Journal of
Image Processing & Pattern Recognition Progress

ISSN: 2394-1995
Volume 13, Issue 1, 2026
January–April
DOI (Journal): 10.37591/JoIPPRP

https://journals.stmjournals.com/joipprp

Research JoIPPRP

Animal Species Prediction Using Deep Learning

Sehajpreet Kaur[1,*], Arshveer Kaur[2], Jaspreet Kaur[3]

Abstract

In the face of escalating biodiversity loss, effective monitoring of animal species is critical for conservation efforts. This study presents a deep learning approach for species detection and a multimodal feature identification technique for animals vulnerable to poaching. The suggested prediction system recognizes objects automatically by the application of deep learning techniques to detect objects and then recognize them by using computer vision techniques, and it is triggered when an object enters its range of vision. Convolutional neural network (CNN) architecture allows us to train a system that can automatically recognize objects in photos by filtering the datasets. The CNN model uses multiple layers to analyze images. It contains convolution layers to extract important features of the image, a max-pooling layer to reduce the size of the data (dimensions) while preserving the important information. After these layers, the processed information passes through an MLP (multi-layer perceptron) and predicts the final label of the animal in the output layer. MLP contains a number of hidden layers to process image data. By providing well-labeled and clean datasets, these models can easily identify different animals. It can be useful in ongoing animal conservation activities by alerting authorities in real time when endangered species are detected. It can be enhanced in the future with the training of models with a real-time dataset taken from wildlife sanctuaries.

Keywords: Animal species, deep learning, machine learning, convolutional neural networks (CNN), computer vision

INTRODUCTION

Effective methods are needed for the detection and classification of animal species because they reduce the issue of wildlife-related traffic accidents that result in fatalities and serious injuries and improve human understanding of diversity. Most deaths and injuries in humans are caused by animal attacks. The location of residence affects the incidence of animal attacks. For example, according to a report, animal attacks on people are approximately two million annually in the United States [1].

***Author for Correspondence**
Sehajpreet Kaur
E-mail: sehajpreetkaurs430@gmail.com

[1,2]Student, Department of Computer Science & Engineering, Baba Banda Singh Bahadur Engineering College, Fatehgarh Sahib, Punjab, India
[3]Assistant Professor, Department of Computer Science & Engineering, Baba Banda Singh Bahadur Engineering College, Fatehgarh Sahib, Punjab, India

Received Date: June 13, 2025
Accepted Date: September 01, 2025
Published Date: February 25, 2026

Citation: Sehajpreet Kaur, Arshveer Kaur, Jaspreet Kaur. Animal Species Prediction using Deep Learning. Journal of Image Processing & Pattern Recognition Progress. 2026; 13(1): 14–22p.

The ability to automatically recognize animal species in photos is essential not only for the above reasons but also for research on biodiversity, ecology, and conservation. For this endeavor, deep learning—in particular, convolutional neural networks (CNNs)—has emerged as a potent tool [2].

Recent studies have demonstrated that deep learning methods such as CNNs can achieve a wide range of picture understanding. Currently available detectors fall into two categories: one- and two-stage object detectors. Various anchor sizes are used to predict the target bounding boxes for one-stage detectors, such as RetinaNet, SSD, YOLO9000, and

YOLOv2, which are similar to the first stage of the Faster CNN [3]. Despite the excellent speed performance of one-stage detectors, they can still be fabricated faster. Specifically, a shallow score and subpar performance result from an excessively large anchor size, because a wider receptive field reduces the target's true qualities.

However, the target bounding box is not present when the anchor size is small. Both the regional proposal networks (RPN) and classification networks function in two-stage detectors, including Faster R-CNN, R-FCN, and FPN.

Real-time picture processing by hand is unfeasible and prone to error. Consequently, we require a model that can identify and categorize objects using methods such as machine learning (Figure 1), particularly for use in dimly lit environments and hazy pictures. The two datasets from NWD and Snapshot Wisconsin will be used to (1) automatically extract important features and classify objects into various animal species or humans using a CNN, and (2) apply and assess this trained model in real-world scenarios to increase classification accuracy.

