IJITCE

ISSN 2347- 3657

# International Journal of Information Technology & Computer Engineering


Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 9, Issue 4, 2021

# SPECIES DETECTION USING DEEP LEARNING FOR BIRDS

**Dr.Siddiqui Riyazoddin Alimoddin[1], Md.Abdul Rawoof[2]**
*Professor[1], Asst.Prof[2]*
***Department of cse***
***NAWAB SHAH ALAM KHAN COLLEGE OF ENGINEERING & TECHNOLOGY NEW MALAKPET,***
***HYDERABAD-500024***

## Abstract:

***Some bird species are becoming more uncommon, and even if they are discovered, it may be difficult to classify them. From a human viewpoint, birds may be seen in a variety of sizes, shapes, colours, and angles. Because of this, it is easier to identify birds by looking at their photos, as opposed to listening to their calls. In addition, it is easier for humans to identify birds based on their appearances in photos. As a result, our technique relies on the Caltech-UCSD Birds 200 dataset for both training and testing. To produce an autograph using tensor flow and a deep convolutional neural network (DCNN), a picture is transformed to grey scale using the DCNN technique. The testing dataset is used to compare the various nodes, and a score sheet is generated as a result. The highest score on the scoring sheet may be used to determine which bird species is needed. Analyses of the dataset (Caltech-UCSD Birds 200 [CUB-200-2011]) have shown that the system is 80-90% accurate in identifying birds. The experiment is run on Ubuntu 16.04 using a Tensor flow library on a laptop.***

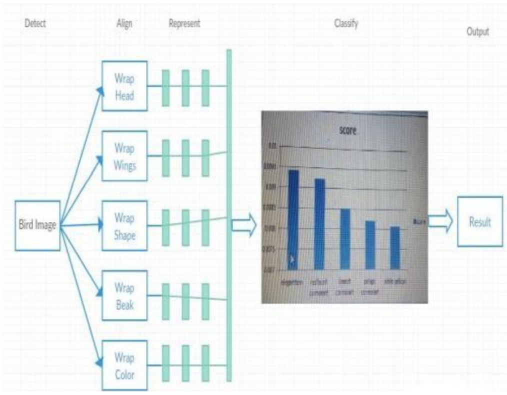
***IndexTerms: Caltech-UCSD; grayscale pixels; Tensorflow Autograph***

## INTRODUCTION

There is a lot of interest in bird behaviour and population trends these days. By promptly responding to environmental changes, birds (e.g. the insects they eat) help humans identify other creatures in the environment [2]. However, acquiring and compiling data on birds necessitates a significant investment of time and resources on the part of humans. Researchers, government entities, and the like will need a dependable system that can handle massive amounts of bird data and function as a useful tool in this situation. As a result, the identification of bird species is critical in determining which specific picture of a bird belongs to which species. Identification of bird species is the process of determining which group a particular bird belongs to based only on visual cues. An picture, audio, or video recording may be used to identify a person. Birds may be identified by their sounds according to an audio processing approach. Insects, things from the actual world, and other ambient noises hinder the processing of this information. Generally, visuals are more powerful than audios or videos when it comes to persuading people. As a result, a strategy that relies on a picture rather than voice or video to categorise birds is the favoured choice. People and computer algorithms alike find it difficult to identify bird species, and this is true for both humans and computers. Bird species identification has been a concern for ornithologists for decades. To be an ornithologist, you have to know all there is to know about birds, including where they live, how they live, how they reproduce, how they effect the ecosystem, and so on. For the most part, ornithologists use Linnaeus' categorization system to determine the kind of bird they are looking at [1]. It is becoming easier to categorise items using image-based classification systems like Caltech-UCSD, which have a wide range of classifications. A lot of progress has been made in this area recently. Bird photographs from the Caltech UCSD Birds 200(CUB-200-2011) collection are widely recognised [4]. More over half of the birds in the sample are located in North America. Annotations such as 15 Part Locations and 312 Binary Attributes are included in the Caltech-UCSD Birds 200 dataset, which comprises of 11,788 pictures. There are many different ways to classify birds, but one method is used in this work to focus on the difficulty of classifying a large number of classes within one category. Because of the striking resemblances among bird species, it is particularly difficult to assign them to a specific classification. There is also a great deal of variety within classes since birds are non-rigid objects that may deform in several ways. A modest number of

ISSN 2347–3657
Volume 9, Issue 4, 2021

classifications or vocalisations have been used in previous bird categorization studies.




