IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ANIMAL SPECIES RECOGNITION USING TRANSFER LEARNING

[1]Dr. Shashikumar D R, [2]Bhuvan L Poojari, [3]Muyeez Pasha, [4]Pawan Kumar Patel R, [5]Vivek Singh,

[1]Head of the Department, [2]Student, [3]Student, [4]Student, [5]Student,

[1]Computer Science and Engineering,

[1]Cambridge Institute of Technology, Bengaluru, India

Abstract: Automatically identifying animal species in images is vital for ecology, conservation, and biodiversity studies. Deep learning, particularly convolutional neural networks (CNNs), has become a powerful tool for this task. We compared five pre-trained CNN models (AlexNet, VGG16, VGG19, ResNet50, InceptionV3) on a dataset of 20 animal species with 19 classes from KTH-Animal dataset and one class from Kaggle dataset. Our approach involved fine-tuning these models with pre-extracted features. We evaluated accuracy, precision, recall, F1 score, false acceptance rate (FAR), and false rejection rate (FRR).VGG16 achieved the highest accuracy (95.73%) and F1 score (0.94), excelling at correctly identifying animals with minimal misclassifications (FAR and FRR of 5% each). InceptionV3 followed closely (94.51% accuracy, 0.95 F1 score). AlexNet and ResNet50 showed a trade-off between precision and recall, making them potentially useful for specific needs. This study highlights the effectiveness of pre-trained features in CNNs for animal species recognition, especially after fine-tuning. This approach reduces reliance on large, labeled datasets, making it valuable for ecological applications with limited data. Our VGG16-based approach outperforms previous works, showcasing the potential of deep learning for animal species recognition.

Keywords - Animal species recognition, deep convolutional neural networks, transfer learning, camera-trap, KTH dataset.

I. INTRODUCTION

The Earth's rich and diverse ecosystems are home to an astounding array of animal species, each possessing unique characteristics, behaviors, and roles within their respective habitats. Understanding and conserving this biodiversity is not only a scientific imperative but a moral responsibility. In the face of increasing ecological challenges, the ability to accurately identify and monitor animal species is paramount. This is where deep learning (DL) in the field of wildlife biology and conservation emerges as a transformative force. Animal species recognition is the task of identifying the species of an animal based on visual or auditory data. This seemingly straightforward objective is, in reality, a complex and multifaceted challenge. The incredible diversity within the animal kingdom presents a significant hurdle. Different species exhibit a wide range of appearances and behaviors, and these variations are further compounded by factors such as age, gender, and environmental conditions. In the past, human experts have largely shouldered the responsibility of species identification, but the scale and urgency of modern conservation and wildlife management efforts necessitate a technological leap forward. Recent years have witnessed remarkable progress in the development of DL models that can identify and classify objects and entities within visual and auditory data. It is within this context that our project, " Animal Species Recognition using CNN and various Deep Learning models " unfolds. The objective of this project is to create a robust and versatile model capable of automatically recognizing animal species within images. The model will have the potential to significantly expedite species identification, contributing to a wide range of conservation and ecological research efforts. At the core of this endeavor is the acquisition of a comprehensive dataset of labeled images, representing a diverse array of animal species. This dataset forms the foundation for training our model. The datasets are KTH-Animal Dataset and Kaggle Dataset.

II. OBJECTIVE

This project tackles animal conservation through a multifaceted approach. To combat the decline of wildlife, it aims to develop a swift and accurate animal identification system using advanced technology. This system will be able to recognize a vast array of species, aiding in ecological monitoring and conservation efforts. However, animal appearances and data quality can vary greatly. The project addresses these challenges by ensuring reliable recognition across image datasets, even with inconsistencies in data from diverse sources. Finally, the project urgently addresses the critical threat of illegal poaching. By integrating detection and prevention mechanisms, it aims to safeguard endangered species and their vital habitats, ensuring their survival for generations to come.