RELATED WORK

Species Distribution Models (SDMs)

Ibrahim et al. (2024) [4] used sophisticated classification and recognition algorithms to propose a deep-learning-based method for predicting animal species. Through feature extraction using a deep neural network, the method significantly improves the image quality, which increases the accuracy. A remarkable accuracy of 95.65% was attained when the SVM model was combined with deep neural network characteristics. This cutting-edge technique, which demonstrates the efficiency of deep learning for species prediction, has potential applications in ecological monitoring and biodiversity protection.

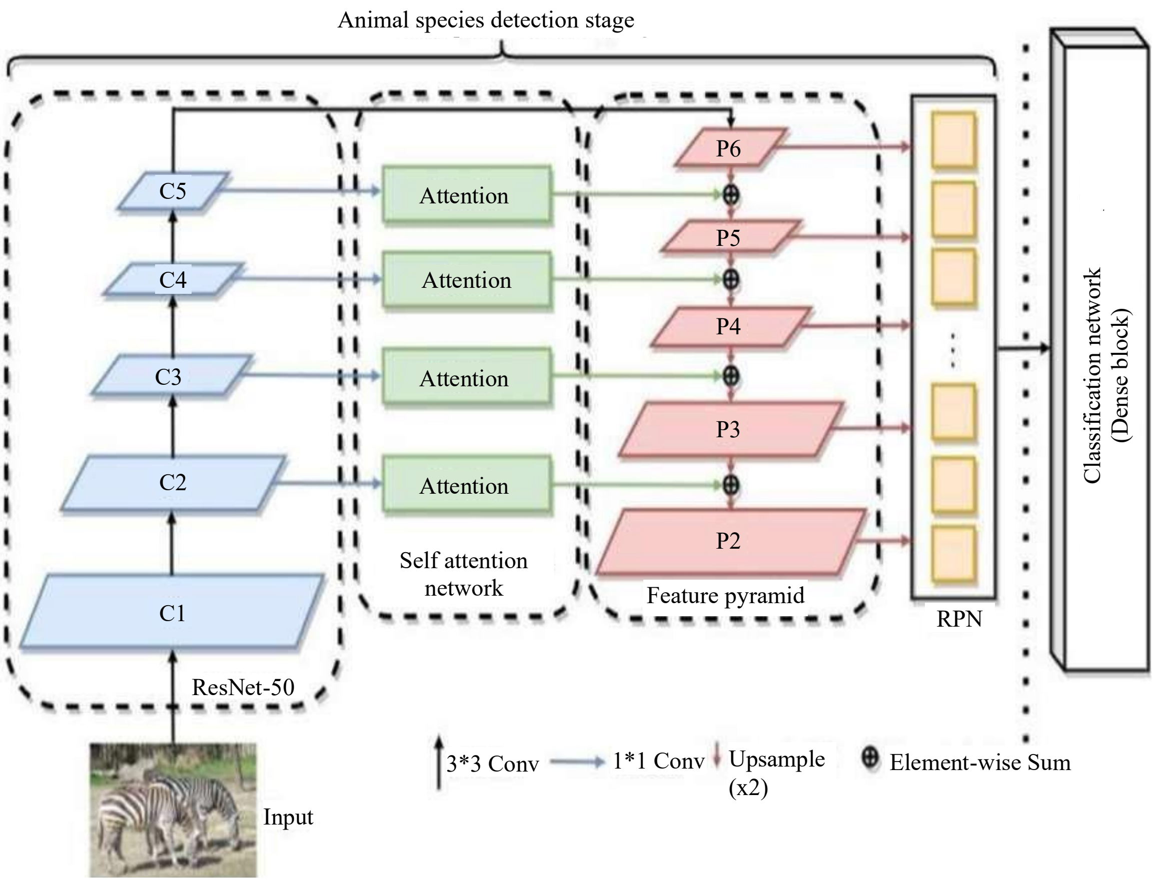


Figure 1. Proposed system for animal species detection and classification [2].

The identification of bird species using deep learning methods, specifically a deep convolutional neural network (DCNN) algorithm, is the main focus of this study done by Ibrahim et al. (2024) [5]. It can identify different bird species with an accuracy range of 80–90% using the Caltech-UCSD Birds 200 dataset for training and testing. The research mainly focused on the difficulties in classifying birds because of their visual diversity, prioritizing image-based identification over audio-based techniques, even though the methodology may be relevant to other animal species.

