



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** X **Month of publication:** October 2022

DOI: <https://doi.org/10.22214/ijraset.2022.47039>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

Bird Species Identifier using Convolutional Neural Network

Ashmita Jange¹, Deepika Kattimani², Prof. Jyothi Patil³

^{1,2}B.E student, ³Associate Professor, Department of Computer Science and Engineering, P.D.A College of Engineering, Kalaburagi, Karnataka, India

Abstract: In our world, there are above 9000 bird species. Some bird species are being found rarely and if found also prediction becomes very difficult. In order to overcome this problem, we have an effective and simple way to recognize these bird species based on their features. Also, the human ability to recognize the birds through the images is more understandable than audio recognition. So, we have used Convolutional Neural Networks (CNN). CNN's are the strong assemblage of machine learning which have proven efficient in image processing. In this paper, a CNN system classifying bird species is presented and uses the Caltech-UCSD Birds 200 [CUB-200-2011] dataset for training as well as testing purpose. By establishing this dataset and using the algorithm of similarity comparison, this system is proved to achieve good results in practice. By using this method, everyone can easily identify the name of the particular bird which they want to know.

Keywords: Deep Learning, CNN Model, Classification and Prediction, TensorFlow, Keras

I. INTRODUCTION

Bird behaviour and populace patterns have become a significant issue now a days. Birds help us to recognize different life forms on the earth effectively as they react rapidly to ecological changes. Be that as it may, assembling and gathering data about bird species requires immense human exertion just as it turns into an extremely costly technique. In such a case, a solid framework that will give enormous scale preparation of data about birds and will fill in as a significant apparatus for scientists, legislative offices, and so forth is required. In this way, bird species distinguishing proof assumes a significant job in recognizing that a specific picture of birds has a place with which categories. Bird species identification means predicting the bird species belongs to which category by using an image. The recognition of bird species can be possible through a picture, audio or video. An audio processing method makes it conceivable to recognize by catching the sound sign of different birds. Be that as it may, because of the blended sounds in condition, for example, creepy crawlies, objects from the real world, and so forth handling of such data turns out to be progressively convoluted. Normally, people discover images more effectively than sounds or recordings. So, an approach to classify birds using an image over audio or video is preferred. Bird species identification is a challenging task to humans as well as to computational procedures that carry out such a task in an automated fashion.

As image-based classification systems are improving the task of classifying, objects are moving into datasets with far more categories such as Caltech-UCSD. Recent work has seen much success in this area. Caltech UCSD Birds 200(CUB-200-2011) is a wellknown dataset for bird images with photos of 200 categories. The dataset contains birds that are mostly found in Northern America. Caltech-UCSD Birds 200 consists of 11,788 images and annotations like 15 Part Locations, 312 Binary Attributes, 1 Bounding Box. In this project, rather than recognizing an oversized number of disparate categories, the matter of recognizing an oversized number of classes within one category is investigated – that of birds. Classifying birds pose an additional challenge over categories, as a result 2 of the massive similarity between classes. additionally, birds are non-rigid objects which will deform in many ways and consequently there's also an oversized variation within classes. Previous work on bird classification has taken care of a little number of classes, or through voice.

II. PROBLEM DEFINITION

Manual identification of bird species is very tedious task as well as very unreliable as his/her knowledge may not be in-depth and limited to the local bird species. This process is a lot of time-consuming and it may contain some errors. There are lots of books that have been published for the process of helping a human incorrectly identifying bird species. The current bird species identification process involved using the bird audio which is recorded and fed into the system. Nevertheless, it requires the hundreds of hours to carefully analyzed and classify the species. Due to such a process, large scale bird identification is almost an impossible task. So, to automate the process is a more practical approach.

