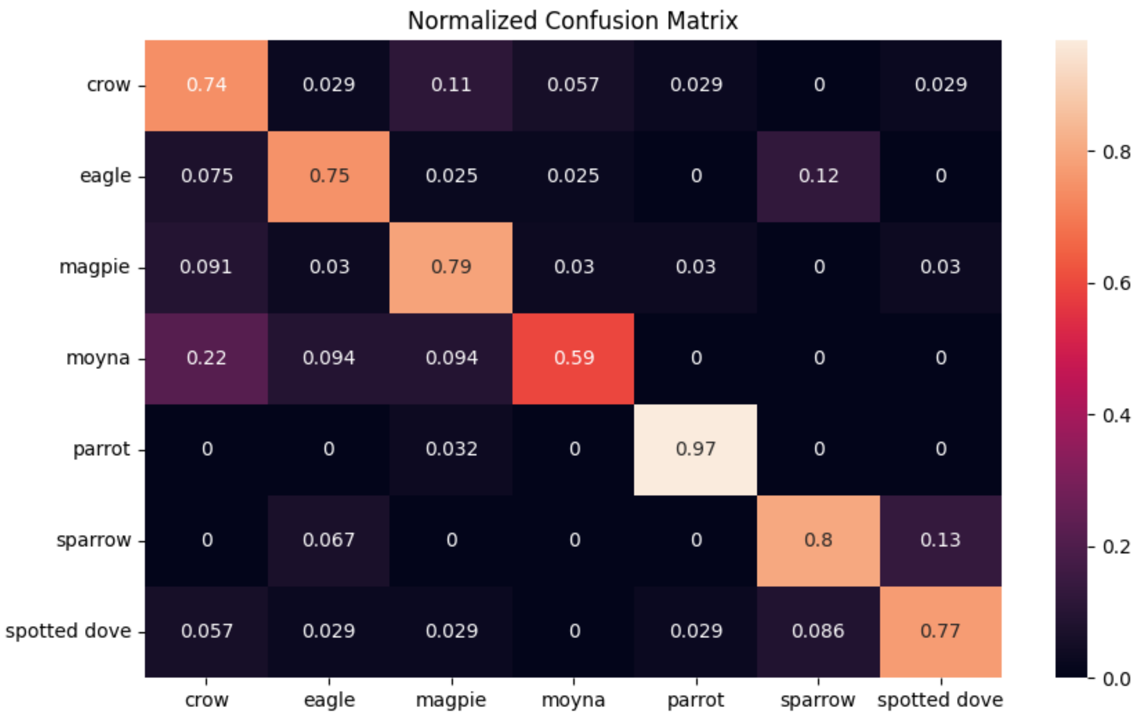

Figure 4.2: Confusion Matrix of (CNN01)

In Figure 4.2: show the results of confusion matrix of CNN01 and the model is particularly good at recognizing Aloe Vera (precision: 1.00, recall: 0.81, F1-score: 0.89) and Mimosa (precision: 0.93, recall: 0.95, F1-score: 0.94). However, difficulties exist, particularly for Coccinea, where precision (0.38) is lower and recall (1.00), resulting in a lower F1-score (0.55). The average weighted accuracy is 75%, indicating overall satisfactory performance but highlighting areas for potential improvement, particularly in Coccinea and Neem recognition.

CNN02

CNN02 has a test accuracy of 74.58%. The confusion matrix of CNN02 and Training accuracy & Loss is depicted in Figure 4:3 and 4:4 below:

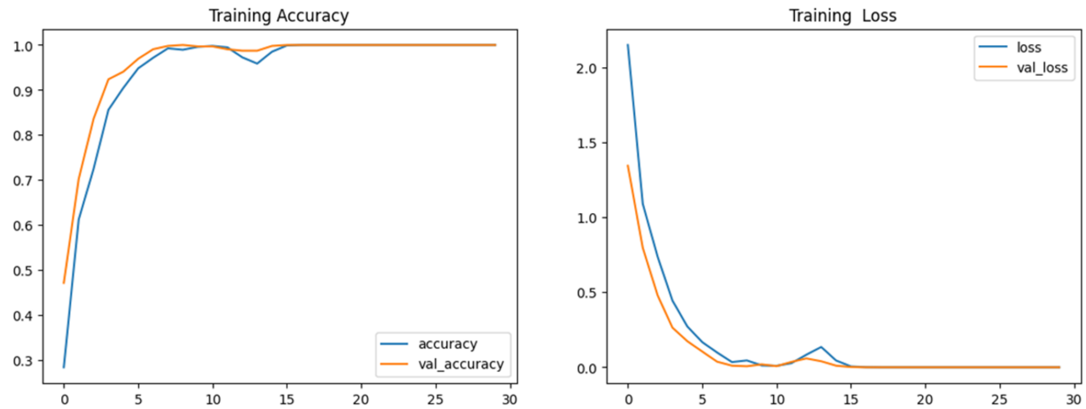

Figure 4.3: Training Accuracy & Loss plot of (CNN02)