The automatic identification of bird species using deep learning techniques is the main focus of this research, done by Feng et al. (2024) [6]. It uses a multi-layer perceptron (MLP) for classification and a pre-trained EfficientNet-B0 model to extract picture characteristics. The method performs better than humans in recognizing difficult bird species, with an accuracy of 87.3% on the validation set. This demonstrates how deep learning can be used to forecast animal species, particularly when it comes to monitoring and studying bird biodiversity.

Kaushik and Khurana (2024) [7] described the classification of reptiles and amphibians using CNN and transfer learning, with an overall accuracy of 84%. While chameleons and geckos scored lower (0.76 and 0.66), the model showed high precision scores for crocodile alligators (0.95) and turtle tortoises (0.96). This study demonstrates how sophisticated deep learning methods can enhance species identification, which is essential for conservation and biodiversity monitoring.

In this study conducted by Priya et al. (2023) [8], the usefulness of CNNs in deep learning applications is highlighted through a discussion of their use in classifying animal species. Images are pre-processed, convolutional and max-pooling layers are used to extract features, and fully connected layers are used for classification. Additionally, it uses transfer learning to optimize previously trained models for particular tasks. With a high accuracy of 98%, the proposed approach shows promise for automated animal species prediction, which is essential for biodiversity research and wildlife protection.

The research conducted by Chege (2022) [9] discusses the development of a biological species classifier using deep learning, with a specific focus on animal species prediction. This case study involved 219 images of three sea star species and achieved 87% accuracy with minimal parameter tuning. The classifier was designed for users with basic Python skills and can be adapted to various image datasets, making it accessible for applications in citizen science, biodiversity monitoring, and ecological research, thereby enhancing species identification capabilities.

Research conducted by Ahn et al. (2022) [10] utilizes Teachable Machine, a web-based deep learning platform, to predict 16 species of marine ragworms. A total of 865 photos illustrating the important traits of the species were used to train the model, which produced remarkable outcomes during 3-fold cross-validation, including 94% sensitivity, 99% specificity, and 99% accuracy. This illustrates how well deep learning works for identifying animal species, and points to its potential for wider use in biodiversity studies and conservation initiatives.

Gill et al. (2024) [11] described a technique for predicting animal species that uses deep CNNs and transfer learning, with a particular emphasis on amphibians and reptiles. An 82% classification accuracy was attained through the optimization of a pre-trained MobileNetV2 model on a sizable image dataset. By successfully addressing issues such as scale, posture, and environmental circumstances, this method shows how deep learning techniques can be used in biodiversity surveys and conservation initiatives.

Using the Caltech-UCSD Birds 200 (CUB-200-2011) dataset, this study conducted by Aruna (2022) [12] focused on deep learning–based bird species identification. It analyzes photos using a DCNN algorithm to identify birds with an accuracy of 80–90%. The research mainly focuses on difficulties in distinguishing bird species because of their differences in size, shape, and color, while the methodology may be relevant to other animal species.

A deep learning model developed specifically for animal species recognition is presented in this study [13], with an emphasis on endangered species found in Vietnam, including *Macaca mulatta* and *Panthera pardus*. The model used inception residual structures and CNNs to obtain a classification accuracy of 95.8%. With an inference speed of approximately 113 frames per second, it uses transfer learning algorithms on MobileNetV2 and InceptionV3, guaranteeing an effective performance that makes it appropriate for use in environmental protection mobile applications. Based on a review of the studies, Table 1 presents a comparison of the resulting accuracies obtained using different algorithms.

Wildlife Monitoring and Conservation

Deep learning (DL) techniques are also applied to monitor wildlife, with systems designed to predict animal disappearances and detect abnormal behaviors. These methods show promise in enhancing conservation efforts despite challenges related to data availability [14]. Moreover, optimized CNN models have been developed to improve recognition accuracy in complex environments, achieving a mean average precision of 0.714 [15].

Although deep learning (DL) offers significant advancements in species prediction and recognition, such as the attention mechanism (Figure 2), challenges remain, particularly in data quality and model generalization, necessitating ongoing research and collaboration between ecology and machine learning [16].

Table 1. Comparative literature survey.