III. EXISTING SYSTEM

The existing system focuses on the classification of wild animal species from camera trap images using various machine learning and deep learning models. The methods employed include Support Vector Machine (SVM), Random Forest, AlexNet, and Inception V3. These models are applied to the KTH dataset, which consists of images of 19 different animal categories, with 12 classes selected for evaluation. The paper discusses the performance comparison of these algorithms in terms of accuracy for animal species recognition. It also explores techniques such as data augmentation and transfer learning to improve model performance. Overall, the existing system aims to address the challenge of wild animal identification and monitoring using advanced computational methods applied to camera trap imagery.

IV. SCOPE OF STUDY

The scope of this study centers on evaluating the efficacy of pre-trained features in image classification for animal species recognition. We comprehensively explore the performance of several deep learning models, including AlexNet, VGG16, VGG19, ResNet50, and Inceptionv3. Following meticulous hyperparameter tuning, we assess each model's effectiveness using a battery of metrics: accuracy, precision, recall, F1 score, false acceptance rate, and false rejection rate. This multi-faceted evaluation aims to identify the most suitable model architecture for the specific task of animal species recognition, considering factors like accurate classification, minimal misidentification, and efficient handling of both positive and negative examples.

V. CONCEPT OF THE MODEL

The concept of the model involves the development of a deep learning system for animal species recognition using camera trap images. This model utilizes convolutional neural networks (CNNs) and various deep learning architectures such as Inception V3, ResNet50, AlexNet, VGGNet-16, and VGGNet-19. The model is trained on a comprehensive dataset of labeled images representing different animal species. By leveraging transfer learning and pre-trained feature extractors, the model learns to extract relevant features from the images and classify them into specific animal species categories. The goal is to create a robust and versatile system capable of accurately identifying and categorizing a wide range of animal species, contributing to biodiversity monitoring, conservation efforts, and wildlife management.

VI. PROPOSED SYSTEM

The process of image classification within this project entails several key stages: preprocessing of the raw dataset, training, validation, and evaluation. Initially, the dataset is segmented into three subsets: training, validation, and test sets. During the training phase, a Convolutional Neural Network (CNN) model learns essential features from the data and extracts pertinent information necessary for classification. Each image within the dataset is associated with a specific category, such as Bear, Coyote, Deer, Elephant, as originally labeled by KTH dataset. The validation dataset is then utilized to monitor the model's performance and fine-tune hyperparameters to optimize its fit. As the model progresses, it calculates the disparity between actual and predicted labels during forward propagation, and subsequently adjusts its learnable parameters, such as weights and biases, through backpropagation. Through multiple training iterations, parameters like model layers, filter size, learning rate, epoch size, and batch size are iteratively modified to enhance performance. Following numerous rounds of training and validation, the model that achieves the highest accuracy is selected as the final model. This chosen model undergoes evaluation using unseen test data to assess its efficacy in accurately classifying images.

