Mini Project Report on

DISCRIMINATIVE LEARNING USING VISION MODELS

Submitted in partial fulfillment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by:

Student Name	University Roll No.
PRAGYA BHATT	2021353

Under the Mentorship of

Prof. Dr. Arun Chauhan

Department of Computer Science and Engineering

Graphic Era (Deemed to be University)

Dehradun, Uttarakhand

January-2025

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project report entitled **"Discriminative Learning using Vision Models"** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Arun Chauhan , Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name : Pragya Bhatt University Roll no : 2021353

Table of Contents

Teacher’s Sign:

Chapter 1

Introduction

1.1 Introduction

Bird species recognition and distinction are key components of preserving biodiversity, ecological study, and wildlife monitoring. Birds have a wide spectrum of visual variation, with distinctive features such as eyes, beaks, heads, claws, and feathers playing an important part in discerning one species from another. Knowing and evaluating these features might provide useful information on evolutionary traits, ecological preferences, and species-specific behaviours.

Object recognition and feature extraction have become crucial in a variety of fields, from driverless vehicles to wildlife protection, in the age of developing artificial intelligence and machine learning. In order to better comprehend the unique traits of different bird species, this research makes use of the YOLOv8, a cutting-edge object detection algorithm, to assess the features of different birds species and also distinguish characteristics between them. The urgent need to distinguish minute yet important variations in the physical characteristics of birds, such as their wings, beaks, claws, and feathers, is what inspired this effort. Ecological studies, conservation initiatives, and species identification can all benefit from accurate differentiation. This work aims to close the gap between automated detection and deep learning by putting in place a deep learning pipeline that incorporates image preprocessing, bounding box creation, and in-depth feature analysis.

In this study, we introduce a YOLOv8-based educational vision-based method for Differential Learning. The main objective is to create a system that can distinguish between different bird species by examining their unique characteristics, particularly their head, eyes, beak, claws, and feathers. It might be difficult for kids, students, and birdwatching aficionados to distinguish minute distinctions between birds. This project efficiently highlights these distinctions through an intuitive, visual learning experience that makes use of object detection and feature comparison.

We created a custom dataset with photos of several bird species that were manually classified to include the areas that matched the five main characteristics. YOLOv8, a cutting-edge object identification model renowned for its accuracy and speed, was used to train the system. You Only Look Once (YOLO) models are quite effective for real - time detection tasks. Here, YOLOv8 makes it possible to locate the specific bird characteristics like eyes, beak, claws, head and feathers inside the pictures.

Following feature detection, the differences between two birds' relevant features are visualized using a comparison method based on heatmaps. The structural similarity index (SSIM) is used to quantify this comparison and highlight areas of notable variance. Through the integration of comparative visualization and detection, the project provides an entertaining and instructive tool for learners to recognize and differentiate important bird traits. This research shows how to use image processing, object detection, and deep learning to improve observational abilities and promote a better comprehension of bird anatomy. It makes studying nature both interesting and enjoyable by fusing interactive education with computer vision technologies.

1.2 Problem Statement

Differentiating between bird species can be challenging due to subtle differences in features like beaks, claws, and feathers. Manual identification is time-consuming and error-prone. This project aims to use YOLO and other techniques to automatically detect and compare bird features using similarity scores and heatmaps. This approach provides a scientific method for studying evolutionary differences and improving bird classification accuracy.

Fig. 1.1 Images of some birds species used in the dataset

Chapter 2

Literature Survey

The research by Nidhi et al. [8] a CNN-based bird recognition system that automates species identification, enhancing accuracy and efficiency compared to traditional methods. It evaluates performance through metrics like accuracy, precision, recall, and F1 score, contributing to biodiversity conservation and ecological research.

Table 2.1 Summary of the key literature.