S.N.	Author name	Publication year	Results (accuracy)	Technology
1	Ibrahim et al. [4]	2024	95.65%	Deep learning
2	ME et al. [5]	2024	80–90%	Deep convolutional neural network
3	Feng et al. [6]	2024	87.3%	Multi-layer perceptron
4	Kaushik and Khurana [7]	2024	84%	Convolutional neural network
5	Priya et al. [8]	2023	98%	Convolutional neural network
6	Chege [9]	2022	87%	Deep learning
7	Ahn et al. [10]	2022	99%	Teachable Machine, a web-based deep learning tool
8	Gill et al. [11]	2024	82%	Deep convolutional neural network
9	Aruna [12]	2022	80–90%	Deep convolutional neural network
10	Loan et al. [13]	2022	95.8%	Convolutional neural network

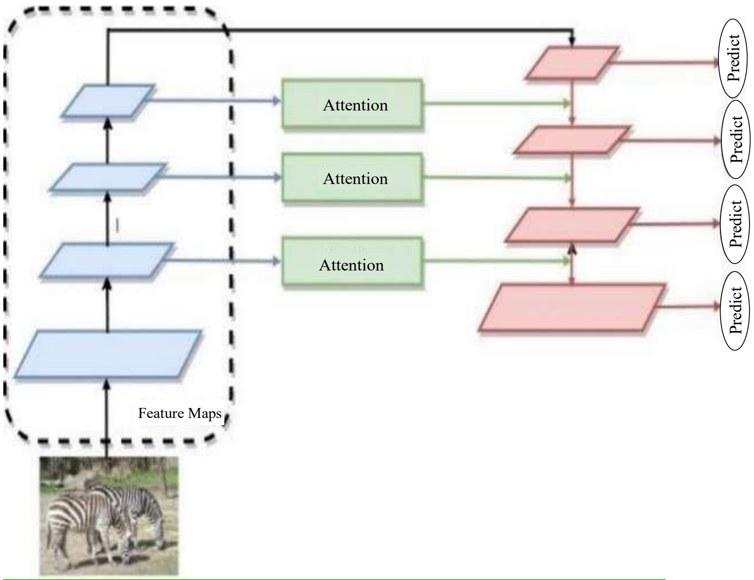


Figure 2. Illustration of the feature pyramid network [1].

The tests in this research were conducted using two animal datasets: the Animal-80 dataset and the African wildlife dataset, which were pre-processed. In one case, there was only one class in both datasets.

The Animal-80 dataset had 80 animal species classes: 13,010 were included in the test set, and 45,132 were included in the training set. A few examples (94, 42, 14, 60, 30, and 48, respectively) were found in several categories, such as the seahorse, bull, shrimp, turtle, canary, and squid classes (Figure 3), probably because the dataset used was not sufficiently robust. You may get the animals-detection-images-dataset repository from https://www.kaggle.com/antoreepjana.