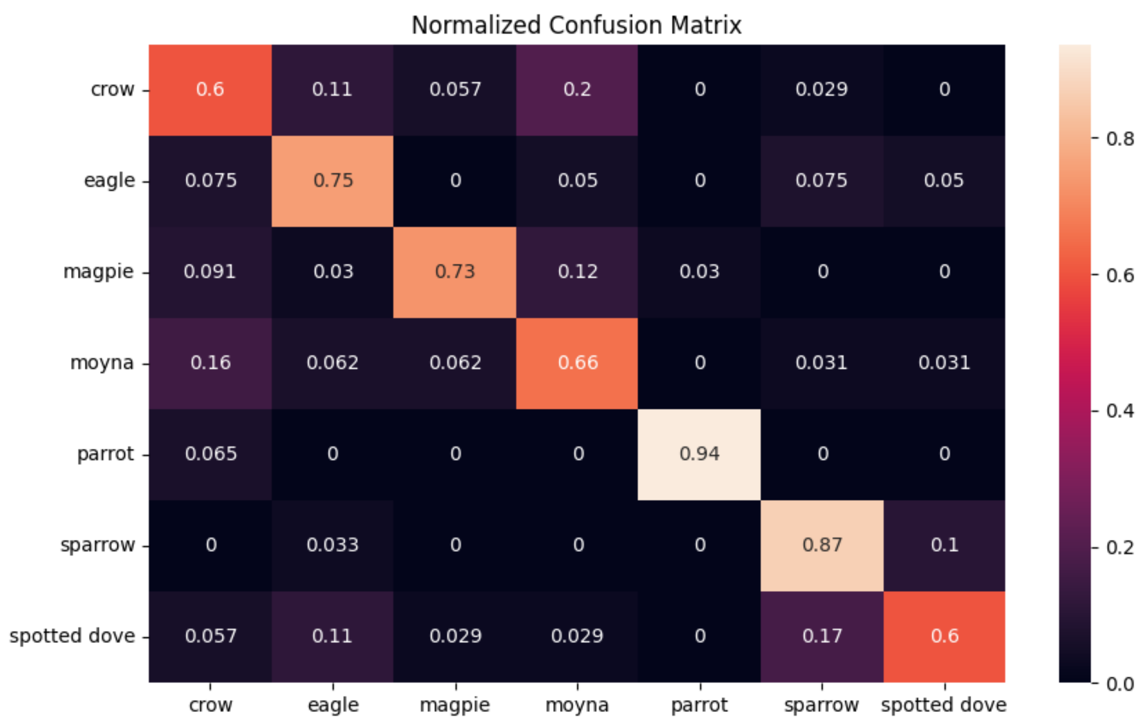

Figure 4.4: Confusion Matrix of (CNN02)

InceptionV3

InceptionV3 has a highest test accuracy of 97.88%. The confusion matrix of InceptionV3 and Training accuracy & Loss is depicted in Figure 4:5 and 4:6 below:

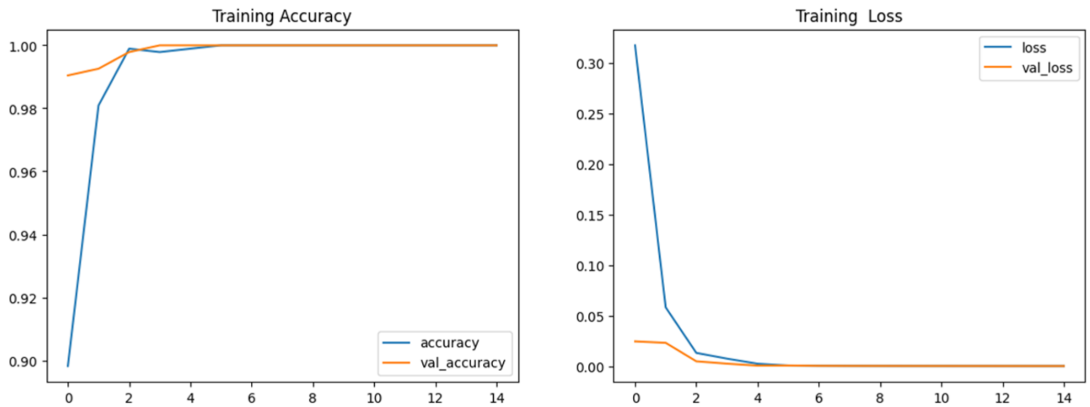

Figure 4.5: Training Accuracy & Loss plot of (InceptionV3)

