Image based Bird Species Identification

Anisha Singh[1], Akarshita Jain[2], Bipin Kumar Rai[3]
[1,2] *(Scholar, Department of Information Technology, ABESIT, Ghaziabad, U.P, India)*
[3]*(Professor, Department of Information Technology, ABESIT, Ghaziabad, U.P, India)*

Abstract: *Nowadays, birdwatching is a common hobby but to identify their species it requires the assistance of bird books. To provide birdwatchers a handy tool to admire the beauty of birds, we developed a deep learning platform to assist users in recognizing species of birds using a software based on the concept of image recognition.*

This software would recognize the input image by comparing the model with a trained model and then predict the bird species. The details would be given out as an output. Also, it will help us to build the dataset if any image captured or uploaded by the user is unavailable in the dataset then the user can add that image to the dataset.

Keywords: *Bird image recognition, Convolutional neural network (CNN), Machine learning, Deep learning.*

I. INTRODUCTION

Birdwatching is a recreational activity that can provide relaxation in daily life and promote resilience to face daily challenges. It can also offer health benefits and happiness derived from enjoying nature.

Identification of bird and insect's species is a challenging task often resulting in ambiguous labels. Even professional bird and insect watchers sometimes disagree on the species given an image of a bird and insect. It is a difficult problem that pushes the limits of the visual abilities for both humans and computers. Although different bird and insect species share the same basic set of parts, different bird and insect species can vary in shape and appearance. Interclass variance is high due to variation in lighting and background and extreme variation in pose.

Neural Network (NN)- Neural network has a gained a great attention.[1] It is known that, mammal's brain, which consists of many interconnected neurons, that can deal with complex and computational tasks, like face recognition, body motion, and muscles activities control.

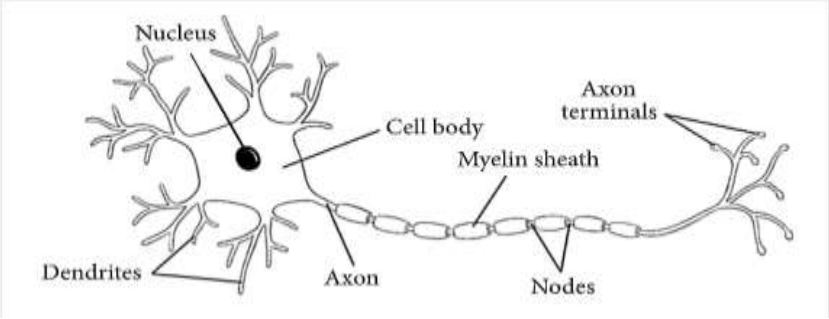

Figure 1: Biological Neural Network [1]

The model of NN consists of three layers, that is, input layer, hidden layer, and output layer. A neural network is a series of algorithms that tries to recognize underlying relationships in a set of data through a process in the exact way the human brain works. In this way, neural networks refer to systems of neurons, either organic or artificial in nature.

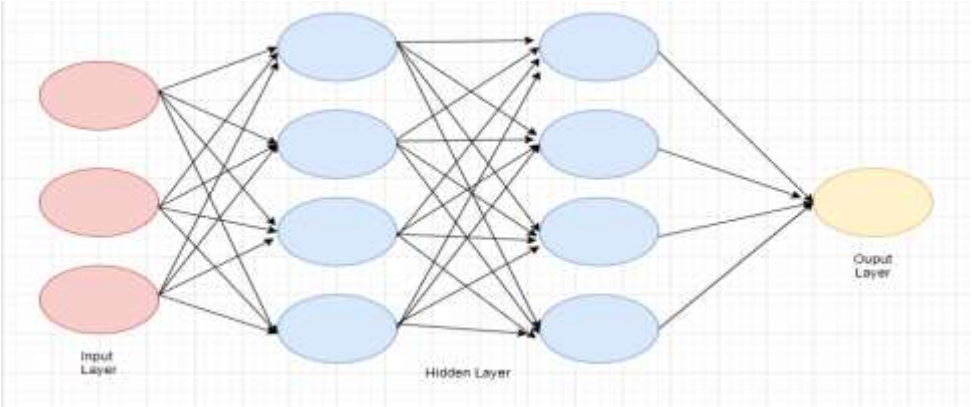

Figure 2: Neural Network

Image processing with NN [17] involves different processes, like:
1. Image preprocessing shows a picture with the same dimensions as the original image. The aim of images preprocessing with NN consists in improving, restoring or rebuilding images.
2. Feature extraction helps to extract a number of features smaller than the number of pixels in the input window. The elements consist in feature extraction are edges, corners, joints, facial features, etc.
3. Segmentation is a division of an image into regions.
4. Recognition involves the determination of objects in an image and their classification.

