Grenze International Journal of Engineering and Technology, June Issue

Image Recognition for Bird Species Identification using Fine Graining

Amruta Tapas Paul[1], Ajinkya Nanivadekar[2], Taha Hussain[3], Swaraj Kadu[4], Aayushi Sonkusare[5] and Utkarsha Pacharaney[6]
[1-5]Yeshwantrao Chavan College of Engineering, Department of Computer Science and Engineering, Nagpur, India
Email: amrutatapas@gmail.com, nanivadekarajinkya@gmail.com, tahahussain.main@gmail.com, swarajunofficial0@gmail.com, sonkusareaayushi26@gmail.com
[6]Faculty of Engineering and Technology, Datta Meghe Institute of Higher Education & Research Wardha
Email: sahildhanvij7@gmail.com

***Abstract*—There exist many millions of different species of birds and animals together on this planet. They live in different habitats, exhibit morphological differences, show variance in their physical features which are indistinguishable by the human eye. Hence, due to a plethora of species the identification of different bird and animal species can be done using fine grained image recognition techniques. Fine grained image recognition is widely used to accurately identify visually similar species showing morphological differences using feature extraction and deep learning techniques with classes showing significant inter class variations and minimal intra class variations. Therefore, a comprehensive study encompassing the ecological statistics of biodiversity as well as endangered species can be done.**

***Keywords*— Fine Graining, Feature Extraction, Specie Identification, Image Recognition, Object Detection.**

I. Introduction

Our earth is home to around 10,000 bird species and 8.7 million animal species. Out of 8.7 million species about 1.2 million species are described formally only and the rest are still unidentified specifically in region like rainforests and deep oceans. Among vertebrates there are around 5,500 mammal species, 32,000 fish species reptilian and amphibian species. These numbers describe our planet earth's incredible biodiversity and increasing complexity to overcome the challenges faced during the identification of these species. Every specie has its own habitat, unique metabolic and catabolic bodily requirements, suitable climatic conditions, morphological differences, and a bunch of physical features. Migratory Bird Species cover tens of thousands of kilometers yearly. Certain species, such as Kiwi in New-Zealand or the hoatzin in South Africa, are restricted to specific regions due to unique ecological conditions. These species are called as Endemic species. Birds show a wide range of habitat diversity. They thrive in various habitats, including toucans in tropical rainforests, Flamingoes in wetlands, sandgrouse in deserts and pigeons in urban environments. They incredibly differ in size from the tiny Bee Hummingbird to massive Ostrich which can grow 2.7 meter all or more.
The conservation of these species is a challenging task and also a threat due to habitat destruction, climate change, poaching etc. Conservation measures like biodiversity monitoring and community-based efforts aim to conserve and also preserve these species.

Grenze ID: 01.GIJET.11.2.280_20

To summarize this our earth's fauna reflects an extraordinary difference of life shaped by millions of years of evolution. The human eye is prone to inaccuracy and errors while distinguishing similar species showing subtle morphological differences, inter class and intra class variations. Fine-grained image recognition (FGIR) plays a revolutionary pivotal role in accurately identifying bird and animal species, addressing the challenges posed by the subtle visual differences. Unlike general image recognition, FGIR focuses on recognizing intricate details such as feather patterns, fur textures, body morphology, and unique coloration. In the identification of bird species, for instance, FGIR models like YOLO, Res Net, and Vision Transformers are trained to detect subtle differences in plumage patterns, beak shapes, and wing markings. Applications of FGIR extend beyond classification to practical use cases such as endangered species monitoring, ecological studies, and biodiversity hotspot mapping. So, in a nutshell, Fine Grained Image Recognition can be done to identify different species of birds and animals thereby monitoring count of endangered species, accurate identification of visually similar species and promotion of ecological studies.