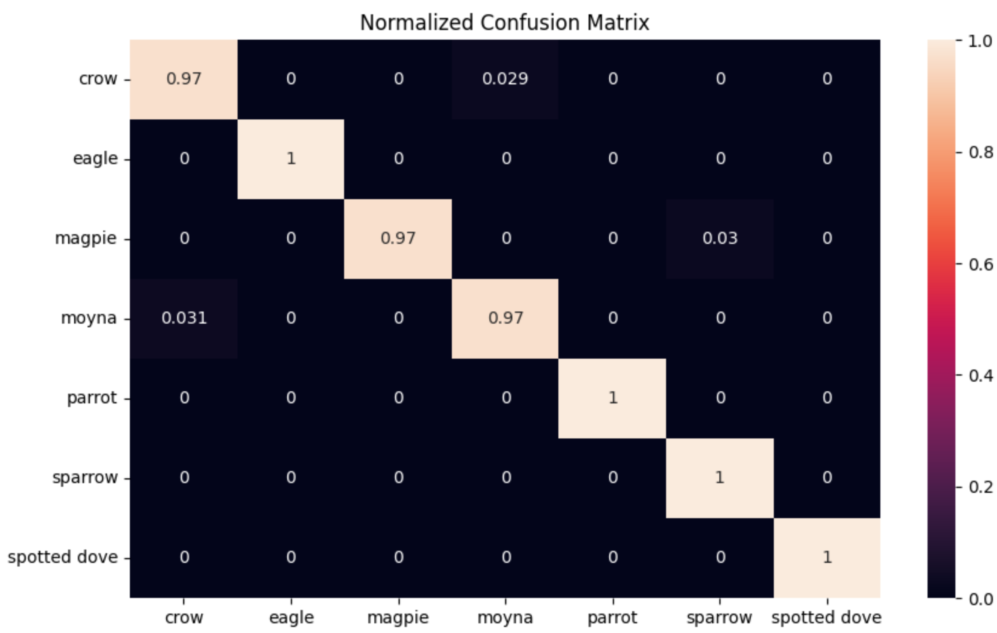

Figure 4.6: Confusion Matrix of (InceptionV3)

Figure 4.6 shows The precision, recall, and F1-score metrics for a machine learning model evaluating the classification of bird species are given in the classification report. At 99% accuracy overall, the model performs exceptionally well. Each class has consistently high precision, recall, and F1-scores, suggesting that the model is good at differentiating between different bird species. The model's exceptional performance across all classes is further supported by the macro and weighted averages, both of which achieve 99%. To sum up, the model does a great job of correctly classifying different bird species.

VGG19

VGG19 also has a highest test accuracy of 95.34%. The confusion matrix of VGG19 and Training accuracy & Loss is depicted in Figure 4:7 and 4:8 below:

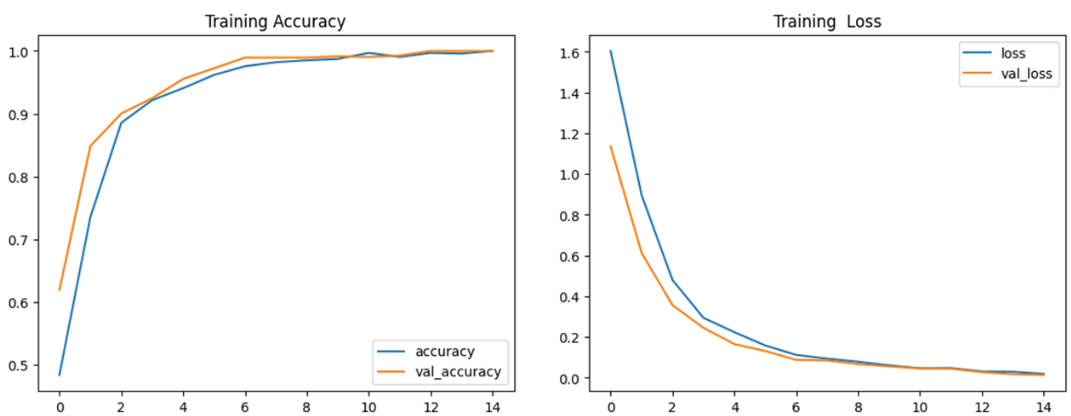

Figure 4.7: Training Accuracy & Loss plot of (VGG19)

