# ANIMAL SPECIES RECOGNITION SYSTEM USING DEEP LEARNING

**SUHAS SUDHEER SHENOY[1], ASHWINI P PATIL[2]**

*[1]Student Dept. of Master of Computer Applications CMR Institute of Technology, Bengaluru*
*[2]Asst.Professor, Dept. of Master of Computer Applications CMR Institute of Technology, Bengaluru*

---

**Abstract -** Animal watching is a common sport, but you must use an animal book to identify the species. We established a deep learning platform to help users in recognising endemic animal species, as well as an app called Imagenet of Animals to provide animal lovers with a handy tool for appreciating the beauty of animals (IoA). To recognise essential traits in animal pictures, a convolutional neural network (CNN) was trained. We initially designed and generated a restricted zone The shapes and colours of the levels of granularity are intriguing. before balancing the distribution of Animal species. To improve feature extraction, the outputs of the previous and subsequent layers were linearly merged using the skip connection technique.

***Key Words***: Convolution neural network (CNN), Tensorflow (TF), Rectified linear unit (ReLU), Fully connected (FC), Imagenet, Central Processing Unit (CPU), Graphical Processing Unit (GPU), GoogleNet Inception V4 network

## 1. INTRODUCTION

In recent decades, pattern classification has become among the most essential areas of artificial intelligence, playing a crucial role in a variety of real-world applications.

Contrary to the natural intelligence exhibited by people and other animals, artificial intelligence (AI), sometimes referred to as machine intelligence, is intelligence expressed by machines. The phrase "intelligent agent" refers to any machine that can perceive Adapt to its environment and take efforts to maximise the chance of success in achieving its goals. The term "artificial intelligence" is used when a machine duplicates "cognitive" activities such as "learning" and "problem solving," which are frequently associated with other human brains.

The Animals Recognizer utilises images as input to learn and recognise the animals, however deep learning normally need a heavy amount of training dataset, and the training process is time-consuming and slow. As a consequence, the GoogLeNet inception v4 network is utilised to extract the Animal features. The animals are then categorised using the attributes that were gathered by assessing how closely they resemble the template. Studies show that our system can accurately identify a specific animal's species after learning around 10 samples of that animal. The deep learning incremental framework gives the self-learning system a wide range of application and customization.

Artificial intelligence (AI) is the study of the way the human brain makes decisions, learns, and thinks. The study also results in clever software systems. The objective of AI is to improve computer abilities that are comparable to those of the human brain, such as reasoning, learning, and problem-solving.

### 1.1 Aim and Objective

The project's goal is to identify different animal species using deep CNN. Because there are so many different kinds of animals, manually recognising them may be difficult. Data mining limits are studied in this part of the research, and an unique technique termed Lightweight Machine Convolution Network for Animal Recognition is devised. It will have low computational expenses and great precision.

### 1.2 Proposed System

The proposed approach employs the ImageNet Inception v4 network and CNN to identify animal traits, as well as TensorFlow to allow our recognizer to recognise any species.

## 2. LITERATURE SURVEY

Yao, S. et al. [1] introduced fine grained visual categorisation, which tries to distinguish between objects of the same kind. With just the original image as its input, this novel description was able to automatically provide visual depictions that are distinct enough for precise visual categorization. Fine-grained visual categorization has a number of drawbacks, but its fundamental drawback is that it is computationally expensive and unsuited for large-scale picture classification. Xie. Et al. [2] proposes that instance search should not produce only nearly identical photos while also providing fine-grained results, which is typically what a user wants. It creates a large-scale database with the reference pictures compressed at consistent bit rate levels using JPEG encoders using various optimization techniques, introducing a baseline system employing fine-grained classification scores. In subjective tests, the comparison approach is used to rank them in order to identify tiny differences. The primary problem with fine-grained findings is that classification of items that belong to the same species duplicates the data. The following paper discusses the template matching [3] algorithm, which is used to pinpoint certain details in a picture that should match the template image. Information regarding the locations of the image's interesting items is given, along with a list of those things.