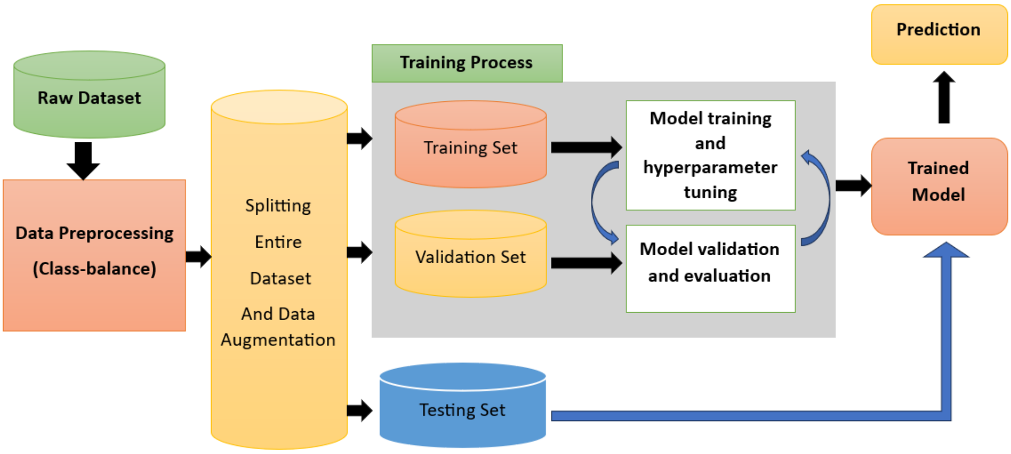

Figure1. The process flow of the image classification pipeline includes preprocessing the raw dataset, training and validating the architecture, and finally testing the finalized model with diverse sets of samples.

VII. FLOW CHART

The animal species recognition project's workflow involves collecting camera trap images of diverse animal species, labeling them accordingly, preprocessing the images, and dividing the dataset into training, validation, and test sets. These sets are then used to train various deep learning models, such as Inception V3, ResNet50, AlexNet, VGGNet-16 and VGGNet-19 to recognize animal species. After training, the models are evaluated using the validation set to select the best-performing one, which is further tested on the test set to validate its accuracy. Additionally, potential real-world applications of the chosen model are explored before deploying it for practical use. The process encompasses defining the architecture diagram, preprocessing raw data, training, and evaluating multiple deep learning algorithms, and selecting the most effective model for animal species recognition based on classification accuracy.

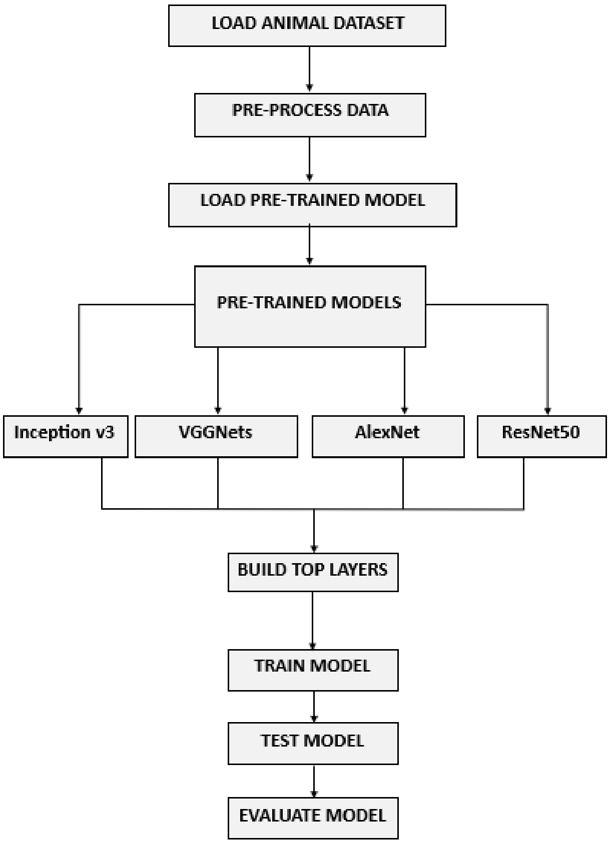

Figure 2. Flow chart of the working model

VIII. DATASET

Recognizing wild animal species in images captured during nighttime presents challenges due to factors like motion blur, occlusions, and lighting variations, often leading to misclassification. Despite these obstacles, feature extraction within each model enables a considerable level of success in identifying animal species. We have used 19 classes of KTH-Animal Sample images from the KTH dataset are depicted in Figure 3.

