IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

IMAGE BASED BIRD SPECIES IDENTIFICATION USING DEEP LEARNING

[1]A V Siva Krishna Reddy, [2]Dr. M A Srinuvasu,[3]K Manibabu, [4]Ch B V Sai Krishna, [5]D Jhansi

[1,3,4,5]B.Tech Student, [2]Associate Professor,

Department of Computer Science & Engineering, Raghu Institute of Technology, Visakhapatnam, India

Abstract: Nowadays, it has become a hobby of watching birds for human beings. Due to the different features of the bird with their color, size, and viewing angle of humans are unable to identify and classify the bird species. In this paper, we present a new image classification method for bird species identification. we are extracting the features of a bird and with novel preprocessing and data augmentation methods, we train a convolutional neural network on the biggest publicly available dataset. Our network architecture achieves a mean average precision score of 0.95 when predicting the main species of each image file and scores 0.96, 0.99 chances of bird species from the dataset.

***Index Terms* - Deep learning, Bird Species, Convolutional Neural Network, Image classification, ResNet152.**

I. INTRODUCTION

Many people are fascinated by nature observation and study, particularly birdwatching. Birdwatching serves to protect the environment by observing bird activity and migration patterns. Due to the similarity of the birds' forms/image background and the lack of experience in this sector among bird watchers, recognizing birds based on photographs remains tough [1].

The bird identification can be performed through images, audio, and video. By collecting the acoustic signal of birds, an audio processing technique allows for identification. However, processing such information becomes more difficult as a result of mixed sounds in the surroundings, such as insects, real-world items, and so on. Humans prefer visuals to music or video in most cases. Bird species identification is a difficult undertaking for both humans and computing methods that perform such tasks automatically [2].

Ornithologists have encountered a variety of obstacles in the identification of bird species for decades. Ornithologists examine the traits and attributes of birds [3-4], distinguishing them based on their habitat, ecological impact, biology, and other factors. The ornithologists use Linnaeus' classification system to classify the birds: Phylum, Kingdom, Order, Class, Family, and Species.

The rest of the paper was organized as follows section 1 is the introduction, section 2 describes about the Background work followed by proposed methodology in section 3. Experimental results as discussed in section 4 and finally conclusion.

II. BACKGROUND WORK

In deep learning, a convolutional neural network (CNN) is a class of deep neural networks used for analyzing visual images. It consists of an input layer and output layer as well as multiple hidden layers. Every layer is a group of neurons and each layer is fully connected to all neurons of its previous layer. The output layer is responsible for the prediction of output. The convolutional layer takes an image as input and produces a set of feature maps as output [5]. The input image can contain multiple\ channels such as color, wings, eyes, and the beak of birds which means that the convolutional layer performs a mapping from 3D volume to another 3D volume. 3D volumes considered are width, height, depth.

The CNN has two components:

1. Feature extraction part: features are detected when the network performs a series of the convolutional and pooling operation.
2. Classification part: extracted features are given to a fully connected layer which acts as a classifier

CNN consists of four layers: convolutional layer, activation layer, pooling layer, and fully connected. The convolutional layer allows extracting visual features from an image in small amounts. Pooling is used to reduce the number of neurons from the previous convolutional layer but maintaining the important information. The activation layer passes a value through a function that compresses values into range. A fully connected layer connects a neuron form one layer to every neuron in another layer. As CNN classifies each neuron in-depth, so it provides more accuracy [5-6].