Methodology	**What has been done, how it has been done**	**Outcomes**	**Scope for further work**
Saxena et al. [1]	The research focused on identifying bird species using object detection algorithms, specifically Faster R-CNN and YOLOv5. The paper describes data collection, annotation, preprocessing, and the training of the models, which aligns with the methodology outlined.	The results of the YOLOv5 model achieving a high mAP score compared to Detectron2 are mentioned in the paper, highlighting the effectiveness of the YOLOv5 model for bird identification. The YOLOv5 model was trained for 100 epochs, achieving a mAP score of 0.7 while Detectron2 achieved Intersection Over Union score of 0.001.	Future research could focus on expanding the dataset to include a broader range of bird species from Maharashtra and other regions, enhancing the model's applicability.
Ian et al. [2]	Implemented training and validation model system for classification. Used ResNet152V2,	DenseNet169 achieved 95% accuracy on dataset 2. 85% training and 86% validation accuracy for dataset 1.	Implement additional datasets for improved classification accuracy. Explore other deep learning models for better performance.

Pruthvi et al.[3]	InceptionV3, DenseNet121, and DenseNet169 models. Utilized deep learning models: ResNet50, MobileNetV3, EfficientNet-B0, Wide-ResNet50V2. Assessed model accuracy, training, validation loss, and performance metrics.	EfficientNet-B0 achieved 99.09% accuracy in identification.	Advancing computer vision in biodiversity monitoring. Enhancing the species identification accuracy for conservation efforts.
Kuihe Yang et al. and Ziying Song et al. [4]	The study presents a novel approach to enhance the object detection for fine-grained birds using a modified the YOLOv3 framework, termed YOLOBIRDS.	The algorithm used achieved: Improved the detection accuracy for fine-grained bird species.Reduced model size and faster processing times. Enhanced robustness against environmental variations	Future research directions include: Exploring the additional activation functions for improved accuracy. Testing with more diverse bird image datasets
Suhas Reddy et al. and Veluri Raviram Nikhil et al. and P Bhaskar Reddy et al. [5]	Utilized CNN and SVM for bird classification. Employed hyperparameter tuning and cross-validation for model optimization.	Achieved 98% accuracy in bird species classification. Promoted conservation efforts for endangered bird populations.	A CNN-based bird identification project can be further expanded by incorporating audio recordings and interactive maps, which turn it into a multi-modal tool .
Rashmi et al.[6]	Images converted to grayscale and processed through TensorFlow. Deep convolutional neural network predicts bird species using score sheet.	Accuracy range of 80% to 90% in identification. Utilized the deep convolutional neural network for classification.	Testing with more diverse bird image datasets

Chapter 3

Methodology

From model training and data preparation to feature selection and similarity analysis, there are several steps in the process for detecting and comparing features in bird photos. By using YOLOv8 for object detection, this method makes it possible to accurately identify bird traits including the head, eyes, beak, claws, and feathers. The steps listed below describe the implementation process in detail:

3.1 Dataset Collection:

The project's dataset has been custom-created and organized with a focus on identifying and distinguishing characteristics of birds from different species. As the format indicates, the structure is divided into three primary folders: train, val, and test. Each of these folders is further subdivided into images and labels, ensuring clarity and efficiency for the processing pipeline.

The train folder contains 97 bird images along with their corresponding label files, while the val and test folders each consist of 90 images and their respective labels. Bounding box annotations are included in the labels to draw attention to particular bird characteristics that are crucial for species differentiation and are clearly evident in the photos, such as the eyes, beak, feathers, head, and claws.

3.2 Object Detection (Using Yolo) :

i. Labeling Images Using LabelImg: A common approach begins with using LabelImg, a tool for manually drawing bounding boxes for object detection. Load images into LabelImg and manually draw bounding boxes around the desired features. Assign class labels to each bounding box ("eye", "beak", "claws", "head", "feathers").Save the annotations in YOLO (TXT) format directly from the tool.

ii. Parsing Annotations Using a Custom Script**:** If annotations are saved in XML (Pascal VOC) format, you'll need a custom script to convert them into the YOLO TXT format or a structured JSON file for additional processing

iii. Save the processed annotations in both TXT and JSON formats as needed.