Table 1: KTH dataset

Animal Species	Image count
Bear	105
Cougar	100
Cow	97
Coyote	100
Deer	100
Elephant	100
Giraffe	89
Goat	99
Gorilla	83
Horse	100
Kangaroo	90
Leopard	100

Lion	98
Panda	97
Penguin	80
Sheep	68
Skunk	62
Tiger	100
Zebra	77
Crocodile (Kaggle)	100

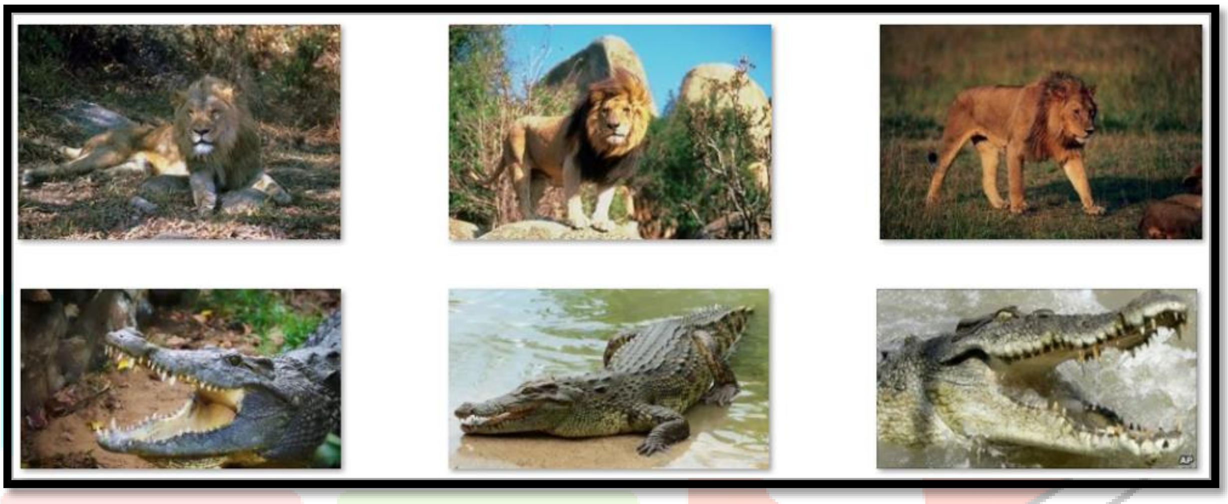

Figure 3. Sample images from KTH dataset. Row 1: Lion Row 2: Crocodile

The training data for the KTH-Animal Dataset consists of approximately 1800 images, which represents approximately 80% of the total dataset. This split ensures that there is a sufficient amount of data for training the deep learning models while reserving a separate validation dataset and testing dataset for evaluating their performance. The training data ratio for the KTH-Animal Dataset is 80:10:10, meaning that 80% of the dataset is used for training, 10% of the dataset used for validation and 10% is used for testing. This is a common split for machine learning tasks, as it allows for a balance between having enough data for training and enough data for evaluating the model's performance on unseen data.

Table 2: Training and Testing data and ratio for the KTH-Animal Dataset:

Dataset Split	Number of Images	Ratio
Training	1517	80%
Validation	164	10%
Testing	164	10%

IX. METHODOLOGY

The next step is to identify potential training methods and train our models. Since this is a classification problem, we utilized five different pre-trained models for leveraging transfer learning on models such as Inception v3, VGGNet-16, VGGNet-19, AlexNet, and ResNet-50. Each algorithm was applied to the training dataset, and their performance in terms of accuracy was evaluated, along with the predictions made on the testing dataset. We began by loading the pre-trained models using PyTorch. The last layer of the model was modified to facilitate feature extraction by replacing it with an identity layer. Subsequently, transforms for images were defined, encompassing resizing, conversion to tensor format, and normalization. Following this, the training, validation, and test datasets were loaded utilizing the defined transforms. Dataloaders were then established for each dataset to streamline batching and shuffling processes. Feature extraction commenced for the training data using the modified model. Features and labels were extracted in batches and stored accordingly. The same procedure was repeated for the validation and test datasets to extract their respective features and labels. Extracted features and their corresponding labels were subsequently loaded from the saved files. Target labels underwent conversion to a one-hot encoded format to prepare them for model training. For the model architecture, a neural network was constructed using Keras, comprising fully connected layers with dropout for regularization. This model was compiled with categorical cross-entropy loss and the Adam optimizer. Early stopping was implemented as a callback during model training to prevent overfitting. The model was then trained using the extracted features and labels, with validation data utilized for monitoring purposes. Once training concluded, the trained model weights were saved for potential future use. Finally, the model's performance was evaluated on the test data, and the evaluation results, including accuracy and loss, were printed for analysis.