II. LITERATURE REVIEW

The paper [1] discusses novel approaches targeted at the improvement of fine-grained image classification without relying on huge labeled datasets. It uses species of animals and birds whose numbers are fewer with smaller dataset size. The proposed dataset uses seen and unseen classes. It learns the association of visual features and semantic attributes. It makes use of a new model called ZIC-LDM, or Zero-shot Image Classification with Learnable Deep Metric. Zero-shot learning is a fundamental concept which allows models to recognize and classify unseen categories based on knowledge learned from previously known categories and performs knowledge transfer. It thus introduces complex metrics so that zero-shot models can accurately keep up with accuracy levels while classifying unseen categories or fine-grained distinctions. Paper [2] addresses the issue of speeding up the processing time and improving the accuracy in identifying species of animals using sophisticated models of R-CNN. It tries to overcome some of the serious problems that are being faced in real life like wildlife–human and wildlife–vehicle conflicts, which demand efficient and accurate detection mechanisms. For this purpose, research is carried out by employing four types of R-CNN models, which are incorporated with D-CNNs. This approach is intended to carry out a through comparison and evaluation of how effective they are in animal species detection. The paper [3] is about decoupled knowledge distillation. Here, the process of distillation is separated into parts to make it better for the transfer of knowledge from a large teacher model to a smaller learner or student model. The model is trained on the CUB-200-2011 dataset. They have used DenseNet121 and ShuffleNetV2 for feature extraction during training and compressed student model respectively. Decoupled Knowledge Distillation (DKD) separates the transfer of target and non-target class knowledge. The paper [4] mainly focuses on advancing the capabilities of fine-grained image recognition specifically for bird species. They have trained their model using the CUB200-2011 dataset. The Res2Net-CBAM modules enhanced multi-scale feature extraction in the backbone, while the CBAM mechanisms add layers of attention for identifying critical features. Survey paper [5] discusses the issue of detecting minor differences in visually similar objects, including species or models. The deep learning algorithms have drastically improved the performance of FGIA, but some challenges that include viewpoint variation, cluttered background, and feature discriminative ability are yet to be met. The paper [6] presented a channel attention mechanism. They utilized a dense set with high-resolution images of birds and other species. It applies a deep learning model along with the channel attention mechanism to well capture and amplify the most relevant features within fine-grained images. The work discusses the application of integrating deep learning with the channel attention mechanism in performing species identification tasks. The next paper [7] revolves around bird identification through Fine Grained Image Recognition, made possible by a newly innovative tool called BirdSnap to identify 500 common North American bird species. It employs one-vs-most classifiers. Birdsnap focuses particularly on North American birds and so is practical for regional bird watchers and researchers. BirdSnap demonstrates how models built with domain-tailored classification methods, coupled with spatial-temporal insights, can be quite accurate and useful in achieving an easy-to-use system for identification.

The paper [8] introduces a PLN combined with DPM for the imbalanced recognition of butterfly species. The proposed PLN-DPM model make use of two ResNet-50 networks that collaboratively adjusted parameters for better recognition of less common species. DPM reduced biases because it focused only on samples whose predictions were consistent between the two networks. In this experiment, the model had an accuracy of 86.2%, better than the top alternatives.

A high-resolution seed dataset to be used in fine-grained classification was presented in paper [9]. The dataset is developed to capture slight differences among species of seeds, which will train deep learning models such as

CNNs for identifying specific patterns and features of each seed. The dataset involves different species and environments, hence ideal for proper classification. This paper [10] employs the Stanford Cars dataset with more than 16,000 high-quality images of 196 car models for fine-grained vehicle classification. Every image is labeled with information such as make, model, and year, assisting in vehicle identification and classification. CNN models trained on this dataset are highly efficient at detecting slight visual differences between car models.

In [11], the paper introduced a dataset of more than 20,000 images, corresponding to 120 dog breeds, with detailed information for each breed. Designed on detailed information for each breed with high-grained classification. It is very frequently implemented throughout machine learning; much so in convolutional neural networks, as this relates to the application into those tasks regarding image classifications as well as detection of objects. The next paper [12] applies state-of-the-art deep learning techniques to precisely determine the maturity stages of bitter gourds. Bitter gourds have minute differences in texture, color, and shape at different maturity stages, hence not easy to determine with traditional methods. Millions of bitter gourds are cultivated and harvested, and their exact maturity at the time of harvesting is important to get a good quality yield. This model enables classification of bitter gourds into unripe, partly ripe, and fully ripened categories based on morphological fine-grained feature. This way, mature bitter gourds are selectively picked up for harvest, which avoids wastage and enhances crop quality. The Paper [13] published in year 2024 is based on how computer vision models are used in wildlife monitoring. It finds the solution for fine graining classification and course-grained classification. A popular object detection system was used for training the model, YOLOv8. The task of the fine-grained model is to classify individual species, while the coarse- grained model grouped species with similar features or roles. Conclusion Whether detailed classifications or group-level classifications must be applied depends upon the aim of the research and also upon the data. The paper [14] uses ecological data to identify and predict high-value forests. It utilizes the indicator bird species of biodiversity, which are species of birds whose existence and abundance mirror the healthiness and quality of the forest ecosystem. The research results of this paper would try to contribute toward goals of conservation by determining what to protect in the forest and how to use that information for sustainable management. Paper [15] presents an ecological relevance of bird diversity while the birds have been declining across the world. Graph pyramid attention convolutional operation is proposed to boost up the performance of YOLOv5 that has allowed birds to be detected in real-time and with precision in complicated environments. A novel backbone classification network, GPA-Net, integrated with a Cross Stage Partial (CSP) structure helps in reducing the model parameters while keeping the high accuracy intact.

