Identification and Detection of Endangered Animals and Their Actions Recognition using Deep Learning Method

R. Nedunchezhian[1]
Professor, Department of CSE
Coimbatore Institute of Technology, Coimbatore, India
nedunchezhian@cit.edu.in

K. Harinisankari[2]
Student, Department of CSE
Coimbatore Institute of Technology, Coimbatore, India
1805014cse@cit.edu.in

B. Ishwariya[3]
Student, Department of CSE
Coimbatore Institute of Technology, Coimbatore, India
1805017cse@cit.edu.in

M. Madhumathi[4]
Student, Department of CSE
Coimbatore Institute of Technology, Coimbatore, India
1805026cse@cit.edu.in

M. Yuvarekha[5]
Student, Department of CSE
Coimbatore Institute of Technology, Coimbatore, India
1805064cse@cit.edu.in

ABSTRACT

Across the world, several ecologists are very much interested in detecting and analyzing the behavior of the endangered animals for further study related to the conservation of ecosystem and wildlife. Though the entire data is available in hand, it is impossible for a person to analyze what an animal does all time. This information can be extracted from the camera trap images. But it is a time consuming and tedious to check the behavior of the animal. From the literature survey, it is true that a deep learning model will serve this purpose. A deep learning model is constructed to detect animals and recognize their actions. For experimental purpose, four animals namely tiger, Cheetah, Hyena, Jaguar are taken to detect them and also recognize four actions such as Eating, Sleeping, Standing, Sitting. Convolutional Neural Network with pretrained Inceptionv3 algorithm is applied to develop the deep neural network model. The model is trained on 4000 image datasets and an accuracy of 99.34 % is obtained for or animal detection and 99.72% for action recognition. This measure highlights the importance of using deep neural network against traditional analysis of camera trap images. The empirical analysis is done on four endangered animals only due to their close resemblances. This model is scalable and can be extended to analyze any other animal of interest. It is fast, accurate and it saves a lot of time in analysis. The performance of the model is very efficient compared to the model trained of VGG architecture.

KEYWORDS: Endangered Animals, Action Recognition, Animal Detection, CNN, Inceptionv3.

I. INTRODUCTION

The conservation of biodiversity and ecosystem is essential for the balanced living of human beings and other species that are interlinked in the food web. The ecosystem can be conserved only when we have adequate knowledge about the organisms in a specified territory and the behavior of the animal under various circumstances is available.

The animals behave differently in different situations. They locomote at certain period. Their preferences vary over time. The overall analysis of animal behavior will help the authorities to take necessary actions to conserve them. In a food web, if one species goes extinct, it entirely collapses the entire cycle. Hence it is the need of the hour to conserve the wildlife organisms and the habitat.

Ecologists are scientists who study about the ecosystem. They will study the connection and interrelationship between various living organisms from microscopic organisms to plants to animals. It is impossible for the ecologists to monitor the animal activity all the time. Instead of monitoring the animal activity physically, the images captured through the camera traps in the wildlife system can be analyzed by the advanced computer vision methods. If the image count is 100, human could analyze it. But when the count is too large, it is a time consuming and a tedious task.

To automate this process, a deep neural network can be used. The trained deep neural network will automatically analyze the images from the input feed and identify the animal present in it. In addition to that it will also recognize the actions of the animal.

The deep neural network is a mathematical model which is trained to process data in a sophisticated manner.

This paper is organized in six sections. In Section II, the relevant work to this research is discussed. It deals with the recent work done in this field. In Section III, the methodologies used in the system is explained. Section IV provides the details of the techniques implemented in the project and the relevant information about their performance of the system. Section V contains the results of the system, while Section VI concludes through some information on the system's future scope.

II. LITERATURE SURVEY

This section outlines several related efforts on deep learning for animal detection and action recognition. It focuses mostly on current research studies on object detection using machine learning methods. Animal detection can be viewed as an object detection in the computer vision process of classifying and localising its shape in an image.