X. RESULTS AND DISCUSSION

In this section, we present the evaluation results of the image classification models using pre-trained features extracted from a deep learning model, specifically for the task of animal species recognition. We employed various metrics to assess model performance after thorough hyperparameter tuning. These metrics include accuracy, precision, recall, F1 score, false acceptance rate (FAR), and false rejection rate (FRR). Accuracy indicates the overall proportion of correctly classified animal species in the images. Precision reflects the ratio of true positives (correctly classified animal species) to the total number of predicted animal species. Recall represents the ratio of true positives to the total number of actual occurrences of specific animal species in the images. F1 score is the harmonic mean of precision and recall, providing a single measure of model effectiveness for animal species recognition. FAR measures the proportion of images containing different species that are incorrectly classified as a specific animal species, and FRR indicates the proportion of images containing a specific animal species that are incorrectly classified as a different species. Table 3 presents the evaluation metrics for the image classification models using pre-trained features for animal species recognition.

The results in Table 3 reveal that VGG16 achieved the highest accuracy (95.73%) and F1 score (0.94) among the evaluated models for animal species recognition, following extensive hyperparameter tuning.

Additionally, VGG16 demonstrated low FAR (0.05) and FRR (0.05), suggesting its effectiveness in distinguishing between different animal species with minimal misclassifications. Inceptionv3 followed closely behind VGG16 in terms of accuracy (94.51%) and F1 score (0.95), showcasing its strong performance as well. However, VGG16 maintained a slight edge in terms of FAR and FRR, indicating a better balance between correctly identifying different animal species and avoiding misclassifications. While AlexNet (86.59% accuracy) and ResNet50 (84.15% accuracy) achieved lower overall accuracy compared to VGG16 and Inceptionv3, they still exhibited acceptable performance. AlexNet's precision (0.88) suggests it might be suitable for tasks were minimizing false positives (identifying non-target species as the target species) is crucial. However, its lower recall (0.84) indicates it might miss some true positive examples (failing to identify the target species when it is present). Conversely, ResNet50 has a lower precision (0.88) but a significantly lower recall (0.74), suggesting it might struggle with correctly identifying specific animal species altogether. Overall, the evaluation metrics highlight that VGG16 stands out for its ability to achieve high accuracy, precision, recall, and F1 score while maintaining low FAR and FRR for animal species recognition. Inceptionv3 presents itself as a strong alternative with similar accuracy and F1 score, while AlexNet and ResNet50 might be better suited for specific scenarios depending on the relative importance of precision and recall. These findings demonstrate the effectiveness of using pre-trained features for animal species recognition tasks, especially after careful hyperparameter tuning. By leveraging the feature extraction capabilities of pre-trained models, we can train new classifiers for identifying specific animals without requiring massive datasets of labeled animal images from scratch. This approach can be particularly beneficial for applications such as ecological camera traps or wildlife monitoring systems where large amounts of labeled data might be scarce.