III. LITERATURE SURVEY

- A. Juha Niemi to detect an image in two ways i.e., based on feature extraction and signal classification. They did an experimental analysis for datasets consisting of different images. But their work didn't consider the background species. In Order to identify the background species larger volumes of training data are required, which may not be available.
- B. Juha T Tanttu et al (2018), proposed a Convolutional neural network trained with John Martinsson et al (2017), presented the CNN algorithm and deep residual neural networks deep learning algorithms for image classification. It also proposed a data augmentation method in which images are converted and rotated in accordance with the desired color. The final identification is based on a fusion of parameters provided by the radar and predictions of the image classifier.
- C. Li Jian, Zhang Lei et al (2014) proposed an effective automatic bird species identification based on the analysis of image features. Used the database of standard images and the algorithm of similarity comparisons.
- D. Madhuri A. Tayal, Atharva Magrulkar et al (2018) , developed a software application that is used to simplify the bird identification process. This bird identification software takes an image as an input and gives the identity of the bird as an output. The technology used is transfer learning and MATLAB for the identification process.
- E. Andreia Marini, Jacques Facon et al (2013) , proposed a novel approach based on color features extracted from unconstrained images, applying a color segmentation algorithm in an attempt to eliminate background elements and to delimit candidate regions where the bird may be present within the image. Aggregation processing was employed to reduce the number of intervals of the histograms to a fixed number of bins. In this paper, the authors experimented with the CUB-200 dataset and results show that this technique is more accurate.

IV. METHODOLOGY

A. Proposed System

Convolution neural network algorithm is a multilayer perceptron that is the special design for the identification of two-dimensional image information.

It has four layers: an input layer, a convolution layer, a sample layer, and an output layer. In a deep network architecture, the convolution layer and sample layer may have multiple.

CNN is not as restricted as the Boltzmann machine, it needs to be before and after the layer of neurons in the adjacent layer for all connections, convolution neural network algorithms, each neuron doesn't need to experience the global image, just feel the local region of the image.

In addition, each neuron parameter is set to the same, namely, the sharing of weights, namely each neuron with the same convolution kernels to the deconvolution image.

B. Convolution Layer:

The convolutional layer is the core constructing block of a CNN. The convolution layer comprises a set of independent feature detectors. Each Feature map is independently convolved with the images.

C. Pooling Layer:

The pooling layer feature is to progressively reduce the spatial size of the illustration to reduce the wide variety of parameters and computation in the network. The pooling layer operates on each function map independently.

The approaches used in pooling are:

- 1) Max Pooling
- 2) Mean Pooling
- 3) Sum Pooling

D. Fully Connected Layer:

Neurons in the fully connected layer have full connections to all activations inside the preceding layer. In this, the output obtained from max pooling is converted to a onedimensional array and that should be the input layer and the process continues the same as the ANN model.

E. Architecture

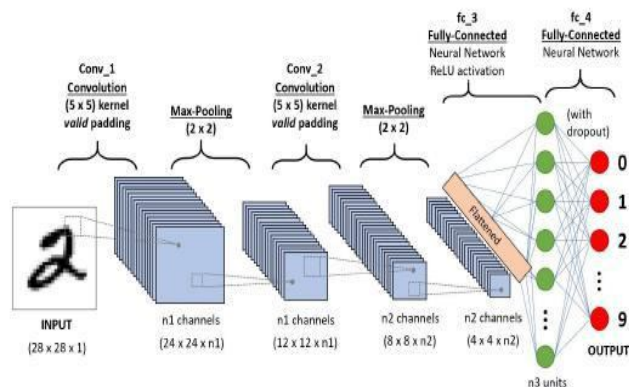


Fig 1. A CNN Sequence

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The preprocessing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

A ConvNet is able to successfully capture the Spatial and Temporal dependencies in an image through the application of relevant filters. The architecture performs a better fitting to the image dataset due to the reduction in the number of parameters involved and reusability of weights. In other words, the network can be trained to understand the sophistication of the image better.

F. Convolution Layer

Filters (Convolution Kernels)

A filter (or kernel) is an integral component of the layered architecture. Generally, it refers to an operator applied to the entirety of the image such that it transforms the information encoded in the pixels. In practice, however, a kernel is a smaller-sized matrix in comparison to the input dimensions of the image, that consists of real valued entries.

The real values of the kernel matrix change with each learning iteration over the training set, indicating that the network is learning to identify which regions are of significance for extracting features from the data.