III. Dataset Description

The CUB-200-2011 (Caltech-UCSD Birds 200) dataset is a comprehensive collection of images designed for fine-grained bird species classification. It includes 11,788 images of 200 bird species, with detailed annotations such as bounding boxes, part locations, and attributes, facilitating the study of fine-grained visual recognition tasks.

IV. Proposed Design and Methodology

The proposed hybrid mechanism is to be more effective for accurate bird and animal species identification. Utilizing the Caltech UCSD Dataset 200 with images of 200 North American bird species, it begins by using YOLOv8 for detecting regions of interest within the input images. The process in YOLOv8 enables fast detection with precision, making the image segmented into isolated birds and noise in the background, enhancing the focus on the right features. The segmented images are then converted into vectors, labeled and then compiled into a data frame for analysis.

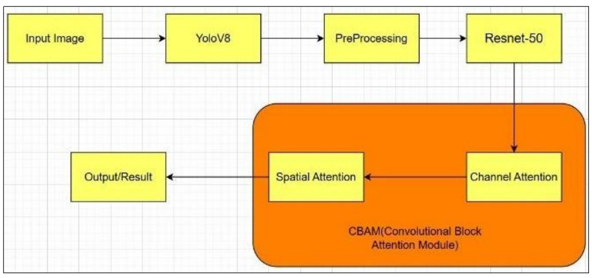


Fig 1: Framework for Image Classification using YOLOv8 and CBAM Attention Mechanism

In the Data Loader, images are resized to 3×224×224 dimensions for ResNet-50. Grayscale images are augmented by stacking two extra layers. The data is processed in batches of 16 to speed up and stabilize convergence of loss. The backbone, being ResNet-50, abstracts generalized features from images, and Convolutional Block Attention Module CBAM further refines the information. It involves channel and spatial attention, such that it helps to pay its attention on feature and regions respectively. This will facilitate the system in recognizing fine grain details so as to get very accurate classification of the species.

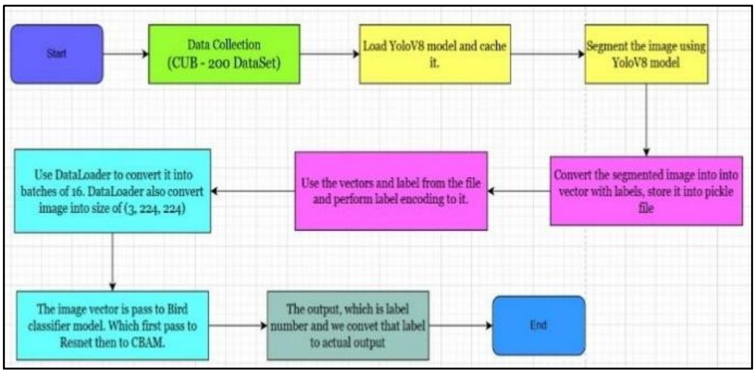


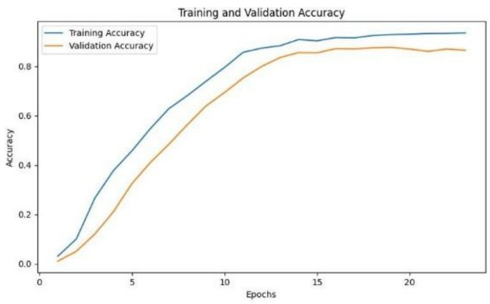
Fig 2: Step-by-Step Methodology for Bird Image Classification

V. Result and Discussion

We evaluate the performance of the model with various metrics. Accuracy is the major metric for identification of performance of classification. We also calculate metrics like Precision, Recall, and F1-Score computed for each class. These give deeper insights into how well the model can differentiate between closely related bird species.

It is important to identify the classes where the model is struggling, thus letting us fine-tune it further. Therefore, our model achieves a 88 % accuracy on Fine Grained Recognition of Bird Species using Res-Net 50 and CBAM.

