Bird Species Classification using Machine Learning

Project Progress Report

1. Introduction and Problem Statement

Birds play a critical role in global biodiversity, but many species are endangered. It is essential to accurately track and estimate bird populations, especially for local indigenous species. Unintentionally, wind farms have killed rare bird species in a number of nations. Due to differences in lighting and camera angles, it can be difficult to identify different bird species. Accurate models for species recognition are challenging to develop because of their variety in sizes, forms, colours, and physical traits. Through the use of supervised and semi-supervised learning techniques, this study attempts to identify bird species that are in danger of extinction by using decision trees and deep neural networks in machine learning [4, 6, 9].

The selection of the data set [1] is centred on the difficulty in identifying and classifying different bird species due to visual differences in size, form, colour, and other physical traits. The varying lighting conditions, various camera angles, and the complexity of the various bird species' visual traits are all difficulties in tackling this issue. Additionally, challenges are presented by the numerous bird species and the requirement for accurate identification. The selected method entails using deep learning models, such ResNet and CNNs, which are capable of learning detailed characteristics and patterns from the photos, to overcome these issues.

The use of pre-training on a dataset containing several bird species classes, followed by fine-tuning and training from scratch using stochastic gradient descent (SGD) optimisation, is one potential approach to overcoming the difficulties. A variety of criteria, including accuracy, precision, and confusion matrix analysis, can be used to assess the models. Plots of the model's performance that demonstrate diminishing loss and rising accuracy over time throughout training can shed light on the models' capacity for learning.

2. Proposed Methodologies

2.1. Preprocessing the Data

Image Datasets often come in a variety of sizes and aspect ratios. The photos must be resized while retaining the aspect ratio to a standard size. This procedure lessens computing complexity during training and aids in ensuring homogeneity across the dataset. We have used the python script to resize and remove the invlid images that have incorrect sRGB format, While the original images consists of 1200*800 resolution, but some of them also have different resolution, so we have resized all the images into 224*224 while mainting the original aspect ratio. We have divided the whole dataset into training(70%), validation(15%) and testing(15%) dataset. The original dataset consists of 25 bird species with 22.6k photos, while we have reduced the number of bird species to train our model, we have accounted 10 Bird species.

2.2. Architectures Selection and Training

We utilized ResNet [10], a highly influential deep learning model developed by Microsoft Research, in our bird classification project. ResNet effectively addresses the challenge of vanishing gradients in deep neural networks by incorporating residual connections that enable the direct flow of gradients from earlier layers to later layers. This architectural design allows for the training of deep networks without suffering from performance degradation.

Our ResNet model was trained from scratch using stochastic gradient descent (SGD) [5] optimization. SGD is a popular optimization technique in deep learning that approximates the true gradient by computing gradients on randomly selected mini-batches of training data. This approach ensures efficient updates of the model's parameters, enabling effective learning.

By combining the hierarchical feature learning capability of ResNet with SGD optimization [3], we aimed to develop a robust bird classification model. The ResNet architecture with residual connections allowed our model to effectively learn intricate features, while SGD optimization facilitated the iterative refinement of the model's parameters. The utilization of ResNet and SGD in our project contributed to accurate bird classification results.

3. Attempts at solving the problem

In this study, we employed a deep learning approach using Convolutional Neural Networks (CNN) for bird species classification. Initially, we utilized a pre-trained model that was trained on a dataset comprising 25 different bird species classes, totaling 22,600 images. The training of this pre-trained model took approximately 3 hours, resulting in a commendable accuracy of 92.6%.

However, in order to explore the impact of training from scratch, we also employed an untrained model. Surprisingly, the untrained model exhibited a significantly lower accuracy of only 11% during training. This outcome highlights the importance of pre-training and fine-tuning in deep learning.

To address the issue of lengthy training times and to focus on a more manageable subset of bird species, we decided to reduce our dataset to 10 bird species classes. This reduction in the dataset size enabled us to mitigate the training time while still maintaining a reasonable level of accuracy.

During the training process, our reduced dataset achieved an accuracy of 89.56%. For testing, the model yielded an accuracy of 82.19% with 10 epochs, indicating its capability to generalize well to unseen data. This outcome demonstrates the effectiveness of the model in accurately classifying bird species, even with a reduced dataset. Here are the evaluation matrices for the reduced dataset size model performance.