## Figure No.1: Process of classification

Detection of a bird from a photograph is shown in Figure 1. The picture is initially uploaded, and then several alignments, such as head, body, colour, beak, and full image, are taken into account. The convocational network used for each alignment is a deep convocational network that extracts features from various levels of the network After then, the image's representation will be taken into account. It will then be used to create a classification result, and the species of bird detected using that classification result. The format of this document is as follows: While identifying a bird visually, there are a number of variables that may be taken into account. Here, the proposed system's development methodology is laid out in further detail. Section IV provides an in-depth look at the system's general operation.

## BACKGROUND

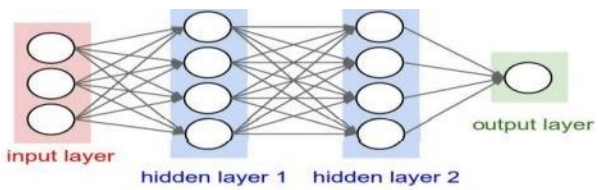
In most cases, a bird's identity may be determined visually or audibly. There are a number of visual elements that make up a bird's appearance. However, the time of year must be taken into account while evaluating the criteria, since the size of a bird's wings fluctuates as it grows. Birds' songs and calls are made up of acoustic components [7]. The distinguishing features of a bird, such as breast patches, wing bars, eye rings, crowns, and eyebrows, are also valuable. A bird's distinctive beak shape is one of its most distinguishing features. Birds are generally identified by their physical traits, such as their form and posture. Experts are usually able to identify a bird based on its silhouette since it is hard to alter. The tail of a bird may also be used to distinguish one species from another. In addition to being long and pointed, the tail might be notched or rounded. Legs may also be used to identify images that are either long or short [10]. A precise result cannot be obtained by just taking into account one parameter. As a result, various variables must be taken into account in order to get the desired result. The size of a bird in a picture is affected by a number of variables, including the resolution, the distance between the birds and the camera, and the lens's focal length. Consequently, a huge number of photos have been observed to discriminate images based on their colour, which is made up of distinct pixels. Researchers have discovered that better picture quality means more accuracy. A CUB 200 dataset with more than 6,000 photos and 200 distinct categories was used to undertake a series of comparisons for the automated bird species identification for bird photographs project [6]. Two colour spaces (RGB and HSV) and a variety of species were examined in this study. The output accuracy ranged from 8.82% to 0.432% when the picture included more than 70% of the pixels [1].

## METHODOLOGY

A variety of approaches have been used in the process of creating the system. These are the names of them: Database (UCSD-Caltech) The system was built using a variety of methods. The following is a list of each of them: Deep Convolutional Neural Network, Unsupervised Learning Algorithm, etc. are some of the tools used. Algorithm: Unsupervised learning algorithms were utilised in this experiment since the input picture definition was unknown. There is no labelling of the data that is presented to an unsupervised learning system; just the input variables (X) are provided. An algorithm discovers intriguing patterns in the data on its own through unsupervised learning. To put it

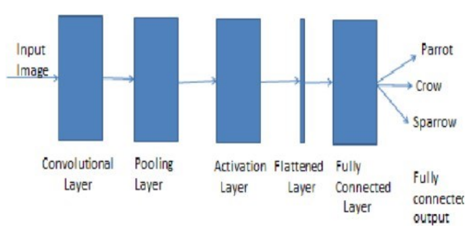
International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 9, Issue 4, 2021

another way, clustering is a method of grouping data into distinct categories. Deep learning techniques have been used to locate a large number of neurons in greater detail. As a picture passes through successive layers of neural networks, deep learning algorithms get a better understanding of it. Neural Networks are used to categorise. Layers of neural networks for feature extraction are shown in Figure 2. Many machine learning techniques are built on top of the neural network foundation. neural networks are made up of a weighted vector and a bias vector (B).




### Figure No. 2: Three layers of Neural Network

Visual image analysis is the primary function of this particular kind of deep neural networks. Multiple hidden layers are included in the system, including an input and output layer. There are groups of neurons in each layer, and each layer is completely linked to the preceding layer's neurons. It is the output layer's job to forecast what will happen. An picture is fed into the convolutional layer, which outputs a collection of feature maps [2]. The convolutional layer performs a mapping from a 3D volume to another 3D volume when the input picture has several channels such as colour, wings, eyes, and beak of birds. A 3D volume's width, height, and depth are examined. There are two main parts to CNN: First, the network executes a series of convolutional and pooling operations in order to identify features. A fully connected layer is used as a classifier for the retrieved characteristics.