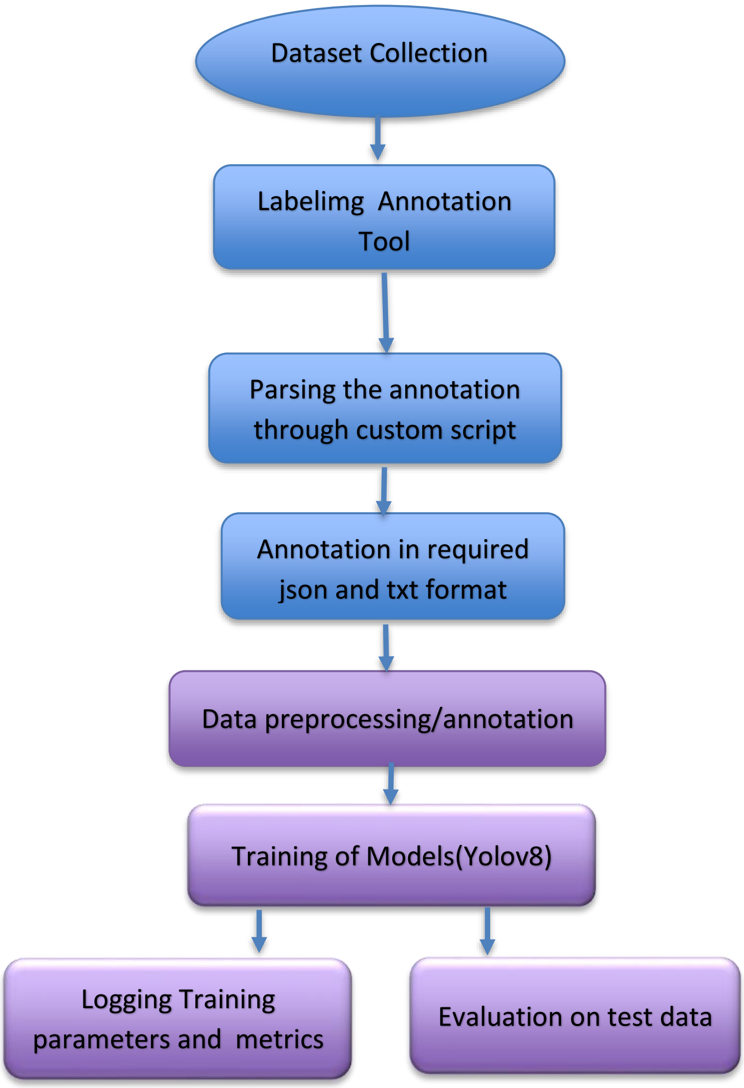
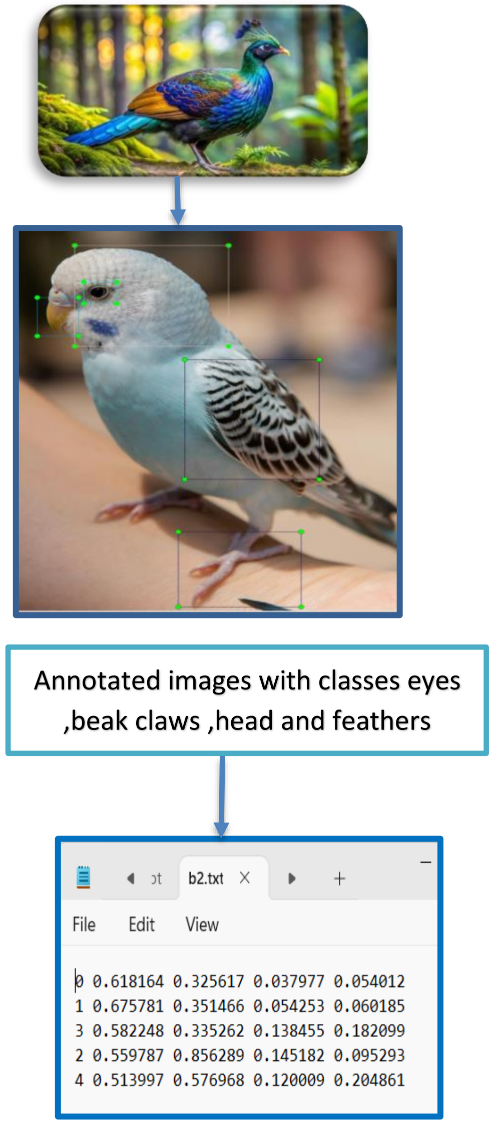