A model using Resnet-18 neural network has been proposed [1]. Three actions were recognised using the Resnet architecture. The accuracy of the model ranging between 88.4% to 94.1%. This model classifies only four types of animals and three types of their actions. Moreover, this model is trained on a scarce dataset.

A deep CNN model was built to identify, count and describe the animal behaviour[2]. VGG 16 architecture was used and a huge dataset of 1.2 million was used to train the model. The accuracy of the trained model is 96.8%.

Another paper proposed that Deep Convolution Neural Network is better for animal detection problems [3]. In this paper several approaches were analysed to observe the locomotive behaviour of animals and deep CNN was found a better approach. The trained model gives an accuracy of 91%. The trained model classifies the animal in both day and night time.

The importance of animal detection problem in real time to prevent animal-vehicle accidents was researched [4]. Convolution Neural Network algorithm was used to detect the animals with an accuracy of 99%. The model proposed just works with individual species and will not work well for multiclass species identification.

A deep learning approach for automatically identifying and isolating the species-specific activity from the still images was developed[5]. Two different architectures of CNN were used and trained on 8368 image dataset. An accuracy of 98.05% was achieved.

III. METHODOLOGY

1.CONVOLUTION NEURAL NETWORK

Convolutional Neural Network is a deep learning technique that is primarily used for the classification of image datasets. The figure 1 shows the architecture of the convolution neural network. Each input image will pass through a series of convolution layers with filters, Pooling, fully connected layers and apply SoftMax function to classify an object with probabilistic values between 0 and 1. The feature extraction component in this architecture would be the combination of '**convolution + pooling**'. These feature extractors, first extract low-level features (like edges, lines) then mid-level features as shapes or combinations from several low-level

features and eventually high-level features. Finally, these layers are flattened and connected to the output layer using an activation function.

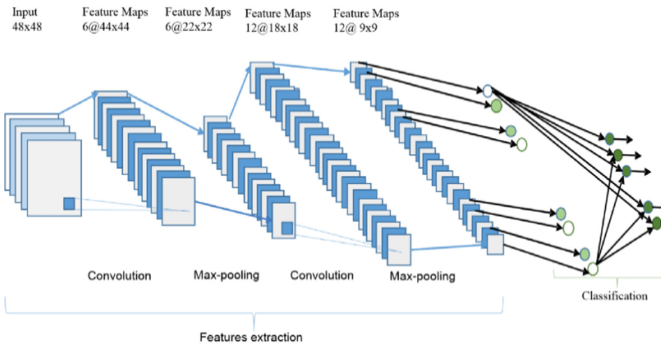

Figure 1 : Convolution Neural Network Architecture. Source[6].

2. INCEPTIONV3

Inceptionv3 is the third version in the series of deep CNN architecture. It was trained on about 1 million ImageNet dataset. The architecture diagram of Inceptionv3 is illustrated in figure 2. Inception v3 has 9 inception modules linearly stacked making it 22 layers deep. Inception Modules are image blocks that allows to use multiple filter types which concatenates the output and pass it on to next layer. It has very high computation power which is very expensive to achieve.

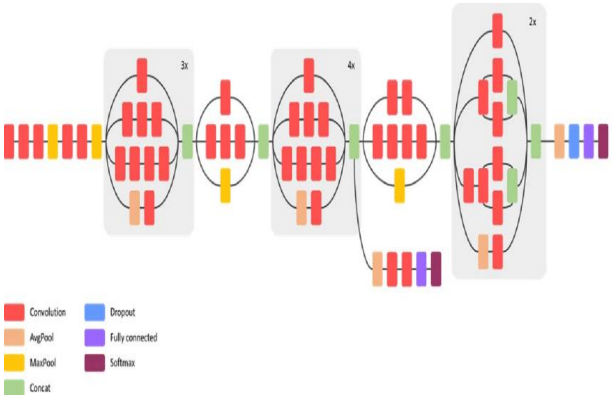

Figure 2: InceptionV3 Architecture. Source [7].