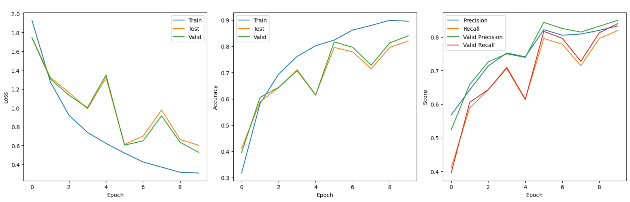

Figure 1. ResNet Model Performance over each epoch with respect to Loss, Recall, Precision and F-Score

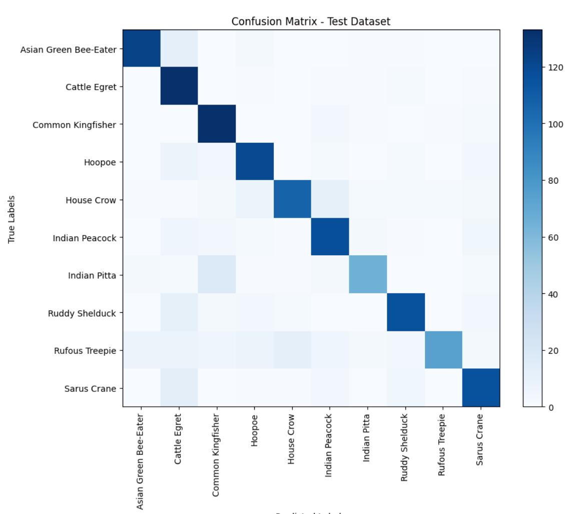

Figure 2. Confusion Matrix to visualize the difference of predicted and true label for ResNet

The visualizations in Figure 1 and Figure 2 provide a comprehensive assessment of our deep learning model's performance. The decreasing loss, increasing accuracy, and precision observed in Figure 1 indicate the model's learning capabilities. Furthermore, the confusion matrix [8] in Figure 2 offers valuable insights into the model's classification behavior, enabling us to identify any confusions between classes and potential areas for enhancement.

Overall, our findings highlight the importance of pre-training and the influence of dataset size on the performance of deep learning models. By leveraging pre-trained models and optimizing the dataset size, we were able to achieve notable accuracy levels in bird species classification using CNNs.

4. Future Improvements

For future work, we plan to leverage the power of t-SNE (t-Distributed Stochastic Neighbor Embedding) [7], t-SNE is a dimensionality reduction technique that has proven to be valuable in bird species classification using deep learning CNN models. By applying t-SNE to the extracted features of bird images, we can visualize the data in a lower-dimensional space (typically 2D or 3D) where the relationships between different bird species become more apparent. This allows us to gain insights into the similarity or dissimilarity of bird species based on their visual features. By examining the t-SNE plot, we can observe how bird species with similar visual characteristics tend to form clusters or groupings.

In our pursuit of improving the performance of our training model, we have outlined two key strategies to be implemented and evaluated. Firstly, we intend to increase the number of epochs during the training process. By extending the training duration, we aim to provide our model with more opportunities to learn and refine its representations of the bird species dataset. Secondly, we plan to explore the utilization of Decision Tree algorithms for both supervised and semi-supervised learning models [2]. Decision Trees are powerful and interpretable machine learning algorithms that partition the data based on feature values, enabling effective classification. By employing Decision Trees, we aim to leverage their capabilities in capturing complex patterns and relationships within the bird species dataset.

To evaluate and compare the performance of these models, we will employ two key evaluation techniques: the confusion matrix and t-SNE scatter plot. These evaluations will guide us in determining the most effective approach for achieving improved accuracy and robust classification results.

References

[1] Arjun Basandrai. 25 indian bird species with 22.6k images, 2023. Dataset of 25 bird species for image classification. 1

[2] Baowen Qiao; Zuofeng Zhou; Hongtao Yang; Jianzhong Cao. Bird species recognition based on svm classifier and decision tree. *2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS)*, 2018. 2

[3] Anshu Raina Chirayu Garg, Abhinav Garg. Analysis of gradient descent optimization algorithms on resnet, 2017. Department of Computer Sciences, University of Wisconsin-Madison. 1

[4] Piyush Bidwai; Vaibhav Mahalle; Eshan Gandhi; Sharda Dhavale. Bird classification using deep learning. *International Research Journal of Engineering and Technology (IRJET)*, 7, 2020. 1

[5] Ayush Gupta. A comprehensive guide on optimizers in deep learning, 2021. Blog on Deep Learning Optimizers - Stochastic Gradient Descent (SGD), Adam. 1

[6] Yo-Ping Huang and Basanta Haobijam. Recognition of endemic bird species using deep learning models. *IEEE Access*, PP:1–1, 07 2021. 1

[7] Geoffrey Hinton Laurens van der Maaten. Visualizing data using t-sne. *Journal of Machine Learning Research 9 (2008) 2579-2605*, 2008. 2

[8] Sarang Narkhaede. Understanding confusion matrix. 2018. 2

[9] Benedict Neo. Bird species classification with machine learning. 2022. 1

[10] Yucheng Li; Hua Zhou; Yili Zhao. The fine-grained recognition of bird images based on joint semantic components and resnet. *2022 5th Internation Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2022. 1