### Figure No. 3: Convolutional Neural Network Layers

These four layers are convolutional, activation, pooling and completely linked. A convolutional layer may be used to extract tiny bits of visual information from a picture. Reduce the number of neurons in prior convolutional layers while retaining relevant information by using pooling. The activation layer compresses data into a range by passing them via a function. Every neuron in one layer is coupled to every neuron in another layer in a fully connected layer. In part, this is due to CNN's ability to classify each neuron in great detail. Two methods are used for picture categorization in machine learning: In this case, the grey scale is used. 2) Utilizing the RGB colour space Grayscale is the most common way to represent data. The computer assigns values to each pixel depending on the pixel's value in a grey scale algorithm. To categorise the data, the computer creates an array with all of the pixel values. Tensorflow is a Google-developed open-source software library. It allows programmers to fine-tune the characteristics of each "node," or neuron, in the system to get the most possible performance out of the system. Tensorflow comes with a slew of image classification libraries pre-installed [3]. Using a set of processing nodes, Tensorflow creates an autograph. Operations such as mathematical operations and connections between nodes are represented in the graph by processing nodes. Tensorflow's python language makes it possible for a coder to carry out these tasks. A dataset is a collection of information. We will be using the Caltech-UCSD Birds 200 (CUB-201-011) dataset for doing any bird-related actions. More than twice as many photos are included in each class as in the

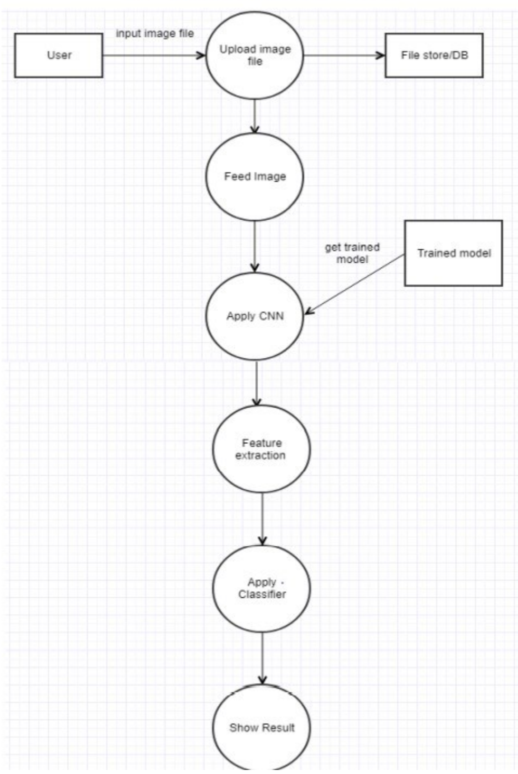
ISSN 2347–3657

Volume 9, Issue 4, 2021

original CUB-200 dataset, and the part locations have been annotated to improve precision [8]. The following is a breakdown of the dataset's information: There are 200 different categories to choose from.

**Figure No. 4: Caltech-UCSD Dataset**

IV. PROPOSED APPROACH




## Figure No.5: Flow of System

The planned system's real flow is shown in the aforementioned picture no. 5. To create such a system, you'll need a dataset that has been used to train the system to recognise images. The two portions of the trained dataset are the results of the training and the results of the testing. Using retrain.py in Google Collab, the dataset may be retrained to improve its identification accuracy. Because more steps lead to greater accuracy, the training dataset is constructed using 50000. The training dataset has a 93% accuracy rate. Approximately a thousand photos with an accuracy of 80% make up the dataset being tested. In addition, a dataset is verified with a 75 percent accuracy in order to improve system performance. An picture is temporarily kept in a database whenever a user uploads an input file to a website. The system feeds this data into the CNN, which is subsequently paired with a training dataset. There are several convolutional layers in a CNN. There are a variety of alignments/features that are taken into account while classifying a bird to ensure optimal accuracy. These properties are extracted from the network's many levels using a convocational network. It is then classified using an unsupervised approach known as deep learning, which uses a neural network (NN). Furthermore, the picture is classified pixel by pixel using a grey scale approach. These characteristics are then combined and sent to a classification algorithm. A trained dataset is used to produce probable outcomes from the input. During categorization, a network of nodes is created, which eventually becomes an autograph. A score sheet is built on the basis of this network, and output is generated using the score sheet.