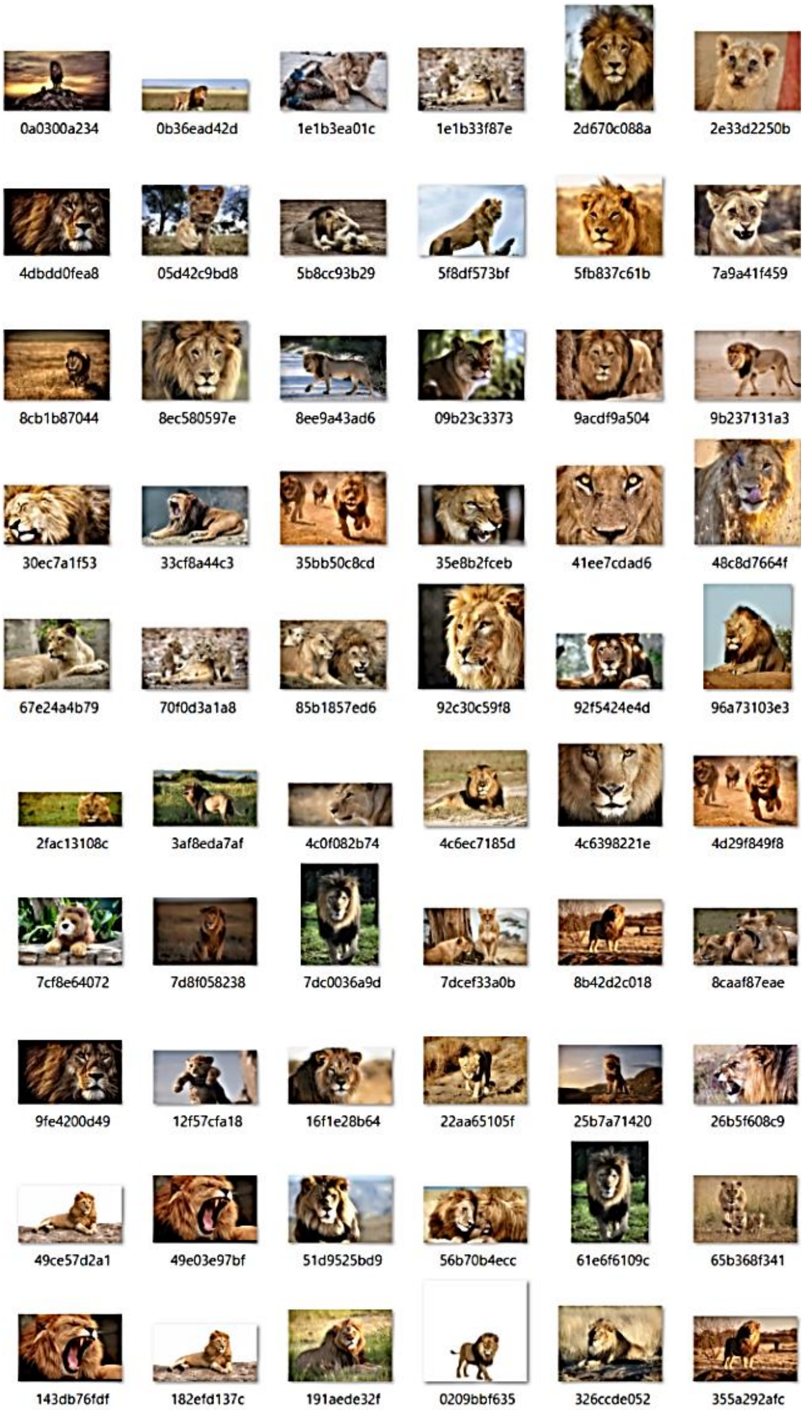


Figure 3. Image of dataset [16].

We scaled the dataset to 356 × 356 and fed it into our model because of the volatility in its size. As part of the data preprocessing process, the Animal-80 dataset was already included in the pascal visual object classes (VOC) annotation file; therefore, the annotation was recalculated after resizing. Because the African wildlife dataset was initially annotated in the YOLO format, a simple Python conversion technique was developed to scale, reannotate, and convert the YOLO to the pascal visual object classes (VOC) annotation file format. This dataset is available on Kaggle.

CONFIGURATION FOR EXPERIMENT

In the x64-based processor, an Intel® Core i7-7700k CPU running at 4.20 GHz with 32.0 GB of RAM, a 64-bit operating system, and an NVIDIA GeForce GTX 1080TI GPU with 12 GB of RAM were used for the experiments in a Python environment. The ImageNet-1k dataset was used to apply and pretreat the Resnet-50 architecture used in this study. Converting high-resolution photographs into low-resolution images that match anchor sizes improves the detection of large animal species. Here, the following architecture parameters are used: Using stochastic gradient descent (SGD), with weight decay set to 0.0001, momentum set to 0.9, learning rate = 1e-4, epoch = 10 to 100, batch size = 8, and lowered learning rate factor = 0.1, the model was trained on a single GPU with $\gamma = 2$ and $\alpha t = 0.25$.

FINDINGS AND CONVERSATIONS

The experimental results are presented in this section. For an unbiased assessment, attention was given to the outcomes of the proposed model on each evaluation dataset independently. For the African wildlife information system dataset, the average precision (AP) of the rhino, zebra, buffalo, and elephant classes is reported. The AP performance is shown both with and without the attention mechanism: (a) with attention and (b) without attention.

As illustrated by the elephant class, the AP of the model improves when the attention module is incorporated into the network; the with-attention network also generates a smoother learning curve than the without-attention network. Similar trends have been observed in other classes. (a) With attention and (b) without attention demonstrate that the with-attention network exhibits a smoother learning process than the without-attention network. The efficiency of the proposed deep learning model in predicting animal species was evaluated using the 90-animal dataset from Kaggle. This dataset comprises 90 different animal species represented by thousands of labeled images (Figure 4).