As a result, it took time. A way for motion animal detection using the advantages of global patterns of pixel motion is proposed by Fang, Y., et al. in [4]. In this study, segmented areas were employed, and a different threshold was applied to screen out unfavourable candidates who may be from the background. The intricacy is the primary negative. In order to develop a classifier that optimises the margin and consistency over all data and employed labelling techniques, J. Tanha et al. [5] offer a modestly learning multiclass technique that applies a similarity function to all data and a basic classifier. There were labels for each and every feature. The primary flaw was the laborious and difficult labelling process. In order to construct a system for automatically observing animals,H. Nguyen et al. [7] provide a framework for establishing automatic wild animal identification. An advanced deep convolutional neural network architecture and a single-labeled dataset were used to train a computer system that can recognise species and filter animal photographs on its own. Since the records are not updated, the biggest drawback of the wildlife spotter project's datasets is that hybrid animals cannot use them. J Deng et al[8] .'s deep research of ImageNet in its present state, which consists of 12 subtrees, 5,247 synsets, and 32 lakh images, is presented. It talks about a brand-new data-base called "ImageNet," which is a sizable ontology of images constructed using the worldnet architecture.It discusses a brand-new database called "ImageNet," which is a substantial concept of images built on top of the worldnet architecture. ImageNet aims to contribute 500-1000 clear, high-resolution photographs to each of the worldnet's 80,000 synsets. Akshay Kapoor et al. [9] propose several design principles and alterations offered in the GoogLeNet and inception networks. The computational efficiency of these modifications is investigated, and network features and performances are compared on the ImageNet dataset. A critical assessment of inception networks is also provided.

## 3. METHODOLOGY

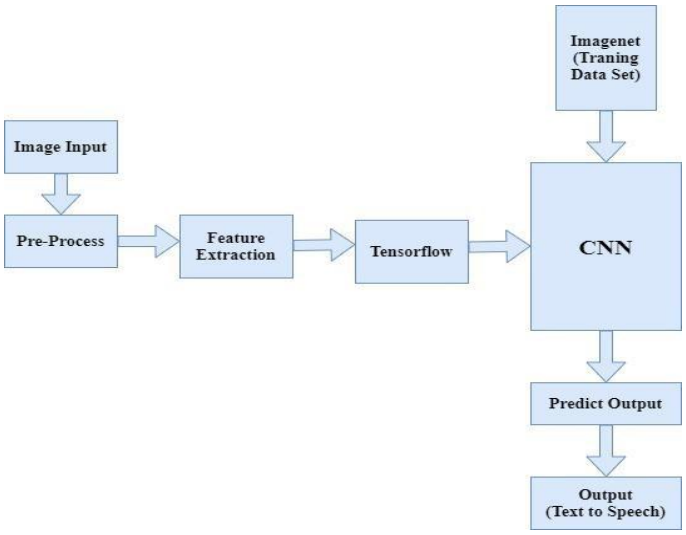


**Fig -1:** System Architecture of Animal species recognition system

At the beginning user has to upload image i.e., any animal image including land animals, aquatic animals and birds. Preprocess is stage where raw image given by the user has to be processed before it is fed into learning purpose, to enhance the image for accurate prediction. The entire image is divided into tiny objects for which features are extracted using deeplearninglog4j. Then, Tensorflow is used to convert the extracted features into a graph values or pb(protobuf) values. This conversion is needed as the ImageNet datasets contains the image in form of pb values. CNN performs comparisons between the input data and the training dataset and predicts the output of the animal species with the accuracy percentage. Text to speech converter is used for converting output of the recognized animal species into human like voice.

## 4. MODULES

**4.1. CNN (Convolution Neural Network)** Convolutional neural networks are sophisticated artificial neural systems that are frequently used to classify images, group them based on similarities, and identify animals in context. In this module based on the extracted features of animal image given by the user , CNN classifies the animal image and then identifies the species of the animal. The Fig-2 The three-layer, The CNN architecture with convolution, sharing, and complete connectivity is shown. The convolution layer employs a number of filters to generate a wide range of activation characteristics on the input image. Pooling layers, like convolution layers, have a specific function, such as max pooling, which takes the value at its greatest level in a specified filter zone, or average pooling, which takes the value at its average level in the same region. These are typically employed to make the network's dimensions smaller. In the picture above, maximum pooling is used. Fully connected layers are inserted before the CNN's classification output to flatten the data prior to classification.

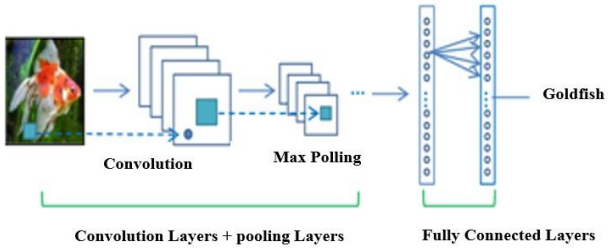


**Fig -2:** Working of CNN

### 4.2 Feature extraction based on DeepLearningLog4j

Feature is extracted for the animal image which is preprocessed using Deep Learning Log4J algorithm and helps in classification of the image based on the extracted features. Deep neural networks have the ability to extract more robust and discriminative features, which allows

recognition systems based on these characteristics to train more quickly and with higher recognition rates. The feature extraction speed and identification accuracy both increase when the deep neural network architecture is used. GoogLeNet inception v4 is chosen to conduct the comparison experiment and extract photo features since it performs well at identifying animals. On ImageNet, these networks have prior training.