## V. EXPERIMENTAL ANALYSIS

In order to test the suggested method for classifying bird species based on colour traits and other factors such as size, shape, etc., using the Caltech-UCSD Birds 200 dataset. Among the 11,788 annotated photos of birds in this collection are those with a rough segmentation, a bounding box, and binary attribute annotations for each of the 200 species of birds. You may upload photos from your computer's hard drive or from your Google Drive and use Google-Collab as a platform to train your dataset. A labelled dataset is available for image processing classifiers after training. All five species are represented in a dataset that includes around 200 example photos for each species, which contain environmental elements such as grass and trees as well as the animals themselves. Because of the emphasis on size, form, and colour, a bird may be recognised in almost any situation. Prior to histogram histogram segmentation, these criteria are taken into account. The picture has been reduced to a pixel count using the grey scale approach, which creates a value for each pixel and generates nodes based on that value, also known as neurons. The structure of the matched pixels is determined by these neurons in a way that is similar to a network of linked nodes. Tensorflow is able to categorise the picture based on the signature created by the nodes in the image. To construct a score sheet, a classifier compares the picture to the pre-trained dataset photos from Caltech UCSD. The highest score sheet matching value is the outcome of a bird species, and the score sheet comprises the top five match results. The Caltech UCSD has been trained to achieve an accuracy of 80% in this experiment.

Figure No.6 is an example of an image that the algorithm uses to classify a bird from the Northern Hemisphere. Let's have a look at how it's being rated now.

After categorization, the algorithm provides the scorecard below, which shows the likelihood that the bird in question belongs to which species.

Table No. 1: Score Sheet

| Sr.No | Species | Score Obtained |
|---|---|---|
| 1 | Elegant tern | 0.00943 |
| 2 | Red faced cormorant | 0.00924 |
| 3 | Brant cormorant | 0.0085 |
| 4 | Pelagic cormorant | 0.0082 |
| 5 | White pelican | 0.00808 |

The scorecard produced by the system is shown in table no.1. After analysing these results, it has been determined that the species with the highest scores are necessary. The graph below illustrates this outcome.

## VI. CONCLUSION

Using the dataset (Caltech-UCSD Birds 200) for picture classification, the current work studied a technique for identifying bird species using a Deep Learning algorithm (Unsupervised Learning). It has a total of 11,788 images spread over 200 categories. Using a user-friendly website, a user may submit a picture for identification purposes and get the required result. According to the suggested approach, a component is detected and CNN features are extracted from several convolutional layers. The classifier uses the aggregated data to make classifications. The technology has been able to accurately forecast the presence of 80 percent of bird species based on its findings.

## VII. FUTURE SCOPE

Instead of a website, consider developing an android or iOS app. Second, the system may be built utilising the cloud, which can store a big number of data for comparison as well as supply a significant amount of computational power (in case of Neural Networks).

## REFERENCES

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

**[2] Fagerlund, S., 2007. Bird species recognition using support vector machines. EURASIP Journal on Applied Signal Processing, 2007(1), pp.64-64.**

**[3] 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).**

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

**[5] Andr´eia Marini, Jacques Facon and Alessandro L. Koerich Postgraduate Program in Computer Science (PPGIa) Pontifical Catholic University of Paran´a (PUCPR) Curitiba PR, Brazil 80215–901 Bird Species Classification Based on Color Features**

**[6] Image Recognition with Deep Learning Techniques ANDREIPETRU BĂRAR, VICTOR-EMIL NEAGOE, NICU SEBE Faculty of Electronics, Telecommunications & Information Technology Polytechnic University of Bucharest.**

**[7] Xception: Deep Learning with Depthwise Separable Convolutions François Chollet Google, Inc.**

**[8] Zagoruyko, S. and Komodakis, N., 2016. Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer. arXiv preprint arXiv:1612.03928.**

**[9] Inception-v4,Inception-ResNetand the Impact of Residual Connectionson Learning Christian Szegedy, Sergey Ioffe,VincentVanhoucke ,AlexandrrA.Alemi**

**[10] Stefan Kahl, Thomas Wilhelm-Stein, Hussein Hussein, HolgerKlinck, Danny Kowerko, Marc Ritter, and Maximilian Eibl Large-Scale Bird Sound Classification using Convolutional Neural Networks**

**[11] Thomas Berg, Jiongxin Liu, Seung Woo Lee, Michelle L. Alexander, David W. Jacobs, and Peter N. BelhumeurBirdsnap: Large-scaleFine-grainedVisualCategorizationofBirds**

ISSN 2347–3657

Volume 9, Issue 4, 2021

**[12] Bo Zhao, Jiashi Feng Xiao Wu Shuicheng Yan A Survey on Deep Learning-based Fine-grained Object Classification and Semantic Segmentation**

**[13] Yuning Chai Electrical Engineering Dept. ETH Zurich,VictorLempitsky Dept. of Engineering Science University of Oxford, Andrew Zisserman Dept. of Engineering Science University of Oxford BiCoS:ABiSegmentation Method for Image Classification.**