Fig. 3.1 Flowchart for Object detection using Yolo



<u>**Label of a image**</u>

3.3 Feature-Centric Methodology for Bird Image Differentiation

This methodology outlines the steps involved to differentiate features in bird images. The process includes feature cropping, preprocessing, and comparing features using advanced techniques like heatmaps and Structural Similarity Index (SSIM).

i. **Cropping of Features :** Images are first examined to highlight the features unique to each bird, including the head, wings, eyes, beak, and claws. This trimming makes sure that the model ignores the background noise and highlights the important features.

ii. **Image Preprocessing:** All cropped images are resized to a uniform dimension (e.g., 224x224 pixels) to standardize input sizes. Pixel values are normalized to a standard range ([0, 1] or [-1, 1]) for consistent and efficient processing.

iii. **Data Splitting:** The dataset is split into three distinct sets to facilitate training, validation, and testing:

- **Training Set**: Used to train and optimize model parameters.
- **Validation Set**: Used to monitor and fine-tune the model during training, preventing overfitting.
- **Test Set**: Used for final evaluation to measure the model's ability to generalize to unseen data.

iv. **Similarity Analysis Using SSIM and Heatmap Visualization:** The Structural Similarity Index (SSIM) is applied to compare specific features between pairs of bird images. SIM quantifies the similarity in terms of luminance, contrast, and structural alignment, providing a score between -1 and 1. Heatmaps are generated to highlight regions of interest that contribute most to feature differences. These visualizations aid in interpreting similarity scores by showcasing the exact regions where variations occur.

v. **Similarity Score Calculation:** Based on the SSIM results, a similarity score is assigned to each image pair. These scores are used to identify distinct features, highlighting differences between bird species.

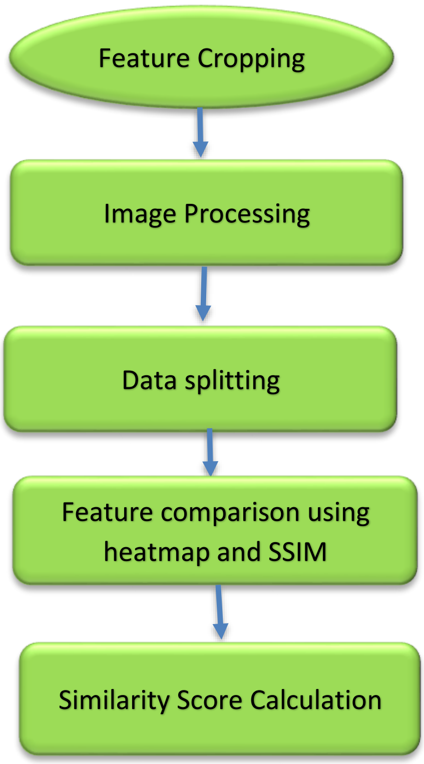


Fig. 3.2 Flowchart for feature differentiation

Chapter 4

Result and Discussion

4.1 Results

First, we used a Structural Similarity Index (SSIM)-based method to examine how similar the bird photos were to one another. The photos in the dataset were feature-isolated (eyes, beak, etc.), and each feature was examined separately to measure variations. The following was discovered by the analysis:

4.1.1. Full Image Comparison: The SSIM method, the structural similarity between the two full images was calculated. The low score suggests distinct visual characteristics, such as differing patterns, textures, or features across the entire image while high SSIM score suggests similarity is quite good .