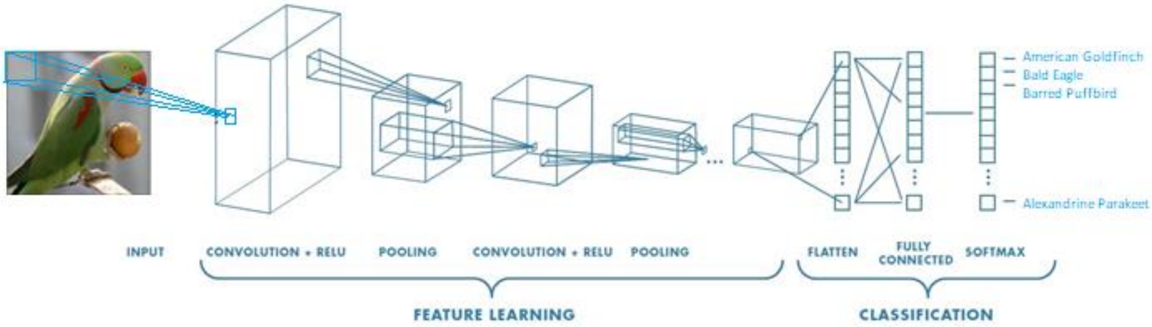

Figure 1: Working of Convolutional Neural Network

The most popular neural network model being used for image classification problems is Convolutional Neural Networks. The CNN model conjuration for bird species identification utilized a stack of convolution layers comprising an input layer, two fully connected layers, and one final output softmax layer [5-6]. Convolutional layers apply a convolution operation to the input and this passes the resulting information on to the next layer. The most popular neural network model being used for image classification problems is Convolutional Neural Networks. The CNN model conjuration for bird species identification utilized a stack of convolution layers comprising an input layer, two fully connected layers, and one final output softmax layer [5-6]. Convolutional layers apply a convolution operation to the input and this passes the resulting information on to the next layer.

Residual Network

Residual Network is shortly called ResNet. ResNet is a specific type of neural network that was familiarized in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun in their article 'Deep Residual Learning for Image Recognition' [7]. Mostly in order to solve a complex problem, we stack some additional layers in the Deep Neural Networks which results in improved accuracy and performance. The intuition behind adding more layers is that these layers progressively learn more complex features.

There are two types of residual connections:

The identity shortcuts (x) can be directly used when the input and output are of the same dimensions. When the residual block function input and output dimensions are not same

$$y = f(x, \{W_i\} + x \qquad (1)$$

When the dimension's change,

A. The shortcut still performs identity mapping, with extra zero entries padded with the increased dimension.
B. The projection shortcut is used to match the dimension (done by 1*1 conv) using the following formula

$$y = f(x, \{W_i\} + W_s x \qquad (2)$$

When the residual block function input and output dimensions are not same

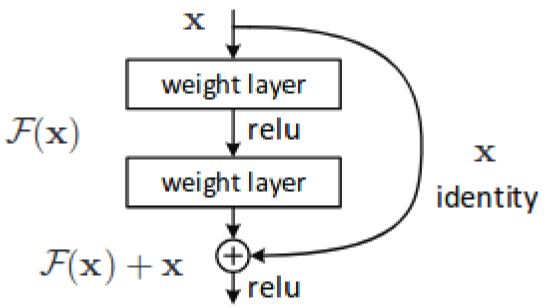

Figure 2: Building Block of Residual Learning

III. Proposed Methodology

The most popular neural network model being used for image classification problems is Convolutional Neural Networks. The CNN model conjuration for bird species identification utilized a stack of convolution layers comprising an input layer, two fully connected layers, and one final output softmax layer [5-6]. Convolutional layers apply a convolution operation to the input and this passes the resulting information on to the next layer.

For the bird image recognition, we required a solid dataset on which the identification system can be trained, tested, and validated. The dataset contains 275 different categories of bird species. The dataset [8] was split into a training set, testing set, and validation set. It is very important to keep the testing set completely separate from the training set since it needed to be sure that the classification model will perform well in real-world scenarios. Then it can be used for training the classification model. More than 70% of data was allocated to the training set and the rest of the data was allocated to the testing set and validation set. The training set and validation set are randomly selected from the dataset for the fine-tuning process.