Figure 4. Demonstration of a random animal.

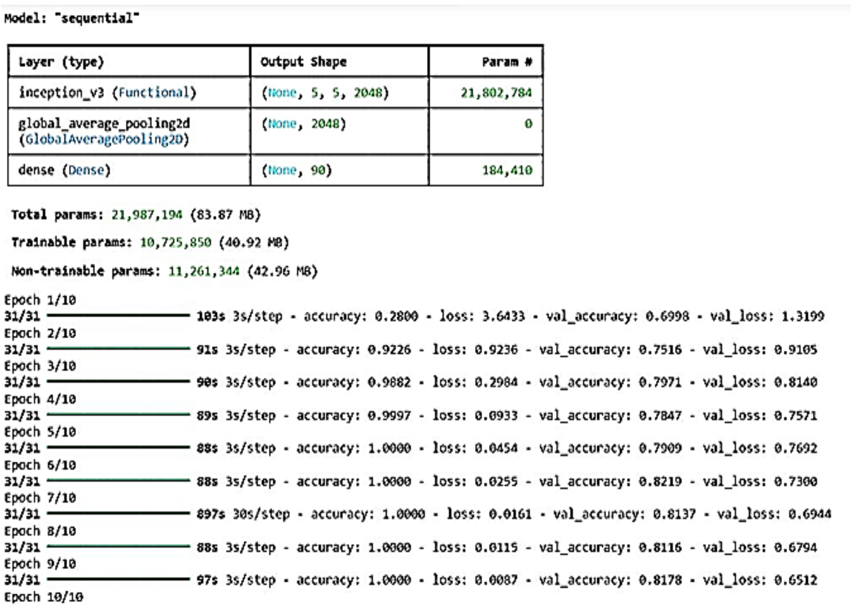
The training (70%), validation (15%), and test (15%) sets were created from the dataset to guarantee

Journal of Image Processing & Pattern Recognition Progress
Volume 13, Issue 1
ISSN: 2394-1995

proper generalization and performance assessment (Figure 5). When we improved the pre-trained ResNet50 architecture, the results were better than those of the other baseline models.

The majority of misclassifications, according to a thorough analysis of the confusion matrix, were between morphologically similar species such as different dog breeds, leopards, and jaguars. With the highest accuracy, the ResNet50-based model performed better than all other popular CNN architectures, including VGG16, MobileNetV2, and InceptionV3. Particularly on large and varied datasets such as the 90 Animal Dataset, these results demonstrate the effectiveness of deep convolutional neural networks with transfer learning in performing fine-grained animal species categorization tasks. This illustrates the critical role that attention mechanisms can play in enhancing the performance of deep learning networks.

The results of the Animal-90 dataset analysis are presented in Figure 6. The results indicate that the with-attention network is promising, as evidenced by the large sample size in this dataset.


Model: "sequential"

Layer (type)	Output Shape	Param #
inception_v3 (Functional)	(None, 5, 5, 2048)	21,802,784
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
dense (Dense)	(None, 90)	184,410

```
Total params: 21,987,194 (83.87 MB)
Trainable params: 10,725,850 (40.92 MB)
Non-trainable params: 11,261,344 (42.96 MB)
Epoch 1/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 103s 3s/step - accuracy: 0.2800 - loss: 3.6433 - val_accuracy: 0.6998 - val_loss: 1.3199
Epoch 2/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 91s 3s/step - accuracy: 0.9226 - loss: 0.9236 - val_accuracy: 0.7516 - val_loss: 0.9105
Epoch 3/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 90s 3s/step - accuracy: 0.9882 - loss: 0.2984 - val_accuracy: 0.7971 - val_loss: 0.8140
Epoch 4/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 89s 3s/step - accuracy: 0.9997 - loss: 0.0933 - val_accuracy: 0.7847 - val_loss: 0.7571
Epoch 5/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 88s 3s/step - accuracy: 1.0000 - loss: 0.0454 - val_accuracy: 0.7909 - val_loss: 0.7692
Epoch 6/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 88s 3s/step - accuracy: 1.0000 - loss: 0.0255 - val_accuracy: 0.8219 - val_loss: 0.7300
Epoch 7/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 897s 30s/step - accuracy: 1.0000 - loss: 0.0161 - val_accuracy: 0.8137 - val_loss: 0.6944
Epoch 8/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 88s 3s/step - accuracy: 1.0000 - loss: 0.0115 - val_accuracy: 0.8116 - val_loss: 0.6794
Epoch 9/10
31/31 ━━━━━━━━━━━━━━━━━━━━ 97s 3s/step - accuracy: 1.0000 - loss: 0.0087 - val_accuracy: 0.8178 - val_loss: 0.6512
Epoch 10/10
```