Table 4.1 SSIM score used for showing features comparison

SSIM Score Range	Outcome	Description
> 0.8	Very Similar	The full images or features are almost identical, with very few differences.
0.5 < score ≤ 0.8	Some Differences	The full images or features have some noticeable differences but share significant similarities.
0.3 < score ≤ 0.5	Quite Different	The full images or features exhibit considerable differences
≤ 0.3	Very Different	The full images or features are highly distinct with minimal structural similarity.

4.1.2. Feature-Level Comparison: The difference heatmap shows significant variations in the "eyes" region, highlighting distinct eye shapes and colours between birds. The below eyes of two birds are different.

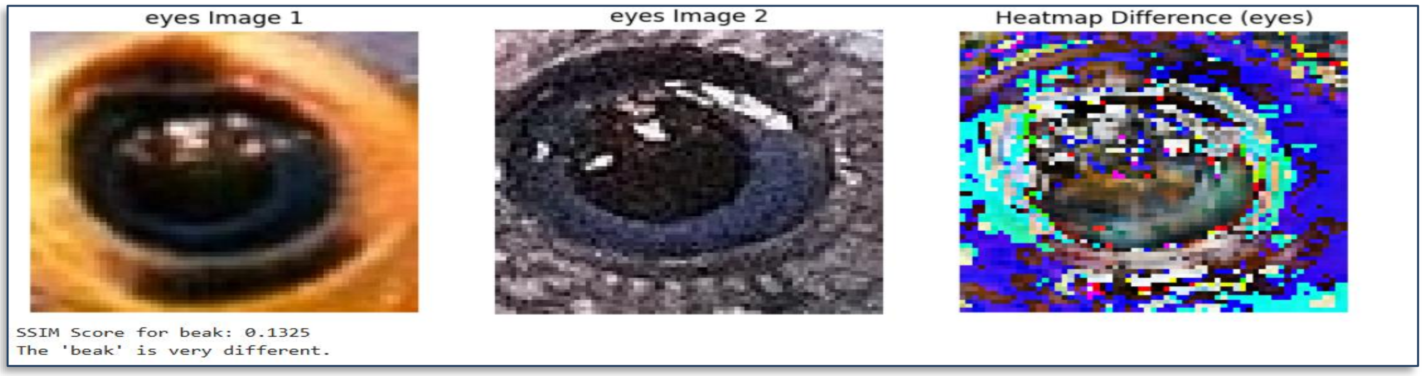


Fig. 4.1 Similarity score for eyes and heatmap difference

- Similarly ,other features such as "beak," "claw," "feathers," and "head" can be similarly compared using the code. The below output primarily focuses on the "beak" and "claws" feature.