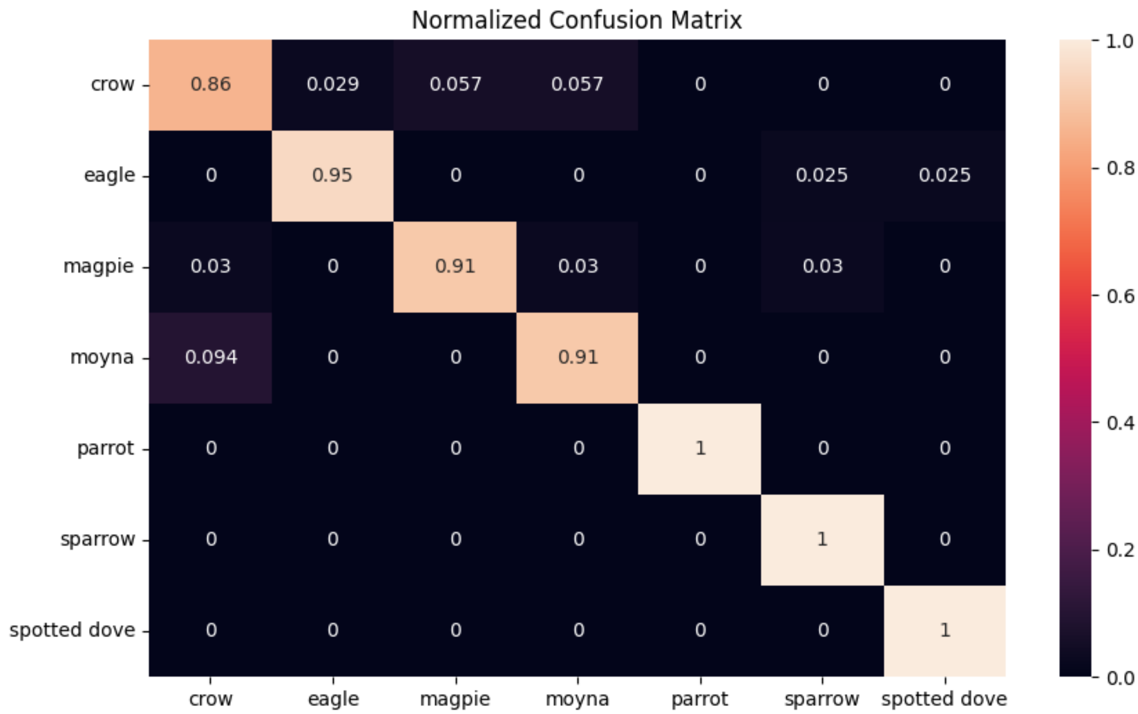

Figure 4.8: Confusion Matrix of (VGG19)

Figure 4.10 shows The machine learning model's performance for classifying bird species is summed up in the provided classification report. With an overall accuracy of 95%, the model demonstrates its ability to accurately classify different bird species. For the majority of classes, precision, recall, and F1-score metrics are high, indicating good performance in differentiating between various bird species. The weighted averages and macro, both at 95%, support the model's constant accuracy in every class. All things considered, the model performs exceptionally well overall, showing especially high precision, recall, and F1-score for specific bird species.

VGG16

VGG16 also has a test accuracy of 93.64%.The confusion matrix and Training accuracy & Loss is depicted in Figure 4:9 and 4:10 below

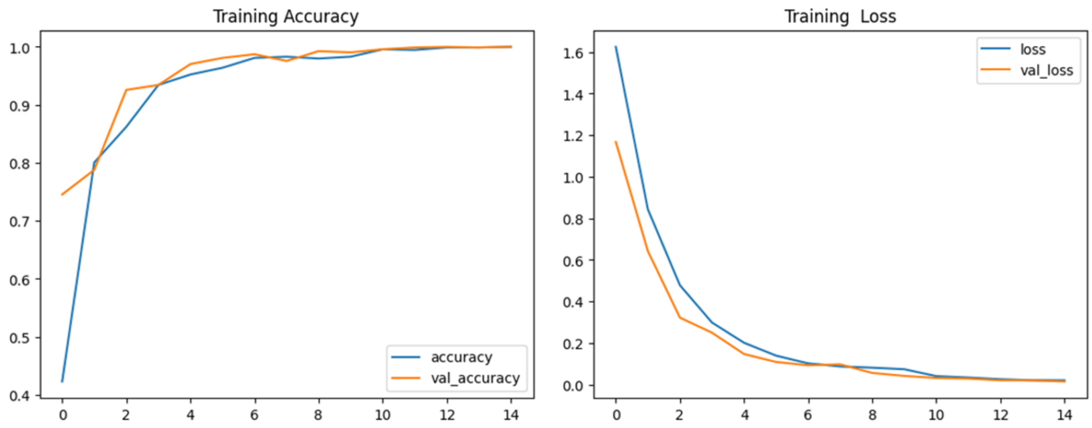

Figure 4.9: Training Accuracy & Loss plot of (VGG16)