A. Evaluation Metrics



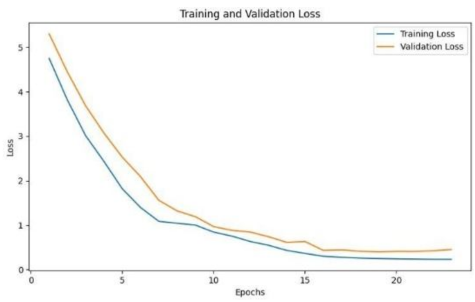


Fig 3: Training and Validation Accuracy

Fig 4: Training and Validation Loss

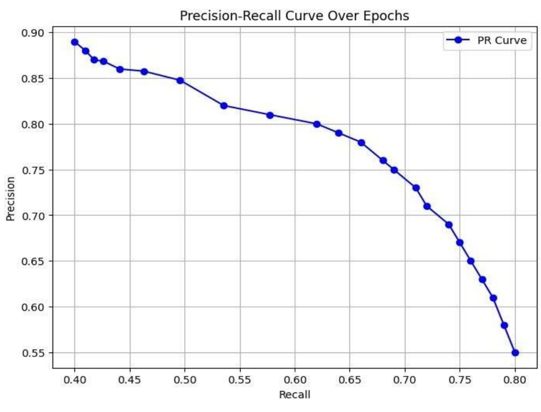


Fig 5: Precision-Recall Curve Over Epochs

V. CONCLUSION AND FUTURE SCOPE

To summarize, fine-grained image recognition has proven to be a powerful technique in identifying subtle differences between bird and animal species. Those are helpful in conservation, ecology research, and in raising public awareness about environment safety, given its various advanced models, datasets, and new methodologies. There lies an opportunity for integration into real-time applications and multimodal data collaborations with other areas, thus lending it an impressive, perhaps priceless capability, for addressing environmental issues and scientific challenges. Fine Grained Image Recognition thus becomes a versatile and sometimes revolutionary device for the effective identification of subtle differences among bird and animal species. FGIR supports biodiversity conservation and ecological research while fostering public awareness about environmental preservation. With decades of habitat loss and species extinction challenges piling up, FGIR serves as the yet-unduplicated office to address timely critical environmental and scientific concerns, as well as help secure a sustainable future with biodiversity.

REFERENCES

[1] Guo, Jingcai, Zhijie Rao, Zhi Chen, Jingren Zhou, and Dacheng Tao. "Fine-grained zero-shot learning: Advances, challenges, and prospects." arXiv preprint arXiv:2401.17766 (2024).

[2] Ibraheam, Mai, Kin Fun Li, Fayez Gebali, and Leonard E. Sielecki. "A performance comparison and enhancement of animal species detection in images with various R-CNN models." AI 2, no. 4 (2021): 552-577.

[3] Wang, Kang, Feng Yang, Zhibo Chen, Yixin Chen, and Ying Zhang. "A fine-grained bird classification method based on attention and decoupled knowledge distillation." Animals 13, no. 2 (2023): 264.

[4] Yi, Xiaomei, et al. "Research on fine-grained image recognition of birds based on improved YOLOv5." Sensors 23.19 (2023): 8204.

[5] Wei, Xiu-Shen, et al. "Fine-grained image analysis with deep learning: A survey." IEEE transactions on pattern analysis and machine intelligenxsce 44.12 (2021): 8927-8948.

[6] Yi, X., Qian, C., Wu, P., Maponde, B.T., Jiang, T., and Ge, W., "Research on Fine-Grained Image Recognition of Birds Based on Improved YOLOv5," Sensors, vol. 23, no. 19, p. 8204, 2023. doi: 10.3390/s23198204

[7] T. Berg, J. Liu, P. Belhumeur, A. Y. Torralba, and J. R. Sweeney, "BirdSnap: Large-Scale Fine-Grained Visual Categorization of Birds," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 2014, pp. 2014- 2021. doi: 10.1109/CVPR.2014.5118.

[8] Xu, Chudong, et al. "Fine-grained butterfly recognition via peer learning network with distribution-aware penalty mechanism." Animals 12.20 (2022): 2884.

[9] Zhao, X., Chen, Y., and Liu, J., "A dataset for fine-grained seed recognition," Scientific Data, vol. 11, no. 1, p. 150, 2024. doi: 10.1038/s41597-024-01234-8

[10] Krause, J., Stark, R., Deng, J., and Fei-Fei, L., "3D Object Representations for Fine-Grained Classification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1-8.

[11] A. Khosla, N. Jayadevaprakash, B. Yao, and L. Fei-Fei, "Novel dataset for Fine-Grained Image Categorization," in Proc. First Workshop on Fine-Grained Visual Categorization (FGVC), IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011

[12] Sheng Jiang, Jiangbo Ao, Hualin Yang, Fangnan Xie, Ziyi Liu, Shanglin Yang, Yichen Wei, and Xijin Deng, "Fine-grained recognition of bitter gourd maturity based on Improved YOLOv5-seg model," Scientific Reports, vol. 14, no. 1, p. 4097, 2024. doi: 10.1038/s41598-024-61635-w.

[13] M. Tam and C. Kay, "Comparing Fine-Grained and Coarse-Grained Object Detection for Ecology," arXiv preprint arXiv:2407.00018, 2024.

[14] Developing fine-grained nationwide predictions of valuable forests using biodiversity indicator bird species, Ecosphere, vol. 12, no. 12, e2505, 2021. doi: 10.1002/eap.2505.

[15] Chen, B., Yan, Y., and Liu, X., "Fine-grained Bird Diversity Recognition Based on Convolutional Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 7, pp. 2174-2186, 2024. doi: 10.1109/TNNLS.2024.1234567