IV. PROPOSED SOLUTION

A deep learning model using Convolution Neural Network and pre-trained Inceptionv3 algorithm is constructed to classify the animals and to recognise the animal actions. The project is split into two different phases as animal detection and action recognition. The final model is a combination of both the phases. The diagram in figure 3 depicts the flow of the solution proposed in this paper.

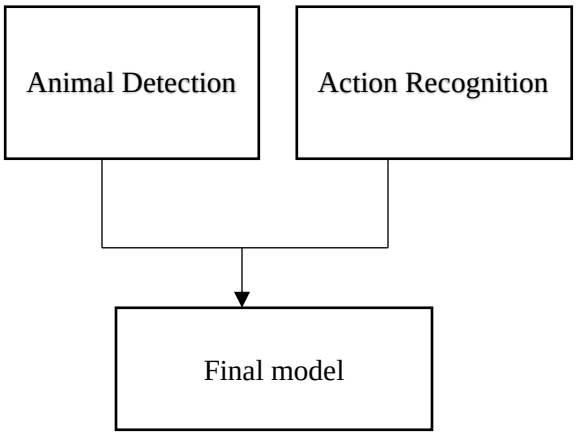

Figure 3: Phases of the model

PHASE 1: ANIMAL DETECTION

The scope of the project is to detect and identify four animals – tiger, cheetah, jaguar, and hyena. The four animals have close resemblances which is very challenging for a human to classify and hence they were chosen for the project. For the animal detection phase a dataset containing 2800 images of the above-mentioned animals were collected.

Figure 4: Images from the animal detection dataset showing cheetah, hyena, tiger and jaguar(Left to right). Source[8]

Some of the sample images of the collected dataset was given in the figure 4. The data was then augmented to generalize the model and make it learn better. The CNN model was then trained with the collected and augmented dataset. An accuracy of 99.36% was attained on training and the validation accuracy was 99.72%. The figure 5 gives the training accuracy plot. When the model was trained using VGG16 architecture, the accuracy of the trained model was 96%. But when the model is trained using Inception v3, the accuracy was improved to 99.36%. The reason for the improvement in accuracy is, in the VGG16 architecture we have 16 layers. But in the case of Inception v3, we have 42 layers.

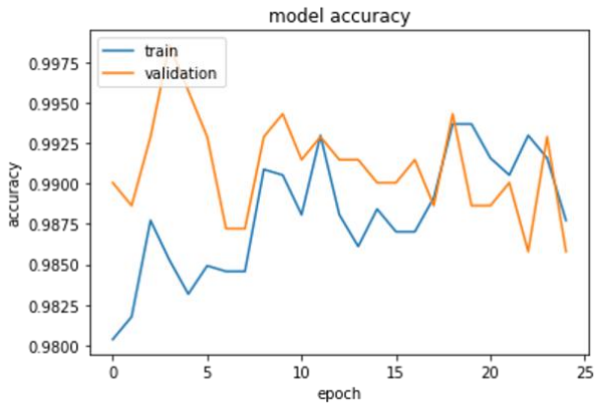

Figure 5: Accuracy of the animal detection model.

PHASE 2: ACTION RECOGNITION

The following are the actions recognized by the model – Eating, sleeping, standing, and sitting. The model could be extended to recognize any actions. For the action recognition phase, a dataset containing the same animals with four actions were collected and augmented. The dataset contains 4000 images of the animal actions. The dataset is then augmented by rescaling, zooming and shifting the images. This makes the model not to learn a pattern instead it will be able to learn the features.

Figure 6: Images from the action recognition dataset showing sitting, standing, sleeping and eating(Left to right).Source[9]

Some of the images of the collected dataset was given in the figure 6. The CNN model was then trained with the collected and augmented dataset. The model was trained for 25 epochs. An accuracy of 99.72% was attained on training and the validation accuracy was 96.86%. The figure 7 gives the accuracy plot of training and validation.