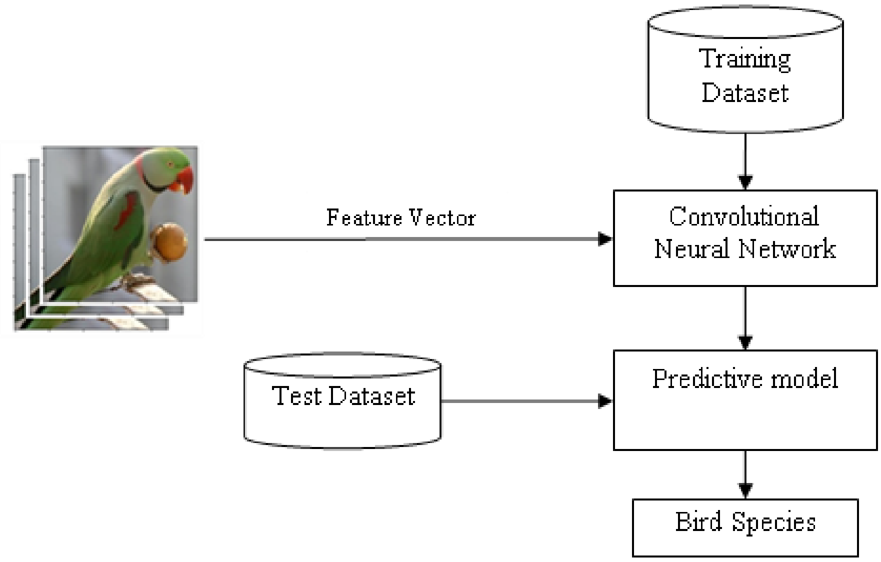

Figure 3: Flow diagram for identification of Bird Species

IV. RESULTS AND DISCUSSION

In this section, the experimental analysis was done on 275 bird species image data set which was openly available at the kaggle repository. We found that the training images count is 39364, the test image count is 1375 and the validation image count is 1375. This experiment was prepared on windows operating system and python 3 versions by using libraries like pandas, NumPy, TensorFlow [9], Keras, sklearn and for plotting visualizations mathplotlib. ResNet 152 v2 model was used to train the image dataset. We observed that test accuracy with 95.71% and loss function with 15.06%. The graph plotted for accuracy and loss function as shown in below figures.

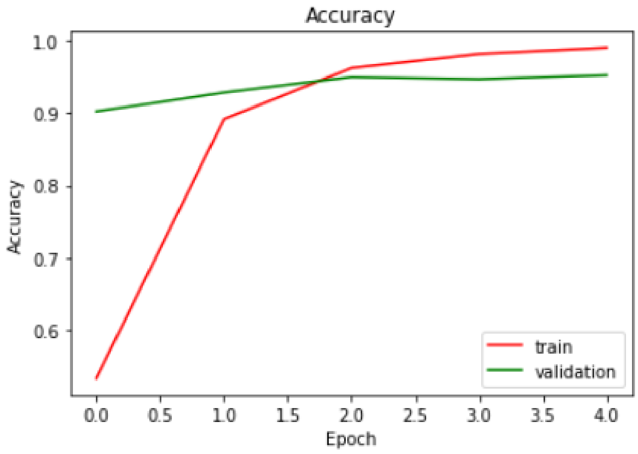

Figure 4: Test Accuracy Curve

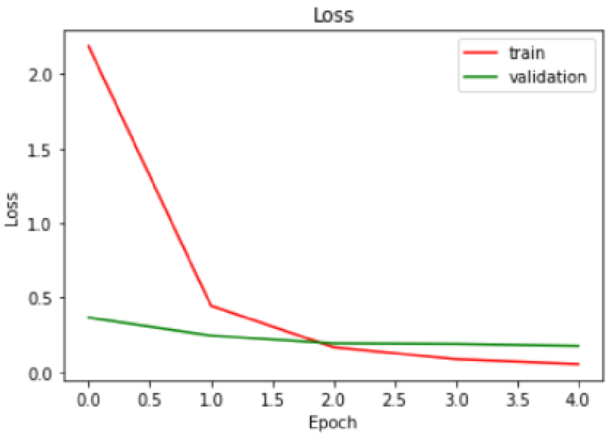

Figure 4: Loss Function Curve

On a pre-trained model of the ResNet, we have given a few bird images and it classified the bird species as 96.89 % chances are there that the bird is American Goldfinch, 100.0 % chances are there that the bird Is Bald Eagle, and 99.62 % chances are there that the bird is Barred Puffbird, and 99.96 % chances are there that the bird is Alexandrine Parakeet. The visualization graph was shown in the below figure.