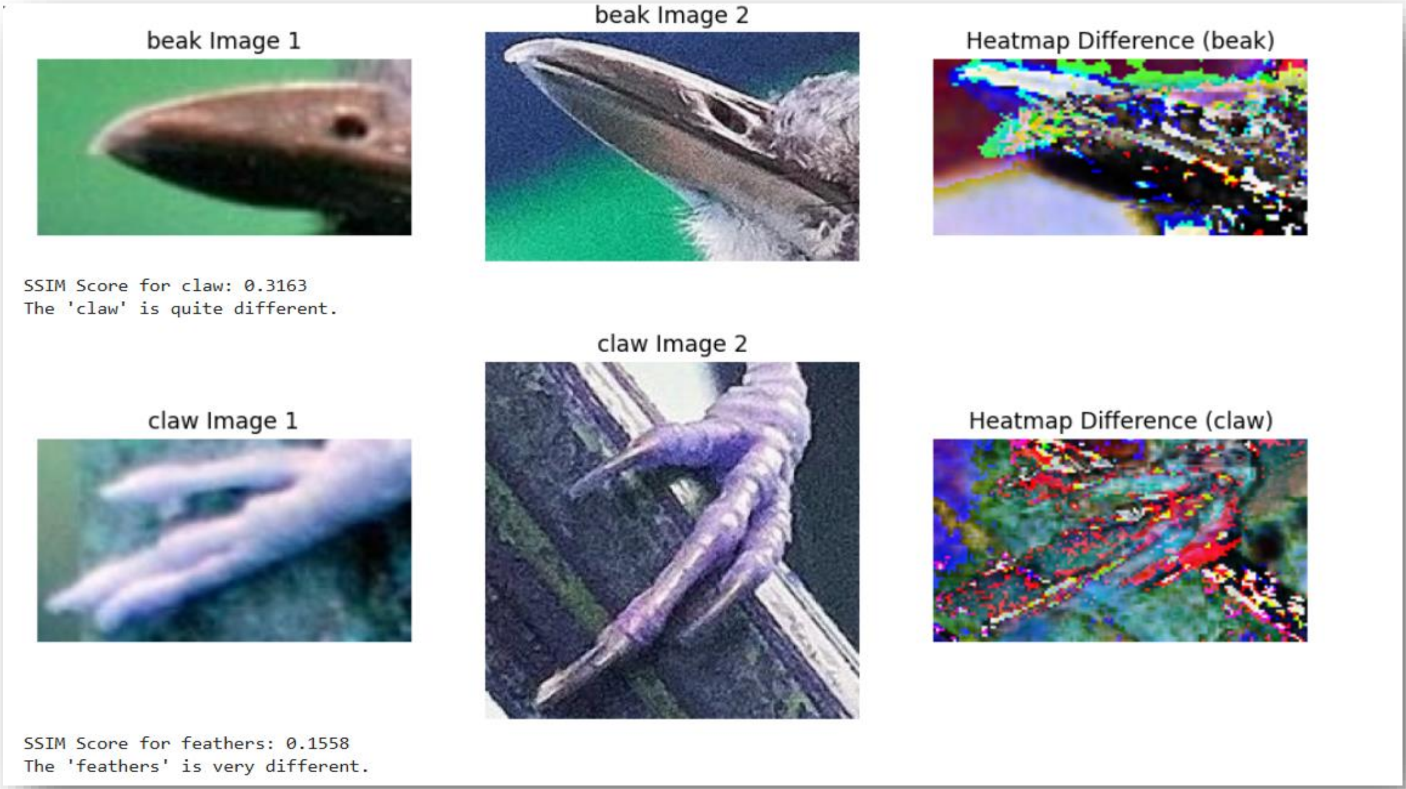


Fig. 4.2 Structural and Visual Differences in Beak and Claw Features

4.2 Discussion

Effectiveness of SSIM: When it came to measuring picture differences, SSIM worked well. Significant differences are highlighted by low scores (<0.3), which makes it appropriate for feature-based analysis of bird photos.

Feature-Specific Insights: A deeper comprehension of the primary traits that set different bird species apart is made possible by the method's ability to make focused comparisons by isolating particular features (like eyes, beak, claws, feathers and head).

Heatmap Visualization: Using colours to encode feature intensity or differences, a heatmap visualization is a graphical representation that highlights areas of interest in an image. It is frequently used for applications including feature analysis, object location, and picture similarity detection. Brighter or warmer colours (like red or yellow) indicate areas of high intensity or major differences, while cooler or darker colours (like blue or black) indicate areas of low intensity or similarity. Heatmaps operate by superimposing a colour map. For instance, when comparing two photos, a heatmap created with tools like SSIM (Structural Similarity Index) can show the changes between each pixel.

Limitations: While the SSIM approach yielded clear results, it has limitations in capturing complex perceptual or semantic differences. For example, variations in lighting, image quality, or orientation might influence the scores, making it less reliable for highly nuanced comparisons.