G. Pooling Layer

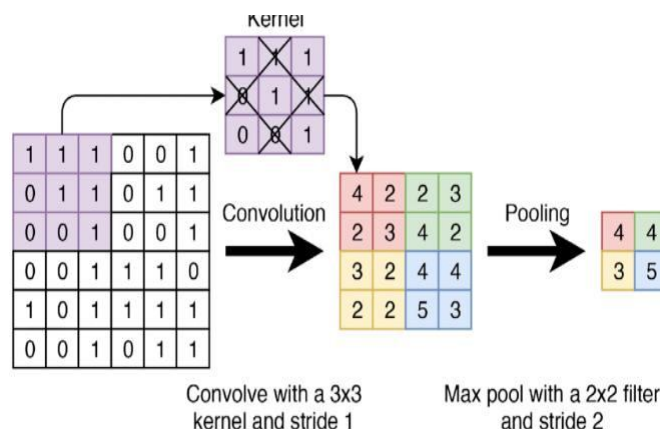


Fig 2. Performing Pooling operation

The Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.

V. PROJECT DESIGN

A. System Implementation

This subsection explains the implementation of the proposed deep learning platform to identify the species of birds. The image of the bird is given as the input to the system. The software system has a trained model with a set of data. The image of the bird gets converted into gray scale and then into matrix format. Various alignments from that image will be taken into consideration and each alignment will be given to the CNN for feature extraction. These extracted features are given to the CNN with trained model and then the resulting values gets compared. According to the compared values the classifier classifies the image into different categories. Then the predictive model predicts the species of that particular bird which is considered to be as the final result. The output layer of the network provides parts of the input image containing the bird.

B. Architecture Diagram

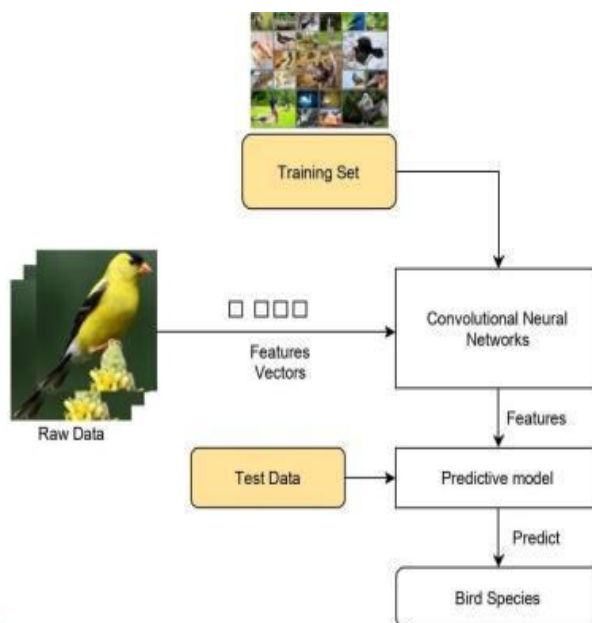


Fig 3. Architecture Diagram

C. Modules Dataset

A dataset is a collection of data. For performing action related to birds a dataset named Caltech-UCSD Birds 200 (CUB-200-2011) is used.

- 1) *Feature Extraction*: This subsection describes the feature extraction process in identification of the bird species using bird images. Primary task is to extract features from raw input images, when extracting relevant and descriptive information for fine-grained object recognition. To extract and learn features, CNNs apply a number of filters to the raw pixel data of an image
- 2) *Predictive Model*: In the feature extraction process, the feature vectors extracted from the raw data (Image of a bird) is given to the CNN which is trained with the training dataset. The extracted features are then passed to the predictive model which compares the features with the test data. Then the model predicts the species of that particular bird in the image

D. Hardware Requirements

- 1) *Processor*: Intel Core I3 and above
- 2) *Processor Speed* : 1.0GHZ or above
- 3) *RAM*: 4 GB RAM or above
- 4) *Hard Disk*: 500 GB hard disk or Above

E. Software Requirements

Operating System: Windows 7/10 or above •Front End : Python •Back End : SQLITE3

VI. IMPLEMENTATION

Training:

Testing:

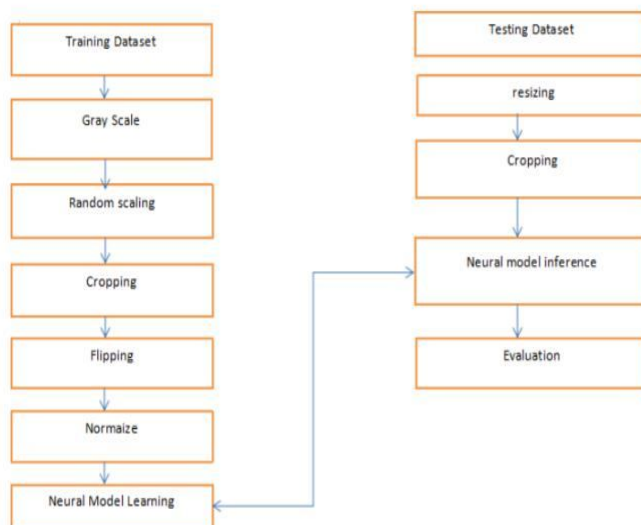


Fig 4.Training and Testing Phase

Deep learning operational working is similar to the human brain. It learns from the data and makes inferences on the data feature based on trained data. Therefore to develop a good neural model having a diverse as well as a huge dataset is necessary. For this purpose, In our research, we are using the data augmentation technique which helps to increase the number of training samples per class and reduce the effect of class imbalance. Relevant image augmentation techniques are chosen so that the neural model can learn from the diverse dataset. Those techniques are Gaussian Noise, Gaussian Blur, Flip, Contrast, Hue, Add (add some values to each channel of the pixel), multiply (multiply some values to each channel of the pixel), Sharp, Affine transform. The large dataset also help to avoid the problem of overfitting which happens quite often in deep network learning. As the image dataset requires higher computational capability as compared to the text-based dataset. In our research, we try to reduce this computational requirement by removing the unwanted part from the image so that the neural model needs to deal with a lesser amount of pixel in the image for processing. So to eliminate background elements or regions and extract features from the only body of the birds, pretrained object detection deep nets are used. For this model, we are using Mask R-CNN to localize birds in each image in training phase as well as in the inference phase. We have used the pre-trained weights of Mask R-CNN, trained on the COCO dataset [6] which contains 1.5 million object instances with 80 object categories(including birds)

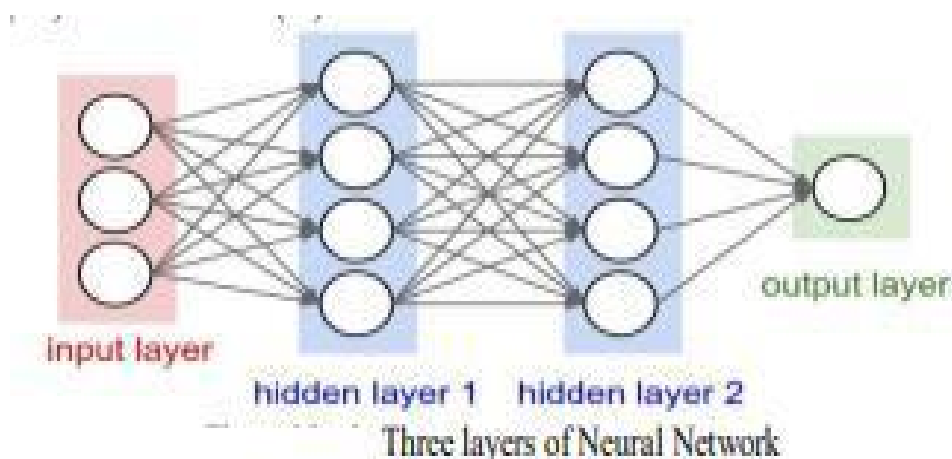


Fig.5 Three Layer of Neural Network

VII. EXPERIMENTAL RESULTS

The evaluation of the proposed approach for bird species classification by considering color features and parameters such as size, shape, etc. of the bird on the Caltech-UCSD Birds 200 (CUB-200-2011) dataset. This is an image dataset annotated with 200 bird species which includes 11,788 annotated images of birds where each image is annotated with a rough segmentation, a bounding box, and binary attribute annotations. In this the training of dataset is done by using Google-Collab, which is a platform to train dataset by uploading the images from your local machine or from the Google drive.

