International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 07 | July 2021 www.irjet.net p-ISSN: 2395-0072

ANIMAL SPECIES RECOGNITION SYSTEM USING DEEP LEARNING

DEEPTHI K[1], VIMALA N[2], YUGALAKSHMI K[3], S VINITHA[4], MAMATHA V[5]

[1,2,3,4]Student Dept. of Computer Science and Engineering BIT, Bangalore
[5]Professor, Dept. of Computer Science and Engineering BIT, Bangalore

---

Abstract - *Animals watching is a common hobby but to identify their species requires the assistance of Animal books. To provide Animal watchers a handy tool to admire the beauty of Animals, we developed a deep learning platform to assist users in recognizing species of Animals endemic to using app named the Imagenet of Animals (IoA). Animal images were learned by a convolutional neural network (CNN) to localize prominent features in the images. First, we established and generated a bounded region of interest to the shapes and colors of the object granularities and subsequently balanced the distribution of Animals species. Then, a skip connection method was used to linearly combine the outputs of the previous and current layers to improve feature extraction. Finally, we applied the SoftMax function to obtain a probability distribution of Animals features. The learned parameters of Animals features were used to identify pictures uploaded by mobile users. The proposed CNN model with skip connections achieved higher accuracy of 99.00 % compared with the 93.98% from a CNN and 89.00% from the SVM for the training images. As for the test dataset, the average sensitivity, specificity, and accuracy were 93.79%, 96.11%, and 95.37%, respectively.*

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

Over the last few decades, pattern classification has become one of the most important fields of artificial intelligence because it constitutes an essential component in many different real-world applications.

Artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals. In computer science AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals. Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving".

The images are fed as input to the Animals Recognizer to learn and identify the Animals but the deep learning generally requires a lot of training data and the training process is very slow and complex. Hence we use the GoogLeNet inception v4 network to extract the Animal features. Then the animals are classified based on the extracted features by measuring the similarity between the Animals and its template. Experiments show that our system can identify the animal species accurately after the system learns about ten samples of a particular animal. The self-learning system has a wide range of applicability and flexibility because of the incremental frame based on deep learning.

AI is a study of how human brain think, learn, decide and work, when it tries to solve problems. And finally, this study outputs intelligent software systems. The aim of AI is to improve computer functions which are related to human knowledge, for example, reasoning, learning, and problem-solving.

1.1 Aim and Objective

The project is aimed to recognize animal species by using deep CNN. Since there are large number of different animals manually identifying them can be a difficult task. In this phase of work, limitation by using data mining is researched and a new method is proposed which employs CNN called Lightweight Machine Convolution Network for Animal Recognition. It will have both good accuracy and minimal computational cost.

1.2 Proposed System

The proposed system is developed on CNN which helps in identifying animal feature and using the ImageNet inception v4 network, TensorFlow helps to achieve our recognizer to identify any species.

2. LITERATURE SURVEY

Yao, S. et al. [1] proposes fine-grained visual categorization, which targets to classify the objects belonging to the same species. This novel description only required the original image as input, but could automatically generate visual descriptions discriminative enough for fine-grained visual categorization. The major drawback of fine-grained visual categorization is it is computationally expensive and not suitable for large-scale image. Xie. Et al. [2] proposes that instance search should not return only near duplicate images, but also fine-grained results, which is usually the actual intention of a user. It introduces a baseline system