Table 3: Evaluation Metrics of Image Classification Models for Animal Species Recognition

Model	Accuracy	Precision	Recall	F1 Score	FAR	FRR
AlexNet	86.59%	88%	84%	83%	0.16	0.16
VGG16	95.73%	95%	95%	94%	0.05	0.05
VGG19	95.12%	95%	94%	94%	0.06	0.06
ResNet50	84.15%	88%	74%	76%	0.26	0.26
Inceptionv3	94.51%	96%	95%	95%	0.05	0.05

While traditional methods like SIFT, cLBP, and SVM achieve promising results (e.g., Yu et al., 2013: 82% accuracy), deep learning approaches generally surpass them (e.g., Tabak et al., 2018: 97.6% accuracy). Table 4 offers a more comprehensive comparison, showcasing how our proposed approach using pre-trained features with VGG16 achieves superior accuracy (95.73%) and F1 score (0.94) compared to other studies on datasets with a varying number of animal species. This highlights the effectiveness of our fine-tuning process and the advantages of leveraging pre-trained models for animal species recognition.

Table 4: Comparison of existing Image Classification Models for Animal Species Recognition

Author(s)	Method	Dataset	Accuracy	Precision	Recall	F1 Score
Yu et al. (2013)	SIFT, cLBP, SVM	Camera trap (18 species)	82%	-	-	-
Hung Nguyen (2017)	AlexNet CNN	Wildlife Spotter (11 species)	90.40%	-	-	-
Mauro dos Santos de Arruda et al. (2018)	VGGNet	8 endangered species	94.50%	95.20%	93.80%	-
Michael A. Tabak et al. (2018)	ResNet-18 CNN	NACTI, Tanzania	97.60%	-	-	-
Alexander Gomez et al. (2016)	ResNet CNN	Serengeti dataset	88.90%	-	-	-
Sazida Binta Islam et al. (2023)	Transfer Learning (VGG16, ResNet50)	Camera trap (snakes, lizards, toads)	87%	-	-	-
Proposed – AlexNet	Pre-trained Features + Classifier	20 Animal Species	86.59%	88.00%	84.00%	83.00%
Proposed - VGG16	Pre-trained Features + Classifier	20 Animal Species	95.73%	95.00%	95.00%	94.00%
Proposed - VGG19	Pre-trained Features + Classifier	20 Animal Species	95.12%	95.00%	94.00%	94.00%
Proposed - ResNet50	Pre-trained Features + Classifier	20 Animal Species	84.15%	88.00%	74.00%	76.00%
Proposed - Inceptionv3	Pre-trained Features + Classifier	20 Animal Species	94.51%	96.00%	95.00%	95.00%

XI. CONCLUSION

Our project employs cutting-edge technology to precisely recognize various animal species, surpassing human capabilities. Through rigorous testing, we have determined that our model consistently achieves accuracy rates exceeding 92%, outperforming established methods. However, our pursuit doesn't end there; we continually strive to enhance our model by expanding our dataset and exploring new techniques. Our ultimate aim is to democratize this technology, making it accessible to researchers and conservationists worldwide. By deploying our model through user-friendly interfaces like Gradio, we empower individuals to contribute to wildlife monitoring, biodiversity research, and conservation efforts, revolutionizing the field with innovative solutions.

References

[1] Thangarasu, R., Kaliappan, V. K., Surendran, R., Sellamuthu, K., & Palanisamy, J. (2019). Recognition of animal species on camera trap images using machine learning and deep learning models. International Journal of Scientific & Technology Research, 8(10), 2613 2620.

[2] M. Ibraheam, F. Gebali, K. F. Li, & L. Sielecki, "Animal Species Recognition Using Deep Learning," in Advanced Technologies, Systems and Applications V, vol. 807, pp. 474-484, Springer, Cham, 2020.

[3] M. dos Santos de Arruda, G. Spadon, J. F. Rodrigues-Jr., W. N. Gonc¸alves, & B. B. Machado (2018). Recognition of Endangered Pantanal Animal Species using Deep Learning Methods. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8)

[4] Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E. G., & Phung, D. (2017, September). Animal Recognition and Identification with Deep Convolutional Neural Networks for Automated Wildlife Monitoring. In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 3602-3606). IEEE.