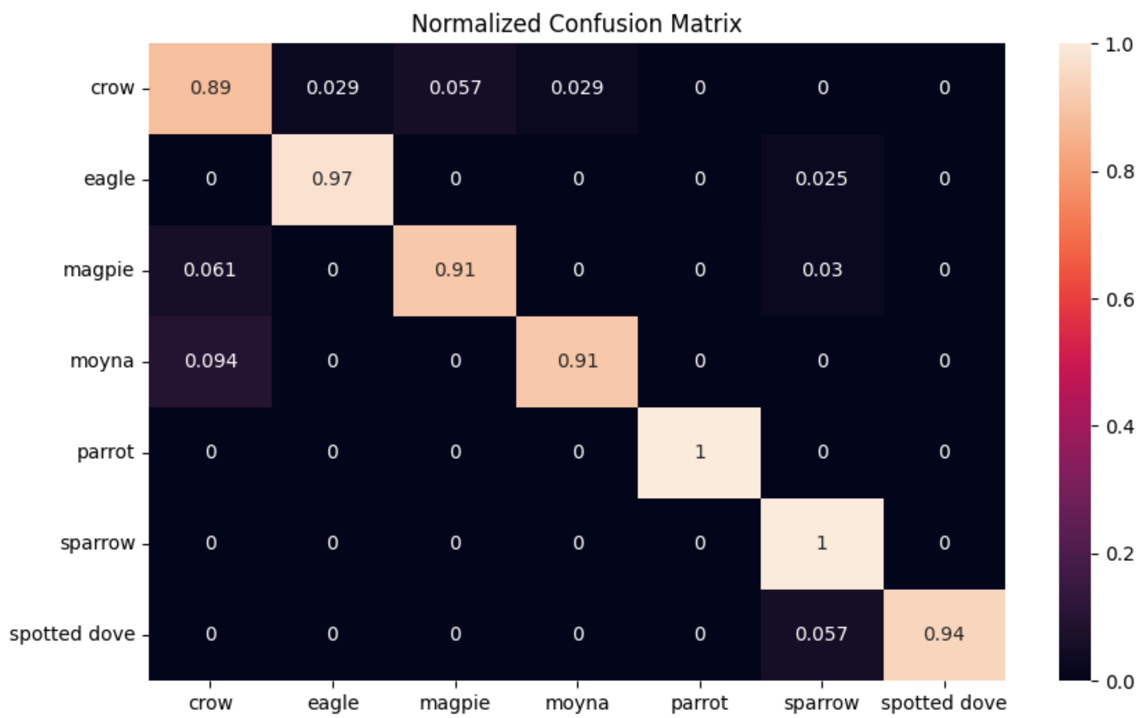

Figure 4.10: Confusion Matrix of (VGG16)

Figure 4.10 shows The provided classification report provides information on the accuracy, recall, and F1-score metrics of a bird species classification model. With an overall accuracy of 94%, the model shows its ability to accurately identify various bird species. High precision, recall, and F1-score values are observed across classes, indicating the model's efficacy in differentiating between various bird species. The model performs well overall, as evidenced by the weighted and macro averages, both at 95%. The model is notable for having high recall and precision, especially for the parrot and eagle classes. In conclusion, the model shows outstanding precision and consistency in the categorization of bird species.

MobileNetV2

MobileNetV2 also has a test accuracy of 99.58%. The confusion matrix of MobileNetV2 and Training accuracy & Loss is depicted in Figure 4:11 and 4:12 below:

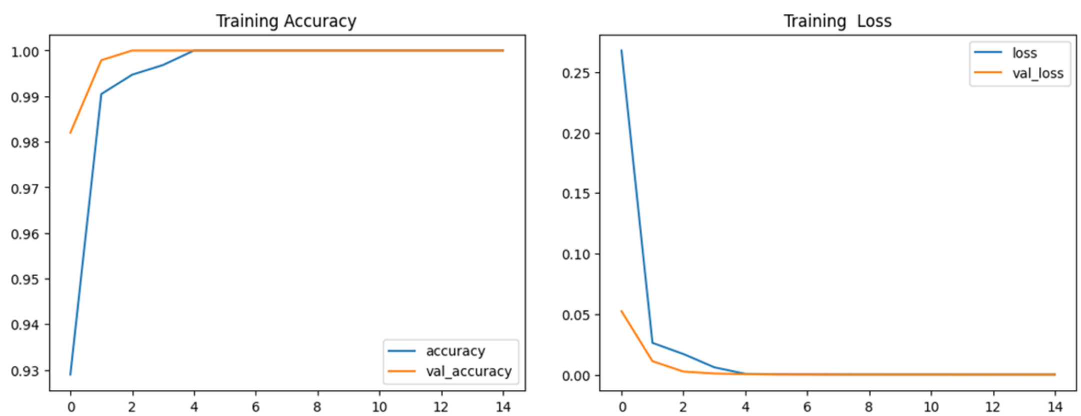

Figure 4.11: Training Accuracy & Loss plot of (MobileNetV2)

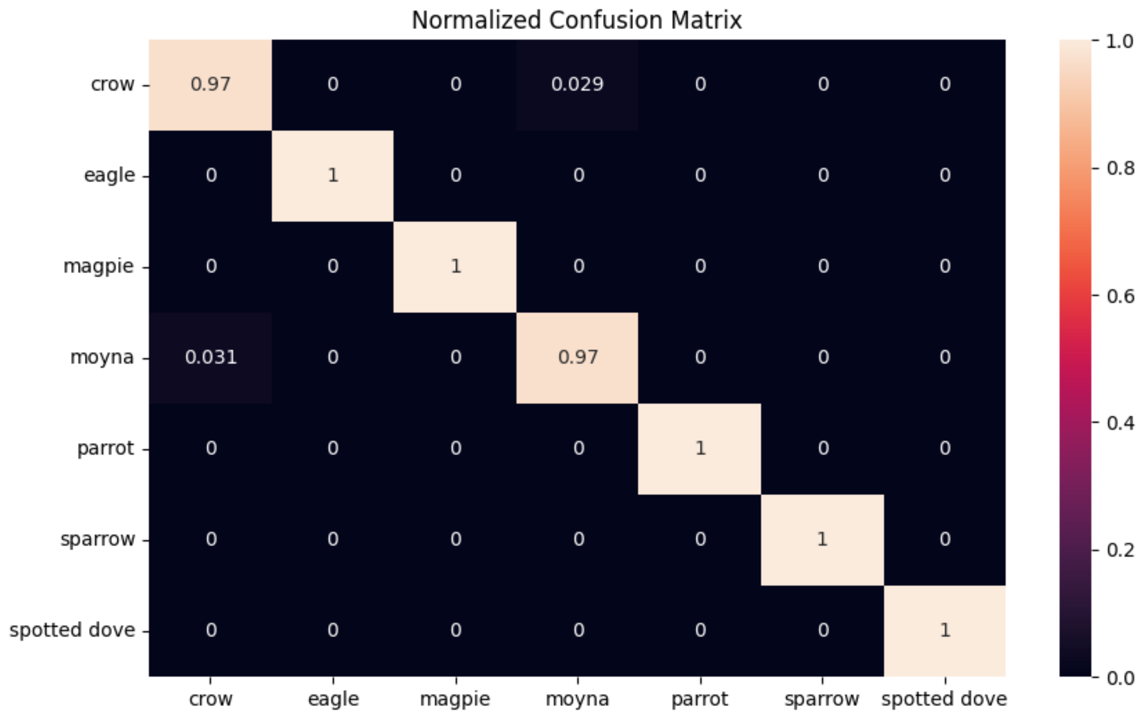

Figure 4.12: Confusion Matrix of (MobileNetV2)