using fine-grained classification scores by constructing large scale database where the reference images are compressed at constant bit rate levels by JPEG encoders with different optimization methods. To distinguish subtle differences, the comparison method is utilized to rank them in subjective experiments. The major drawback of fine-grained results is duplication occurs while classifying the objects belonging to the same species. The next paper is about template matching [3] algorithm used for identifying small parts of an image which should match the template image. The set of interested objects in the image are identified and information about the location of object in the image is provided. So it was time consuming. Fang, Y., et al. [4] proposes an approach for moving animal detection by taking benefit of global patterns of pixel motion. This paper used the segmented regions, another threshold was used to filter out negative candidates, which could belong to the background. The main drawback is the complexity was high. J. Tanha, et al. [5] proposes Multiclass semi-supervised learning algorithm that uses a base classifier in combination with a similarity function applied to all data to find a classifier that maximizes the margin and consistency over all data and used labelling methods. For each and every feature was labelled. The main drawback was labelling was not easy and time consuming. Tu, S. Yin et al. [6] designed a DCNN acceleration architecture called deep neural architecture (DNA), with reconfigurable computation patterns which comprises of a datareuse pattern and a convolution mapping method for different models. The major drawback is its time consuming since pattern has to been compared to every part of image. H. Nguyen et al. [7] proposes a framework to build automated animal recognition in the wild, aiming at an automated wildlife monitoring system. In particular, a single-labelled datasets was used and the state-of-the-art deep convolutional neural network architectures, to train a computational system capable of filtering animal images and identifying species automatically. The major drawback of datasets which is obtained from wildlife spotter project is not applicable for hybrid animal since the datasets are not updated. J Deng et al. [8] offers a detailed analysis of ImageNet in its current state: 12 subtrees with 5247 synsets and 3.2 million images in total . It describes a new database called "ImageNet", a large-scale ontology of images built upon the backbone of the worldnet structure. ImageNet aims to populate the majority of the 80,000 synsets of worldnet with an average of 500-1000 clean and full resolution images. Akshay kapoor et al. [9] proposes the different architecture schemes and the variants proposed in GoogLeNet and inception networks. These variants are analyzed in terms of the computation efficiency and the network features and performances are juxtaposed on ImageNet dataset and critical review on inception networks is provided.

3. METHODOLOGY

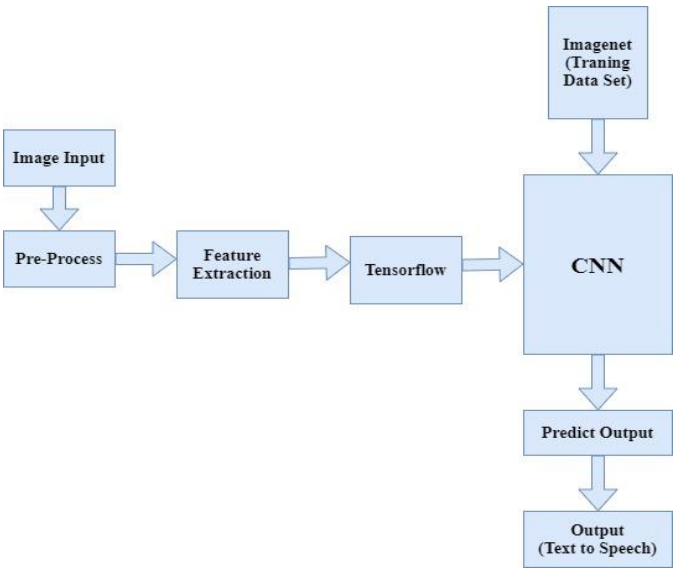

Fig -1: System Architecture of Animal species recognition system

At the beginning user has to upload image i.e., any animal image including land animals, aquatic animals and birds. Pre-process 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 deep artificial neural networks that are used primarily to classify images, cluster them by similarity and perform Animal recognition within scenes. 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 illustrates CNN architecture which employs 3 layers, two consecutive convolutions, pooling layer and a fully connected layer. Convolution layers are the layers where filters are applied to the original image, or to other feature maps in a deep CNN. Convolution layer takes bunch of filters which will be applied to the input image and create different activation features. Pooling layers are similar to convolution

layers, but they perform a specific function such as max pooling, which takes the maximum value in a certain filter region, or average pooling, which takes the average value in a filter region. These are typically used to reduce the dimensionality of the network. In the above figure max pooling is applied. Fully connected layers are placed before the classification output of a CNN and are used to flatten the results before classification.