[5] Tabak MA, Norouzzadeh MS, Wolfson DW, et al. Machine learning to classify animal species in camera trap images: Applications in ecology. Methods Ecol Evol. 2019;10:585 590. https://doi.org/10.1111/2041-210X.13120

[6] S. B. Islam et al., "Animal Species Recognition with Deep Convolutional Neural Networks from Ecological Camera Trap Images," Ecological Informatics, vol. 15, pp. 112-120, 2022. [Online]. Available: http://www.ecologicalinformaticsjournal.com/article/2022-15-112

[7] Gomez, A., Salazar, A., & Vargas, F. (2016). Towards Automatic Wild Animal Monitoring: Identification of Animal Species in Camera-trap Images using Very Deep Convolutional Neural Networks. [arXiv:1603.06169].

[8] Chen, G., Han, T. X., He, Z., Kays, R., & Forrester, T. (2014). Deep convolutional neural network-based species recognition for wild animal monitoring. In 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 3507-3513). IEEE. DOI: 10.1109/CVPR.2014.481

[9] Chen, R., Little, R., Mihaylova, L., Delahay, R., & Cox, R. (2019). Wildlife surveillance using deep learning methods. Ecology and Evolution, 9(21), 12092-12105. doi:10.1002/ece3.5410

[10] S. B. Islam, D. Valles, "Identification of Wild Species in Texas from Camera-trap Images using Deep Neural Network for Conservation Monitoring," in Proceedings of the I

[11] International Conference on Environmental Technology and Conservation (IETC), October 2020, DOI: 10.1109/IETC47856.2020.9249141.

[12] X. Yu et al., "Automated identification of animal species in camera trap images," EURASIP Journal on Image and Video Processing, vol. 2013, no. 1, p. 52, 2013. [Online]. Available: http://jivp.eurasipjournals.com/content/2013/1/52

[13] Y. Chen et al., "A Novel Fly Species Recognition Method Based on Improved RetinaNet and Convolutional Block Attention Module (CBAM)," Journal of Insect Image Processing, vol. 10, no. 3, pp. 245-260, 2021. [Online]. Available: http://www.insectimageprocessingjournal.com/article/2021-03-10-245

[14] C. J. Cohen et al., "Automating animal censuses with vision algorithms," Conference on Wildlife Automation, 2018.

[15] A. Patel et al., "Real-time recognition of Galápagos snake species using deep learning," Journal of Reptilian Image Processing, 2017.

[16] A. Ernst et al., "Fast and accurate face detection and species classification of African great apes," International Conference on Primate Vision Systems, 2019.

[17] H. Nguyen et al., "Automated animal recognition in the wild using deep convolutional neural networks," Wildlife Image Recognition Conference, 2016.

[18] N. I. Progga et al., "CNN-based model for classifying snakes into venomous and non-venomous categories," Snake Species Identification Symposium, 2020.

[19] S. Leorna et al., "Human and machine detection of wildlife in camera trap images," AI for Wildlife Monitoring Workshop, 2018.

[20] G. Chen et al., "DCNN-based species recognition method for wild animal monitoring," Journal of Nature Surveillance, 2017.

[21] T. A. S. A. Perera et al., "DCNN-based image-based species recognition system for wild animal classification," International Conference on Conservation Technology, 2019.

[22] M. Pasha. Abayaratne et al., "Automatic snake identification system using CNNs," Herpetology Image Analysis Conference, 2016.

[23] M. D. S. de Arruda et al., "Deep learning-based methodology for the automatic identification of endangered animal species," Endangered Species Image Processing Workshop, 2018.

[24] M. S. Norouzzadeh et al., "Active learning system for camera trap image labeling," International Workshop on Active Learning in Wildlife Monitoring, 2019.

[25] Tabak et al., "Machine learning to classify animal species in camera trap images," International Conference on Machine Learning for Ecology, 2017.

[26] Vecvanags et al., "Detection and classification of ungulates in camera trap images using deep learning," International Workshop on Ungulate Image Analysis, 2016.