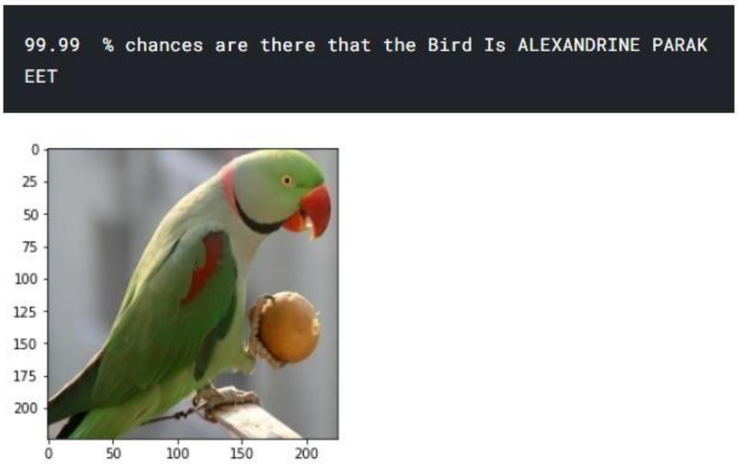

Figure 5: Image have a chance of Alexandrine Parakeet

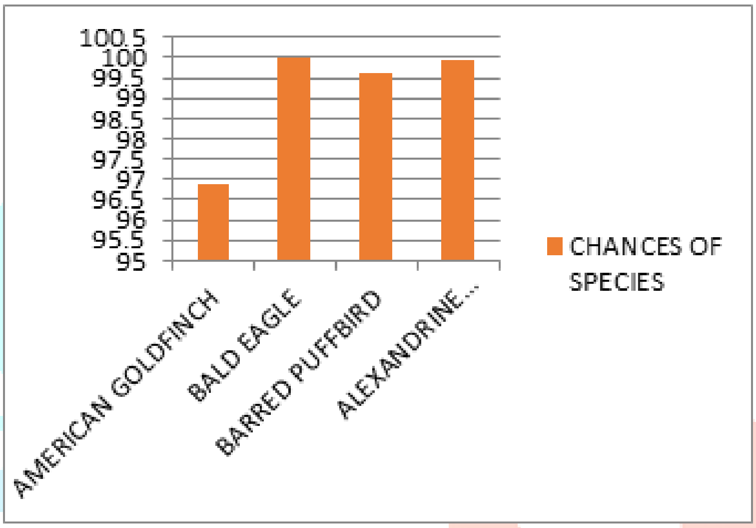

Figure 6: Comparative graph for chances of bird species

V. CONCLUSION

The present study investigated a method to identify the bird species using Deep learning algorithm (Unsupervised Learning) on the dataset (bird species-275) for classification of bird image. It consists of 275 categories or 39,364 photos. The proposed system works on the principle based on detection of a part and extracting CNN features from multiple convolutional layers. These features are aggregated and then given to the classifier for classification purpose. On basis of the results which has been produced, the system has provided the 95.71% accuracy in prediction of finding bird species.

REFERENCES

[1] P. Gavali and J. S. Banu, "Bird Species Identification using Deep Learning on GPU platform," *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, 2020, pp. 1-6, doi: 10.1109/ic-ETITE47903.2020.85.

[2] Tóth, B.P. and Czeba, B., “Convolutional Neural Networks for Large-Scale Bird Song Classification in Noisy Environment.” *2016, September, In CLEF (Working Notes)*,2016, pp. 560-568.

[3] Fagerlund, S., “Bird species recognition using support vector machines.” *EURASIP Journal on Applied SignalProcessing*, 2007(1), pp.64-64.

[4] Pradelle, B., Meister, B., Baskaran, M., Springer, J. and Lethin, R., 2017, November. “Polyhedral Optimization of TensorFlow Computation Graphs.” In 6th Workshop on Extreme-scale Programming Tools (ESPT-2017) at *The International Conference for High Performance Computing*, Networking, Storage and Analysis (SC17).

[5] Gavali, Pralhad, and J. Saira Banu. "Deep Convolutional Neural Network for Image Classification on CUDA Platform." In Deep Learning and Parallel Computing Environment for Bioengineering Systems, pp. 99-122. Academic Press, 2019

[6] Cireşan, D., Meier, U. and Schmidhuber, J., 2012. Multi-column deep neural networks for image classification. arXiv preprint arXiv:1202.2745.

[7] Kaiming, H.,Xiangyu,Z.,Shaoqing, R. and Jian.S., 2015.Deep Residual Learning for Image Recognition. arXiv preprint arXiv:1512.03385

[8] https://www.kaggle.com/gpiosenka/100-bird-species

[9] https://www.tensorflow.org/datasets/catalog/overview.