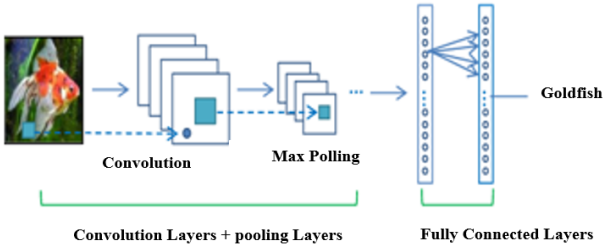

Fig -2: Working of CNN

4.2 Feature extraction based on DeepLearningLog4j

Feature is extracted for the animal image which is pre-processed using Deep Learning Log4J algorithm and helps in classification of the image based on the extracted features. It can extract more discriminable and powerful features by deep neural network, and the recognition system based on these features can learn faster and have a higher recognition rate. The recognition accuracy and the speed of feature extraction when the framework of the deep neural network is selected. GoogLeNet inception v4 are used which have a good performance in Animals recognition, to extract image features, and do the comparative experiment, these networks were pre-trained on ImageNet.

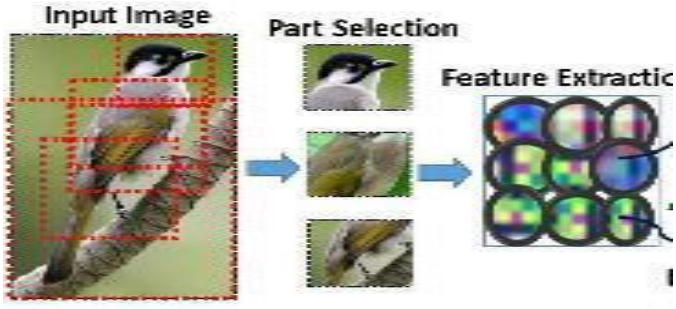

Fig-3 Feature Extraction

4.3 TensorFlow

TensorFlow is a free and open-source software library for data flow and differentiable programming across a range of tasks. It is an open source artificial intelligence library, using data flow graphs to build models. It allows developers to create large-scale neural networks with many layers. TensorFlow is mainly used for: Classification, Perception, Understanding, Discovering, Prediction and Creation. This module will be converting the pre-processed animal image into a graph values so that it can used to identify the animal species.

4.4 ImageNet Training Dataset

The ImageNet Large Scale Visual Recognition Challenge is a benchmark in object classification and detection, with millions of images and hundreds of object classes a large-scale visual recognition challenge. The winner GoogleNet increased the mean average precision of object detection to 0.439329, and reduced classification error to 0.06656, the best result to date. Its network applied more than 30 layers. That performance of convolutional neural networks on the ImageNet tests was close to that of humans. Since there are very large species of animal, we obtain animal images as training datasets to our system with the help GoogleNet inception.

4.5 Text to Speech Converter

Text-to-speech (TTS) is a type of assistive technology that reads digital text aloud. It's sometimes called "read aloud" technology. With a click of a button or the touch of a finger, TTS can take words on a computer or other digital device and convert them into audio. Text-to-voice converter is a additional feature where the output of the recognized animal image is converted to human-like voice.

5. RESULTS AND DISCUSSON

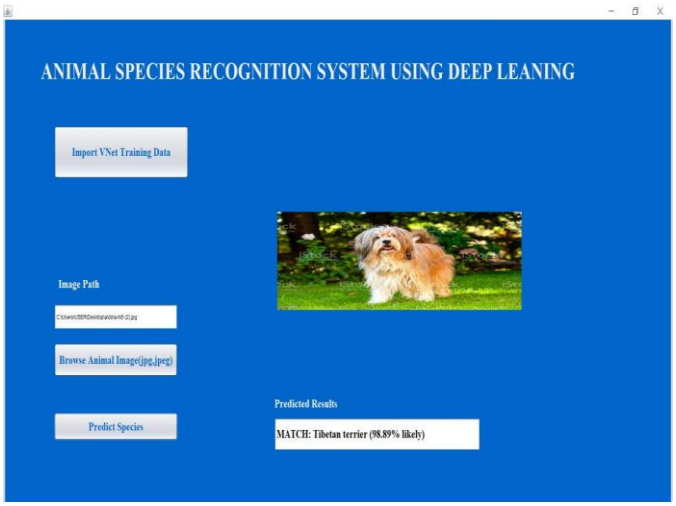

Fig-4 Output Generated for Land Animal

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.