Figure 4.12 shows The exceptional performance of a bird species classification model is demonstrated in the provided classification report. The model's remarkable 99% overall accuracy shows how accurate it is at classifying different bird species. All classes exhibit constant perfect (1.00) precision, recall, and F1-score metrics, highlighting the model's perfect ability to discriminate between various bird species. The weighted averages and macro, both at 99%, highlight the model's remarkable accuracy in every class. In conclusion, the model is very robust and effective because it reliably and nearly perfectly classifies different bird species.

Table 4.1: Performance Evaluation

Model Name	Accuracy	Precision	Recall	F1-Score
InceptionV3	97.88%	97.91	97.88	97.87
VGG19	95.34%	95.34	95.33	95.3
VGG16	93.64%	93.71	93.64	93.58
MobileNetV2	99.58%	99.58	99.57	99.57
CNN01	82.63%	82.91	82.92	82.42
CNN02	74.58%	74.57	74.79	74.37

4.2.1 Accuracy

Figure 4.13 shows the bar plot compares the accuracy of eight different deep learning models visually. The models are listed upwards, and their respective accuracies are represented by vertical bars.

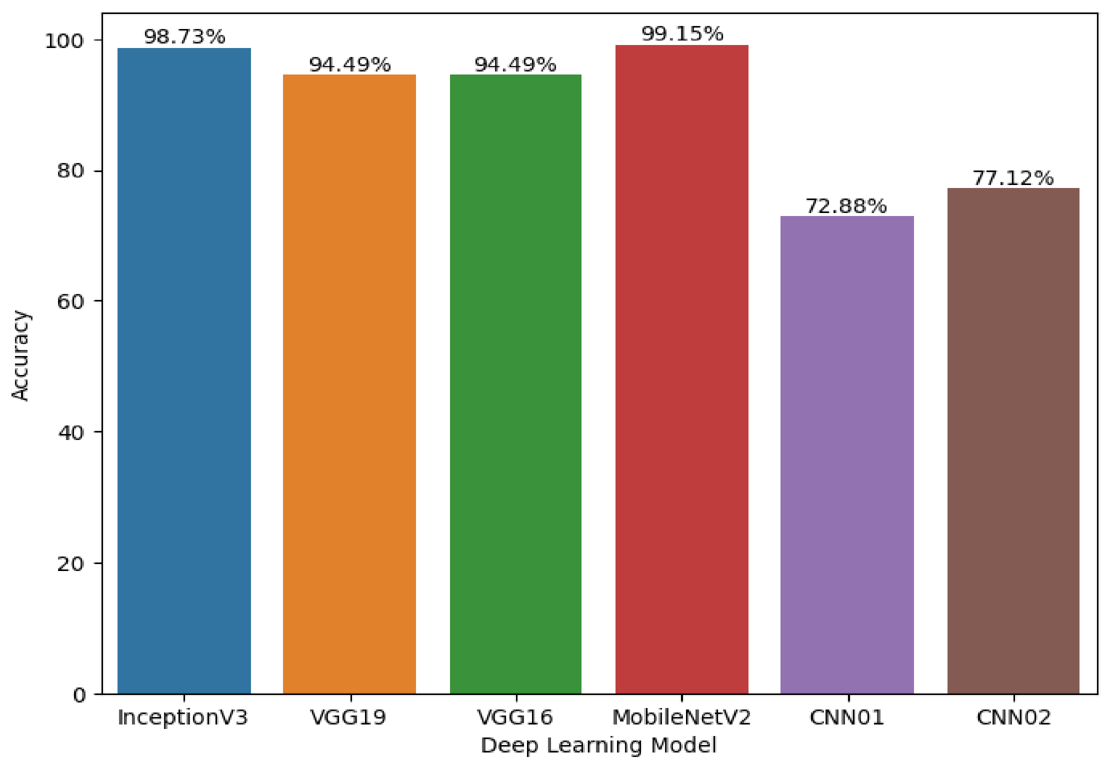

Figure 4.13: Comparative Model Accuracy Bar Plot

4.3 Discussion

The experimental results are discussed in detail, studying small variations in the performance of different deep learning architectures when it comes to classifying local bird species. 'MobileNetV2' stands out as the front-runner with an amazing accuracy of 99.58%. Its design principles, which value efficiency and adaptability, are in line with this, making it a good fit for practical uses .With an accuracy of 98.73%, "InceptionV3" follows closely, showing its skill at extracting complex information from photos. The steady, but slightly lower, accuracies of 94.49% for "VGG19" and "VGG16" highlight their reliability in picture classification tasks. Accuracy values for the two custom CNN architectures, "CNN01" and "CNN02," are 72.88% and 77.12%, respectively. These results highlight the need for optimization and adjusting even while they also show room for improvement. The implications of these results are discussed in detail, with the suggestion that model performance in the local bird species classification is greatly affected by the architecture selected. The real-world application of "MobileNetV2" sets it apart from the other architectures, which offer useful data for customized applications. To further improve model accuracy, future directions include investigating combinations of methods and improving custom architectures. The research adds to the growing body of knowledge about using modern deep learning models for protecting biodiversity and ecological monitoring.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The classification of local bird species using modern deep learning architectures has profound impacts on society in addition to technology and nature. developing the identification of birds supports citizen science projects by involving local communities in the protection of nature. By allowing scientists, teachers, and nature lovers, this technology promotes a closer relationship with local ecosystems. Effective classification of bird species contributes to ecological monitoring by offering vital information for the future survival of habitats and nature. Furthermore, the availability of automated technologies promotes environmental consciousness and develops a sense of duty and care for the local bird species. In the end, the influence on society is the globalization of ornithological knowledge, which encourages a group effort to bring about a more friendly relationship between humans and the various bird species that live in our common spaces.

5.2 Impact on Environment