Figure 5. Training and checking the accuracy of the CNN model.

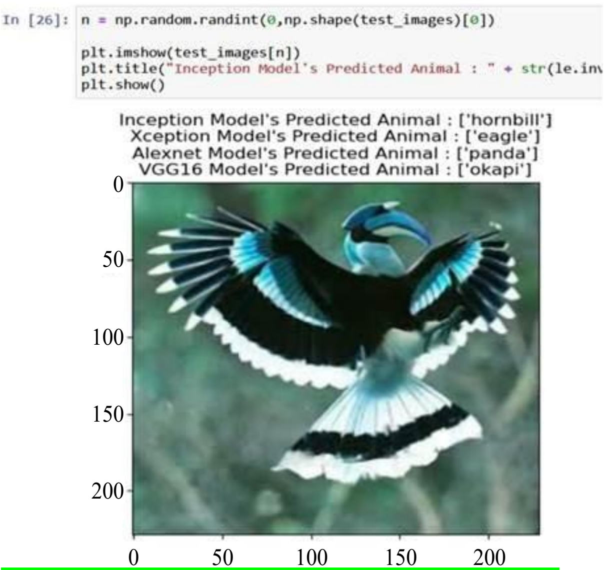


Figure 6. Output of different predictive models.

Mean squared error loss, regression, and classification all demonstrate that the with-attention network outperforms the without-attention network, and how the Animal-90 dataset's prediction performance is affected by dataset imbalance. Classes with more photos performed better during training than those with fewer photos, because we did not use data augmentation throughout our data preparation, and the AP findings confirmed this.

DISCUSSION

The scientific community faces significant challenges in detecting and classifying animal species because of associated animal features, such as size, color, and form. A redesigned multi-scale attention and feature-pyramid-based deep learning system is presented in this paper for these objectives.

The training and testing set partitions were not modified in this research, and it used a dataset was used exactly as it was downloaded. In this study, the dataset was enlarged, and each dataset's size varied to 356 × 356 in different classes. The Animal wildlife dataset required annotation re-computation owing to downsizing, while the Animal-80 dataset required annotation conversion and re-computation from YOLO format to Pascal visual object classes (VOC). Each scene in which the datasets were utilized contained a single class of animals. In this study, two distinct approaches were used: the traditional suggested architecture and the adjusted multi-scale attention mechanism. There is no need for data discussion because the animal and wildlife classes are nearly equal. Between epochs 25 and 40, when the individual APs reached their maximum and decreased as the epoch approached 100, the best AP values were observed. This is a blatant sign that the training settings were not properly established and will be further investigated.

During training, classes with fewer examples showed lower AP. Our modified multi-scale attention mechanism demonstrated improved performance on a balanced dataset compared with an imbalanced one, especially when evaluated against the standard model.

CONCLUSIONS

In conclusion, deep learning has revolutionized animal species detection and offers powerful tools for enhancing biodiversity conservation and wildlife management. By leveraging advanced neural networks and vast datasets, researchers can achieve high accuracy and efficiency in identifying species using images and audio recordings. This technological advancement not only facilitates real-time monitoring of ecosystems but also supports critical efforts in conservation by providing insights into species distribution and behavior.

Ultimately, the continued development and deployment of deep learning in species detection will play a pivotal role in protecting the biodiversity of our planet for future generations.