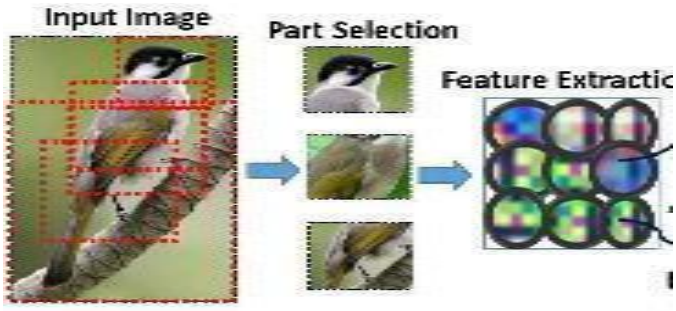


**Fig-3** Feature Extraction

### 4.3 TensorFlow

TensorFlow, a free and open-source software framework, is used for differentiable programming and data flow in a number of activities. It is an opensource artificial intelligence toolkit that uses data,flow graphs to build models. It allows for the creation of multilayered, large-scale neural networks. TensorFlow's main applications include classification, perception, comprehension, discovery, prediction, and creativity. In this module, the pre-processed animal picture will be converted into graph values so that the animal species may be determined.

### 4.4 ImageNet Training Dataset

The ImageNet LargeScale VisualRecognition Contest is the industry benchmark for object recognition and classification, including millions of photos and hundreds of item classes. The winner, GoogleNet, had the best performance to date, bringing the classification error down to 0.06656 and raising the average object detection accuracy to 0.439329. Its network had roughly 30 layers. Convolutional neural networks performed similarly to humans on the ImageNet tests. We utilise GoogleNet to get images of animals as training samples for our program since there are many different animal species.

### 4.5 Text to Speech Converter

Assistive technology is demonstrated via text into speech program, which reads digital text aloud (TTS). It is also known as "read it loud technologes." TTS converts text on a audio from a laptop or desktop device a single click or button push. Using a different function, the result of the animal photo identification is converted into a human-sounding voice.

## 5. RESULTS AND DISCUSSON

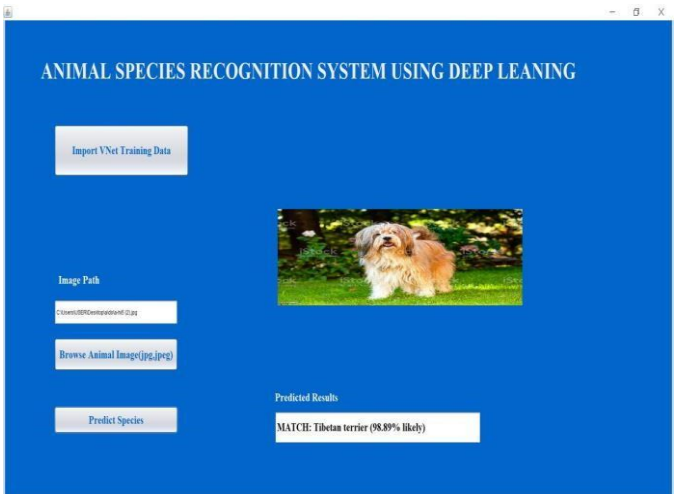


**Fig-4** Output Generated for Land Animal

First the datasets are imported by clicking “Import VNet Training Data” and image is browsed using ”Browse Animal Image” button and the path of the image is displayed in the “Image Path” box. Fig-4 shows the result when a land animal is uploaded. This represents the output generated “MATCH: Tibetan terrier (98.89% likely)” for the land animal along with accuracy percentage after the “Predict Species” button is pressed.