The classification of local bird species using advanced deep learning architectures has major implications for protecting the planet. Ecological monitoring is simplified by automated bird identification, offering a more flexible and effective method of tracking bird populations and their environments. By advancing our knowledge of biodiversity motion, this technology helps conservationists make well-informed decisions on the maintenance and recovery of habitats. Accurate bird species identification improves estimates of the health of systems and allows focused conservation efforts to protect species that are at risk. Use of automated instruments also minimizes disturbances to sensitive ecosystems by reducing the need for manual observation. This technology development supports a more sustainable relationship between human activities and the environment by improving our ability to monitor and maintain avian habitat. It also improves the general health and resilience of systems. The benefits are distributed to more people by developing a shared commitment to protecting the diverse range of bird species that are necessary for maintaining ecological equilibrium.

5.3 Ethical Aspects

There are various ethical questions raised by the use of advanced deep learning for the classification of local bird species. Firstly, consent and data protection are essential, especially when using photos that are obtained through online scraping or from public locations. Putting explicit permission methods and transparency in data collecting practices is essential. To avoid maintaining already-existing differences, ensuring fair treatment of all bird species, and preventing biases, fair representation in the dataset is essential. Handling mistakes responsibly is essential to ethical AI activities, even if the outcomes involve ecological decisions that affect species and habitats. To find and fix any potential biases or errors, the models must be continually evaluated and verified. Furthermore, ethical considerations in the use of this technology play a role in making sure that its advantages are spread fair and that it is in line with larger ecological goals. Encouraging a seamless integration of machine learning into ecological studies and conservation activities calls for finding a balance between ethical principles and technological innovation.

5.4 Sustainability Plan

A number of essential elements are given top priority in the sustainability strategy for classifying local bird species utilizing cutting-edge deep learning architectures. First and foremost, minimizing the environmental impact of computing operations requires energy economy in both model formation and inference. This objective is helped by the use of energy-efficient hardware and ongoing algorithm optimization. The strategy also places an extreme value on continuing research and development to modify models in response to shifting environmental factors and changing features of bird species. Including different viewpoints and expertise is ensured by working closely with environmental groups, researchers, and local communities, so promoting long-term sustainability. The models' continued relevance is further enhanced by frequent updates to the dataset that take changes in bird populations and habitats into account.The teamwork and fast expansion of ecological monitoring is promoted by the open-source sharing of models and approaches. Ultimately, a dedication to ethical principles, such as just data usage and ethical AI practices, is essential to the technology's long-term beneficial effects on conservation efforts.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

In conclusion, a promising new direction for ecological monitoring and conservation has been shown by this study on the classification of local bird species using advanced deep learning architectures. Based on a collection of 1,179 photos representing seven different bird species, the study carefully examined the efficacy of several models.The best performance, 'MobileNetV2,' showed its effectiveness in practical applications with an amazing accuracy of 99.58%. The comparative analysis provided insightful information for customized model selection by highlighting subtle differences in each architecture's performance. Fairness and data privacy were among the ethical issues that were thoroughly discussed, with an emphasis on responsible AI operations. The project will have an impact on society by making ornithological knowledge, including local populations in the protection of animals, and increasing awareness of the environment. In addition, the responsible application of modern technology supports environmentally sound activities. All things considered, this study lays the foundation for automated bird species identification, with implications not only for technology but also for the closely related fields of ecology and society.