After training labeled dataset is ready for classifiers for image processing. There are probably average 200 sample images per species are included in dataset of 5 species which are directly captured in their natural habitat hence also include the environmental parameters in picture such as grass, trees and other factors. Here bird can identify in their any type of position as main focus is on the size, shape and color parameter. Firstly these factors are considered for segmentation where RGB and gray scale methods are used for histogram. That is the image converted into number of pixels by using gray scale method, where value for each pixel is created and value based nodes are formed which also referred as neurons. These neurons relatively defined the structure of matched pixels is simply like graph of connected nodes.

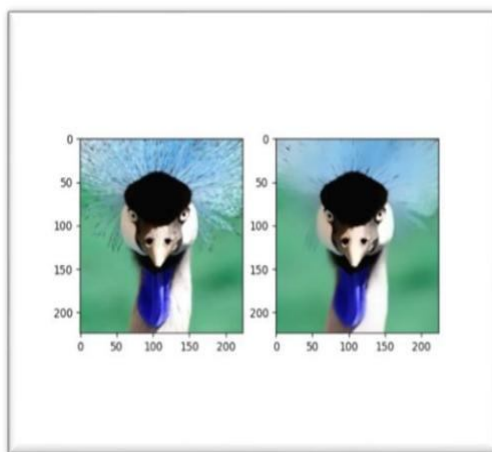


Fig.6 Compare Dataset Screenshot.

The table no.1 shows the scoresheet based on the result generated by the system. After analysis of these result it has observe that, the species those are having the highest score has been predicted as a required species.

VIII. CONCLUSION

The main idea behind developing the identification website is to build awareness regarding bird-watching, bird and their identification, especially birds found in India. It also caters to the need of simplifying the bird identification process and thus making bird-watching easier. The technology used in the experimental setup is Convolutional Neural Networks (CNN). It uses feature extraction for image recognition. The method used is good enough to extract features and classify images. The main purpose of the project is to identify the bird species from an image given as input by the user. We used CNN because it is suitable for implementing advanced algorithms and gives good numerical precision accuracy. It is also general-purpose and scientific. We achieved an accuracy of 85%-90%. We believe this project extends a great deal of scope as the purpose meets. In wildlife research and monitoring, this concept can be implemented in-camera traps to maintain the record of wildlife movement in specific habitat and behaviour of any species.

IX. FUTURE ENHANCEMENT

- 1) Create an android/ios app instead of website which will be more convenient to user.
- 2) System can be implemented using cloud which can store large amount of data for comparison and provide high computing power for processing (in case of Neural Networks).



REFERENCES

- [1] Fagerlund, Seppo. "Bird species recognition using support vector machines." EURASIP Journal on Advances in Signal Processing 2007, no. 1 (2007): 038637.
- [2] Marini, Andrea, Jacques Facon, and Alessandro L. Koerich. "Bird species classification based on color features." In 2013 IEEE International Conference on Systems, Man, and Cybernetics, pp. 4336- 4341. IEEE, 2013.
- [3] Barar, Andrei Petru, Victor Neagoe and Nicu Sebe. "Image Recognition with Deep Learning Techniques." Recent Advances in Image, Audio and Signal Processing: Budapest, Hungary, December 10-2 (2013).
- [4] Qiao, Baowen, Zuofeng Zhou, Hongtao Yang, and Jianzhong Cao. "Bird species recognition based on SVM classifier and decision tree." First International Conference on Electronics Instrumentation & Information Systems (EIIS), pp. 1-4, 2017.
- [5] Branson, Steve, Grant Van Horn, Serge Belongie, and Pietro Perona. "Bird species categorization using pose normalized deep convolutional nets." arXiv preprint arXiv: 1406.2952 (2014).
- [6] Madhuri A. Tayal, Atharva Mangrulkar, Purvashree Waldey and Chitra Dangra. Bird Identification by Image Recognition. Helix Vol. 8(6): 4349- 4352
- [7] Atanbori, John, Wenting Duan, John Murray, Kofi Appiah, and Patrick Dickinson. "Automatic classification of flying bird species using computer vision techniques." Pattern Recognition Letters (2016): 53-62.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)