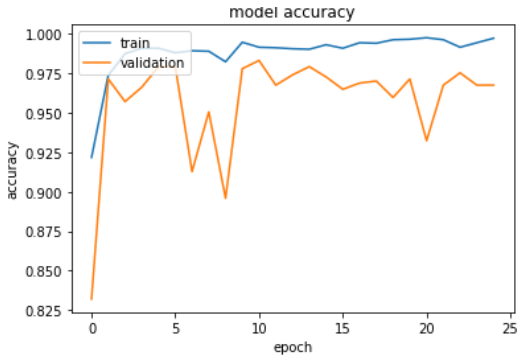

Figure 7: Accuracy of the model action recognition model.

V. RESULT

The trained model was then tested against images which were not used in the training and the model was able to detect and classify the different animals better. About 700 images were used for the testing. The test image was passed to both the trained models and predictions of both the models were saved in the form of a document along with the test image. The figure 8 gives the predictions obtained from the test image.

The InceptionV3 performs well in terms of accuracy and the computation speed was also good when compared with VGG architectures. The VGG architecture couldn't extract complex features whereas the Inceptionv3 is able to do complex tasks. Inceptionv3 has more hidden layers and hence the model is be able to learn better compared with the VGG architecture.

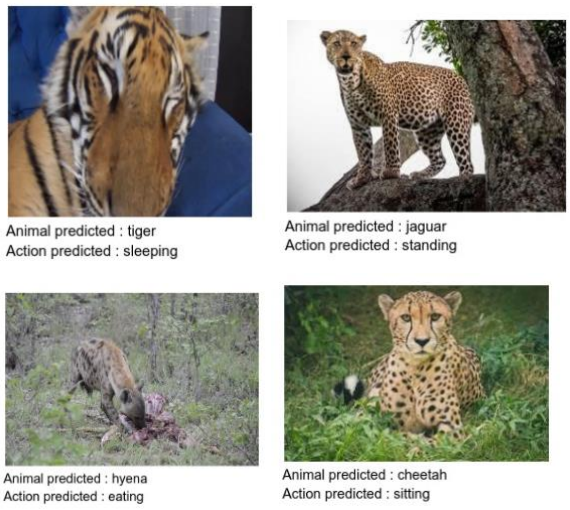

Figure 8: Predictions made by the trained model.

VI. CONCLUSION AND FUTURE WORKS

Although the system results in terms of good accuracy and positive predictions, the following are some of the corrective actions identified:

There is a scarcity of dataset which can be utilised for this system, and it is not varied enough to work in all other possible cases.

The scope of the project is limited to four animals and four actions. It should be extended to many animals.

The model was tested on still images obtained from the camera traps. Further work is to focus on testing it against video stream as input.

REFERENCES

[1] Frank Schindler, Volker Steinhage "Identification of animals and recognition of their actions in wildlife videos using deep learning techniques", Elsevier March 2021

[2] Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson , Meredith S. Palmere , Craig Packer and Jeff Clune "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning", PNAS Nov 2017

[3] Elham Mohammed Thabit A. ALSADI, Nidhal K. El Abbadi "Scrutiny of Methods for Image Detection and Recognition of Different Species of Animals", IJRTE, Nov 2019

[4] N. Banupriya, S.Saranya, Rashmi Swaminathan, Sanchita Harikumar and Sukitha Palanisamy "Animal detection using deep learning", JCR Dec 2019

[5] Ruilong Chen,Ruth Little,Lyudmila Mihaylova,Richard Delahay and Ruth Cox "Wildlife Surveillance using Deep Learning Methods", Ecology and Evolution May 2019

[6] Md. Zahangir Alom, Tarek M.Taha, Chris Yakopcic and Stefan Westberg "A State-of-the-Art Survey on Deep Learning Theory and Architectures", Electronics March 2019.

[7] Mohammad Shakirul Islama, Ferdouse Ahmed Foysala , Nafis Neehal, Enamul Karima and Syed Akhter Hossain, "InceptB: A CNN Based Classification Approach for Recognizing Traditional Bengali Games" , May 2018.

[8]Kaggle-https://www.kaggle.com/datasets/iluvchicken/cheetah-jaguar-and-tiger

[9] www.youtube.com