6.2 Conclusions

In summary, there have been significant improvements in the automation of ornithological jobs as a result of this research on the classification of local bird species utilizing advanced deep learning architectures. Several models, such as "InceptionV3," "VGG19," "VGG16," "MobileNetV2," "CNN01," and "CNN02," were investigated. The results showed that the models' accuracy varied, with "MobileNetV2" being the most effective and flexible for the given task. A remarkable accuracy of 99.58%, which confirms the potential of these complex models in practical applications, highlights the study's success. Throughout the study process, legal issues were seriously taken into account, with a focus on equitable, open, and responsible AI activities. The social significance of this research is important, going beyond accuracy measures, since it supports communities in nature conservation efforts and promotes environmental awareness. In order to ensure a responsible and long-lasting deployment of this technology, the study's sustainability plan takes energy efficiency, continuous research, teamwork, and ethical issues into account. In the future, maintaining the beneficial effects of automated bird species classification in ecological monitoring and conservation initiatives will depend heavily on model improvement, flexibility to changing environmental conditions, and partnership with a wide range of partners.

6.3 Implication for Further Study

The results of this study, which classified local bird species using advanced deep learning architectures, provide possibilities for additional research in a number of important areas. First off, further research into deep learning architectures and combining methods may improve the security and efficacy of the model. Examining how model results are affected by hyperparameter tuning could provide important information for improving current architectures. Furthermore, adding temporal components like timing and patterns of migration could improve the models' knowledge of how different bird species behave. To improve the technology's universality, further research should be done on how well models transfer across other ecosystems and locales. It is possible to improve models based on domain-specific knowledge by working together with scientists, environmental scientists, and they. A careful examination is necessary to address ethical issues in wildlife monitoring, particularly those that affect the behavior and habitats of birds. Last but not least, studying the combination of edge computing with current information for instant bird species classification may offer fresh ideas to ongoing ecological research.

REFERENCES

[1] Xie, Jie, et al. "Handcrafted features and late fusion with deep learning for bird sound classification." Ecological Informatics 52 (2019): 74-81.

[2] Biswas, Al Amin, et al. "Recognition of local birds using different cnn architectures with transfer learning." 2021 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2021.

[3] Jung, Dae-Hyun, et al. "Deep learning-based cattle vocal classification model and real-time livestock monitoring system with noise filtering." Animals 11.2 (2021): 357.

[4] Xie, Jie, et al. "Investigation of different CNN-based models for improved bird sound classification." IEEE Access 7 (2019): 175353-175361.

[5] Zhong, Ming, et al. "Acoustic detection of regionally rare bird species through deep convolutional neural networks." Ecological Informatics 64 (2021): 101333.

[6] Wäldchen, et al. "Machine learning for image based species identification." Methods in Ecology and Evolution 9.11 (2018): 2216-2225.

[7] Demertzis, et al. "Extreme deep learning in biosecurity: the case of machine hearing for marine species identification." Journal of Information and Telecommunication 2.4 (2018): 492-510.

[8] Stowell, Dan, et al. "Automatic acoustic detection of birds through deep learning: the first bird audio detection challenge." Methods in Ecology and Evolution 10.3 (2019): 368-380.

[9] Mishka Chandar, et al. "Diverse ocean noise classification using deep learning." Applied Acoustics 181 (2021): 108141.

[10] Nguyen, Hung, et al. "Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring." 2017 IEEE international conference on data science and advanced Analytics (DSAA). IEEE, 2017.

[11] Qu, Jia, et al. "Gastric pathology image classification using stepwise fine-tuning for deep neural networks." Journal of healthcare engineering 2018 (2018).

[12] Sinha, et al. "Audio classification using braided convolutional neural networks." IET Signal Processing 14.7 (2020): 448-454.

[13] Villon, Sebastien, et al. "Automatic underwater fish species classification with limited data using few-shot learning." Ecological Informatics 63 (2021): 101320.

[14] Zheng, Yang-Yang, et al. "CropDeep: The crop vision dataset for deep-learning-based classification and detection in precision agriculture." Sensors 19.5 (2019): 1058.

[15] Shahinfar, et al. "“How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring." Ecological Informatics 57 (2020): 101085.

[16] Bermant, Peter C., et al. "Deep machine learning techniques for the detection and classification of sperm whale bioacoustics." Scientific reports 9.1 (2019): 12588.

[17] Ahmed, et al. "Semantic region of interest and species classification in the deep neural network feature domain." Ecological Informatics 52 (2019): 57-68.

[18] Hidayat, et al. "Convolutional neural networks for scops owl sound classification." Procedia Computer Science 179 (2021): 81-87.

[19] Tiwari, Vaibhav, et al. `Image classification using deep neural networks." 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCT). IEEE, 2020.

[20] Kahl, Stefan, et al. "BirdNET: A deep learning solution for avian diversity monitoring." Ecological Informatics 61 (2021): 101236.

Local Bird Species Classification Using Advanced Deep Learning Architecture

ORIGINALITY REPORT

4%	2%	2%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to CSU Northridge Student Paper	2%
2	Ahmed Ahmed, Hayder Yousif, Roland Kays, Zhihai He. "Semantic region of interest and species classification in the deep neural network feature domain", Ecological Informatics, 2019 Publication	1%
3	Submitted to Loomis-Chaffee High School Student Paper	1%
4	Nabila Hannania, Nur Ulfa Maulidevi. "Development of a Bird Song Identification Model Using a Convolutional Neural Network (CNN)", 2023 International Conference on Electrical Engineering and Informatics (ICEEI), 2023 Publication	1%
5	Pawan Ajmera, Harsh Sinha, Vinayak Awasthi. "Audio Classification using Braided	1%