Chapter 5

Conclusion and Future Work

The project successfully demonstrates a method for differentiating bird species based on key morphological features using YOLO and SSIM. The YOLO-based detection system effectively localized features like the eyes, beak, claws, feathers, and head, enabling their cropping and comparison. The SSIM metric provided a quantitative measure of the structural similarity between these features across different bird images, offering an in-depth analysis of their differences. This approach highlights the potential for automated feature extraction and comparison, which can be used in evolutionary biology research and species identification tasks. The ability to quantify and visualize morphological differences offers a new perspective on avian phenotypic variation and supports more accurate studies on selective evolutionary processes

This research provides a strong foundation for further advancements in combining deep learning techniques with evolutionary biology studies. The following areas present promising opportunities for future work:

Expansion to Larger, Diverse Datasets: Future models can be trained on larger, more diverse datasets, including a wider range of bird species, habitats, and morphological adaptations. This would increase the system's generalization capabilities and accuracy over a broader variety of avian taxa. Access to larger datasets may also aid in macroevolutionary investigations of morphological variety

Cross-Taxa Generalization: The methodologies created in this project for birds can be applied to other animal taxa, including reptiles, mammals, and insects, to explore comparative morphology and evolutionary adaptations across the tree of life.

Incorporation of 3D Imaging and Point Cloud Processing: In the future, 3D imaging and point cloud processing can be utilized to examine the intricate shapes and structures of birds. Technologies such as LiDAR and depth sensors collect 3D data, offering a more full picture of avian morphology. Advanced models, such as PointNet++ or VoxelNet, can use this 3D data to learn complex features, resulting in more precise shape analysis and a better understanding of evolutionary adaptations. This provides more depth and precision than typical 2D picture processing.

References

[1] Saxena, S., Thorat, S., Jain, P., Mohanty, R., & Baraskar, T. (2022). Analysis of object detection techniques for bird species identification. *Journal of Physics: Conference Series*, 2325, 012054. https://doi.org/10.1088/1742-6596/2325/1/012054

[2] Ian, Ting, Chun, Ing., Nurul, Sabrina, Razali., Mohd, Azhar, Abdul, Razak., Eileen, Su, Lee, Ming., Nor, Aini, Zakaria., Nasrul, Humaimi, Mahmood. (2023). 5. Deep Learning Analysis Using CNN For Bird Species Classification System. doi: 10.1109/icicyta60173.2023.10428729

[3] Pruthvi, Darshan, S, S., Mathew, Jayanth., V., Sangeetha. (2024). 8. Multiclass Bird Species Identification using Deep Learning Techniques. doi: 10.1109/conecct62155.2024.10677184

[4] K. Yang and Z. Song, "Deep Learning-Based Object Detection Improvement for Fine-Grained Birds," in *IEEE Access*, vol. 9, pp. 67901-67915, 2021, doi: 10.1109/ACCESS.2021.3076429.

[5] Pichili, Vijaya, Bhaskar, Reddy. (2023). 9. Endangered Bird Species Classification Using Machine Learning Techniques. International Journal For Science Technology And Engineering, doi: 10.22214/ijraset.2023.5117

[6] C., Rashmi., MD., SHAGUPTHA., Munipalle, Sai, Nikhila., L.PURNIKA. (2024). 3. Birds species identification using deep learning. doi: 10.36893/ijesat.2024.v24i10.016

[7] (2022). 10. Bird Species Identification Using Convolutional Neural Network. Advances in parallel computing, doi: 10.3233/apc220053

[8] Nidhi, Gaur., Jenica, Kodankandath., Tanvi, Mhasal., Jasmitkaur, Mukhtiar., Rexlin, Nadar. (2024). 1. Watch Me: A CNN Based Platform for Bird Species Identification and Educational Resources. doi: 10.1109/i-smac61858.2024.10714737