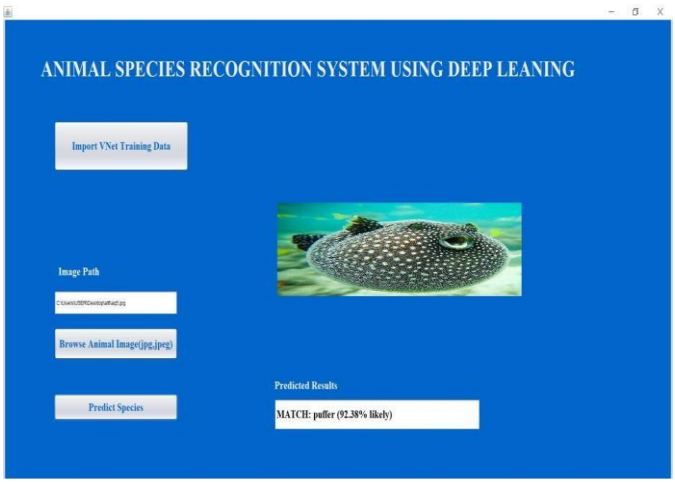


**Fig-5** Result for an aquatic animal

Fig-5 represents the output generated “MATCH: puffer (92.38% likely)” for the aquatic animal along with accuracy percentage after the “Predict Species” button is pressed.

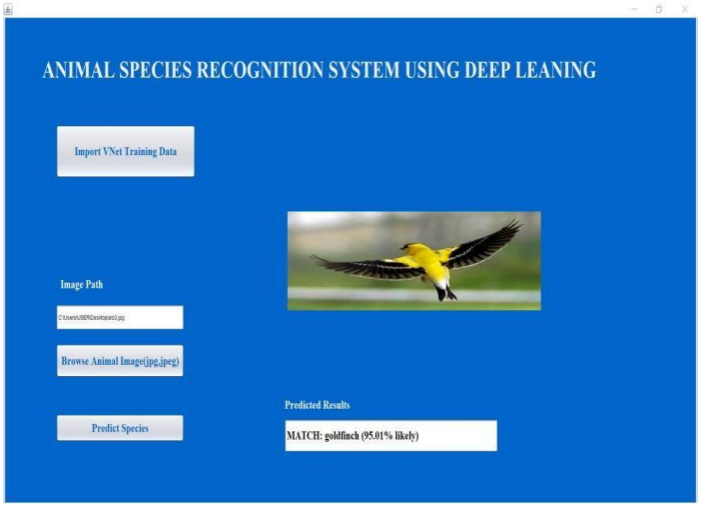


**Fig-6** Output generated for Bird

Fig-6 represents the output generated "MATCH: goldfinch (95.01% likely)" for the bird along with accuracy percentage after the "Predict Species" button is pressed.

## 6. CONCLUSIONS

With the remarkable potential of neural network, the proposed system can recognize animals of all species this includes land animals, aquatic animals and birds. Training datasets obtained from GoogleNet inception model will be in terms of graph values so the graph values for the input image is obtained by using the algorithm specified. Animal species recognition system is implemented with the help of CNN, Google Net inception v4 helps to achieve a higher learning rate and results in faster overall performance.

## REFERENCES

[1] H. Yao, S. Zhang, Y. Zhang, J. Li and Q. Tian, "Coarse-toFine Description for Fine-Grained Visual Categorization," in IEEE Transactions on Image Processing, vol. 25, no. 10, pp. 4858-4872, Oct. 2016

[2] L. Xie, J. Wang, B. Zhang and Q. Tian, "Fine-Grained Image Search," in IEEE Transactions on Multimedia, vol

[3] Prakash, Banupriya, "Animal Detection Using Deep Learning Algorithm"

[4] Fang, Y., Du, S., Abdoola, R., Djouani, K., & Richards, C. (2016). "Motion based animal detection in aerial videos."Procedia Computer Science, 92, 13-17.

[5] J. Tanha, M. V. Someren, M. d. Bakker, W. Bouteny, J. Shamoun-Baranesy and H. Afsarmanesh, "Multiclass Semisupervised Learning for Animal Behavior Recognition from Accelerometer Data," 2012 IEEE 24th International Conference on Tools with Artificial Intelligence, Athens, 2012

[6] F. Tu, S. Yin, P. Ouyang, S. Tang, L. Liu and S. Wei, "Deep Convolutional Neural Network Architecture With Reconfigurable Computation Patterns," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 25, no. 8, pp.2220-2233,Aug.2017, doi: 10.1109/TVLSI.2017.2688340.

[7] Hung Nguyen, S. Maclagan, T. Nguyen, Thin Nguyen, P. Flemons, Kylie Andrews, E. Ritchie and Dinh Q. Phung, "Animal Recognition and Identification with Deep Convolutional Neural Networks for Automated Wildlife Monitoring," 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, 2017, pp. 40-49, doi: 10.1109/DSAA.2017.31

[8] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[9] Akshay Kapoor, Pandit, Tejas, 2020/02/24, Understanding inception network architecture for image classification, Research gate, doi: 10.13140/RG.2.2.16212.35204

[10] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition,"2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, 2016.