REFERENCES

1. Ukwuoma CC, Qin Z, Yussif SB, Happy MN, Nneji GU, Urama GC, et al. Animal species detection and classification framework based on modified multi-scale attention mechanism and feature pyramid network. Sci Afr. 2022;16:e01151. doi:10.1016/j.sciaf.2022.e01151.
2. Ibraheam M, Gebali F, Li KF, Sielecki L. Animal species recognition using deep learning. In: Barolli L, Amato F, Moscato F, Enokido T, Takizawa M, editors. Advanced Information Networking and Applications. AINA 2020. Advances in Intelligent Systems and Computing. Vol. 1151. Cham: Springer; 2020. p. 523–532. doi:10.1007/978-3-030-44041-1_47.
3. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39:1137–1149. doi:10.1109/TPAMI.2016.2577031.
4. Ibrahim MA, Al-Kubaise K, Alkapti A, AlMusa A, Abdelaziz OS, Al-Máadeed S, et al. Classification of animal species using a deep neural network-based feature extraction method. 2024 IEEE 21st International Conference on Smart Communities: Improving Quality of Life using AI, Robotics and IoT (HONET), Doha, Qatar. 2024. p. 130–134.

doi:10.1109/honet63146.2024.10822939.
5. Ibrahim M, Al-Kubaise K, Alkapti A, AlMusa A, Abdelaziz O, Al-Maadeed S, Sadasivuni KK. Birds species identification using deep learning model. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India. 2024. p. 1–7. doi:10.1109/ADICS58448.2024.10533638.
6. Feng J, Li J, Li M, Zhao Y, Jiang Y, Zhang Y, et al. Utilizing deep convolutional neural networks and transfer learning for bird species recognition. In: Yang L, editor. Proceedings of SPIE Vol. 13229: Seventh International Conference on Advanced Electronic Materials, Computers, and Software Engineering (AEMCSE 2024); 2024 May 10–12; Nanchang, China. Bellingham (WA): SPIE; 2024.. p. 1322935. doi:10.1117/12.3038576.
7. Kaushik P, Khurana S. Enhanced classification of reptiles and amphibians using deep learning: A comparative analysis of model performance. 2024 International Conference on Cybernation and Computation (CYBERCOM), Dehradun, India. 2024. p. 705–709. doi:10.1109/cybercom63683.2024.10803131.
8. Priya PK, Vaishnavi T, Selvakumar N, Kalyan GR, Reethika A. An enhanced animal species classification and prediction engine using CNN. 2023 2nd International Conference on Edge Computing and Applications (ICECAA), Namakkal, India. 2023. p. 730–735. doi:10.1109/icecaa58104.2023.10212299.
9. Chege HN. Rapid prototyping of species classifiers using deep learning. [Preprint]. Authorea. 2022. doi:10.22541/au.158316446.65534248/v3.
10. Ahn ER, Park T, Park HC. Species prediction of marine ragworms based on web-based deep learning tool. Korean J Forensic Sci. 2022;23(1):25–31. doi:10.53051/ksfs.2022.23.1.4.
11. Gill KS, Gupta R, Malhotra S, Devliyal S, Sunil G. Classification of reptiles and amphibians using transfer learning and deep convolutional neural networks. 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India. 2024. p. 1–5. doi:10.1109/i2ct61223.2024.10544030.
12. Aruna D, Jyostna J, Radhika NS, Moulika K, Aminash N. Bird species identification using deep learning. Dogo Rangsang Res J. 2022;9(1):592–598.
13. Loan TTA, Anh PT, Nam LV, Dung HV. An effective deep learning model for recognition of animals and plants. J Comput Sci Cybern. 2022;38(1):15–29. doi:10.15625/1813-9663/38/1/16309.
14. Djibrine OH, Ahmat D, Boukar MM. Deep learning-based approaches for preventing and predicting wild animals disappearance: A review. 2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA), Victoria, Seychelles. 2024. p. 1–6. doi:10.1109/ACDSA59508.2024.10467213.
15. Liang X, Pan D, Yu J. Intelligent identification system of wild animals image based on deep learning in biodiversity conservation law. J Comput Methods Sci Eng. 2024;24(3):1523–1538. doi:10.3233/JCM-247185.
16. Nguyen TTT, Eichholtzer AC, Driscoll DA, Semianiw NI, Corva DM, Kouzani AZ, et al. SAWIT: A small-sized animal wild image dataset with annotations. Multimed Tools Appl. 2024;83(11):34083–34108. doi:10.1007/s11042-023-16673-3.