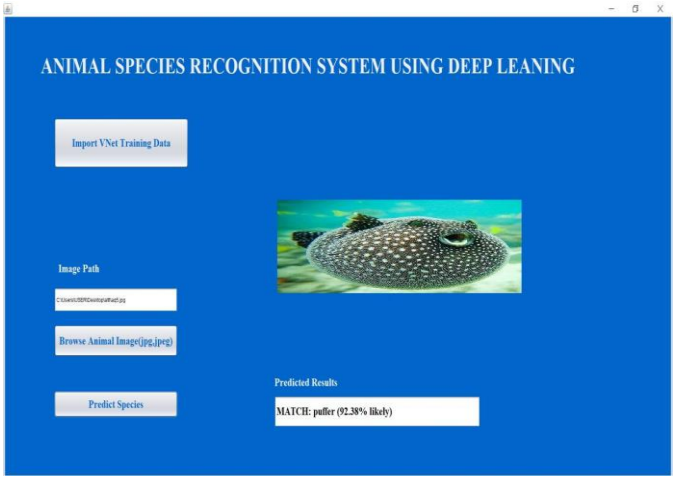

Fig-5 Result for an aquatic animal

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.

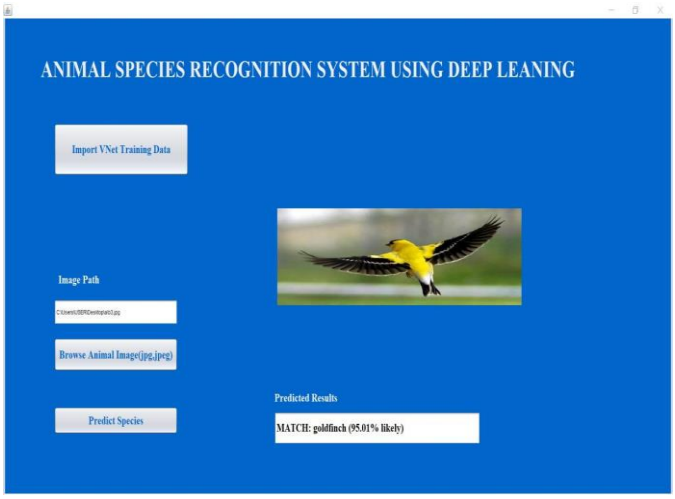

Fig-6 Output generated for Bird

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-to-Fine 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. 17

[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 Semi-supervised 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.

[11] S. U. Sharma and D. J. Shah, "A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique," in IEEE Access, vol. 5, pp. 347-358, 2017

[12] M. Jingyi, T. Zhang, J. Guodong, Y. Wenjun and Y. Bin, "Ground-Based Cloud Image Recognition System Based on Multi-CNN and Feature Screening and Fusion," in IEEE Access, vol. 8, pp. 173949-173960, 2020, doi10.1109/ACCESS.2020.3026364.

[13] F. Garcia, J. Cervantes, A. López and M. Alvarado, "Fruit Classification by Extracting Color Chromaticity, Shape and Texture Features: Towards an Application for Supermarkets," in IEEE Latin America Transactions, vol. 14, no. 7, pp. 3434-3443, July 2016

[14] Verma G.K., Gupta P, "Wild Animal Detection Using Deep Convolutional Neural Network" Proceedings of 2nd International Conference on Computer Vision & Image Processing, 2018, Volume 704

[15] L. Yuan, Z. Qu, Y. Zhao, H. Zhang and Q. Nian, "A convolutional neural network based on TensorFlow for face recognition," 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, 2017, pp. 525-529, doi: 10.1109/IAEAC.2017.8054070.