Convolutional Neural Network- Convolutional neural network (CNN) [2] are feedforward networks in that information flow takes place in one direction only, from their inputs to their outputs.
CNN architectures consist of convolutional and pooling layers. CNN architecture is used for image classification task. An image, input to the network, is followed by convolution and pooling. Thereafter, these operations feed one or more fully connected layers. At last fully connected layer outputs the class label. First, raw input data of numerous semantic parts of a bird were gathered and localized. Second, the features of each generic part were detected and filtered based on shape, size, and color. Third, a CNN model was trained with the pictures in a graphics processing unit for feature extraction with the previously mentioned characteristics, and the classified, trained data were stored in a server to target object. Then, information obtained from an image uploaded by an end-user, captured using a camera, can be navigated to retrieve information and predict from the trained model.

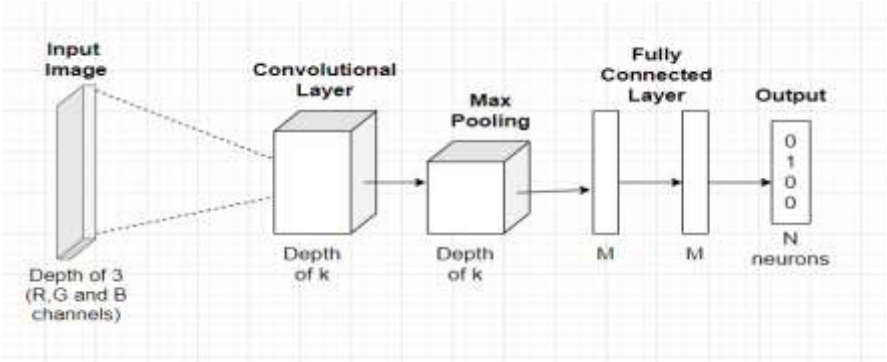

Figure 3: Schematic representation of the architecture of Convolutional Neural Network Related Work

Nadimpalli et al. [4] created a model with the help of image processing technique it recognizes birds in Aqua cultural ponds is a novel concept that enables increased flexibility in distributing predatory birds. Three image processing algorithms image morphology, artificial neural networks, and template matching were designed and tested. They improvised the algorithm to recognize birds in real time conditions and developed necessary algorithms for implementation by using image processing and neural network tool boxes of MATLAB 6.5 version to develop algorithms. The ANN model took three minutes to train the images. However, results were obtained instantly while testing the images.

Christiansen et al. [5] they used digital image processing techniques to automatically detect and study animals on the basis of the video recordings. The thermal radiation of the study animals exceeds the radiation from the background, which makes the animal appear brighter on the video images. But during sunlight periods, the thermal difference between the animal and the background may become smaller and some spots of grass may radiate almost the same temperature as the animals. In that case, Filtering techniques can be implemented to enhance the appearance of the animals. For this purpose, they used the Laplacian of Gaussian filter for pre-processing to enhance the appearance. Under most circumstances, detection rates were close to 100% although dense crops may hamper the detection of animals.

Nadimpalli et al. [6] focuses on bird detection and analyzes the motion detection with image subtraction, bird detection with template matching and bird detection with the Viola-Jones Algorithms. Out of all the methods, bird detection with Viola Jones Algorithm had the highest accuracy (87%) with a low false positive rate. This image processing step would ideally be incorporated with hardware to form a smart scarecrow system. Although the training for the object classifier is slow, the actual detecting is fast that is why there have been some web browser implementation and mobile implementation. The Viola-Jones algorithm can be trained for almost any object as long as there are many similar positive images that can be used for training the classifier.

Moreira et al. [7] represents the state of the art of video detection and tracking of marine vehicles. The marine time environment is very challenging and dynamic. The algorithm for detection and tracking, when used in a marine time environment without proper stability, so it does not produce efficient results. Errors in detection and tracking may occur due to noise, clutter, waves, dynamic and unpredictable ocean appearance, sunlight

reflection, bad environmental conditions and image contrast, presence of objects that float over the ocean, white foam, the great variability of certain marine time vehicle features such as size, maneuverability, appearance, geometric shape and the presence of birds, clouds, fog and aircraft that arises above the horizon. The algorithm seemed not to perform well in some the real time situation when little vessels that have low contrast with the background arise in the camera field of view.

Shalika et al. [8] are concerned with observing animal behavior in wildlife using face detection and tracking. An algorithm for detection and tracking animal in wildlife videos. The detection is based on a human face detection method, using Haar-like features using Ada Boost classifiers. Tracking is done using the Kanade-Lucas-Tomasi tracker and by applying a specific model to the detected face and combining the two methods in a specific tracking model, a reliable and temporarily coherent detection/tracking of animal faces is obtained. In order to the detect particular animal species, the information gathered by the tracker can be used to boost like classification of wildlife videos. The purpose for this project is to make a system for automatically detecting wild animal. This research paper mainly focuses on classification and recognition.

Nguyen et al. [9] uses Wildlife Spotter dataset, consists of a large number of images taken by cameras in Australia, they have developed and illustrated the practicality of a deep learning approach towards constructing automated wildlife monitoring system. With different settings as an experiment for balanced and imbalanced, shows that system is strong, stable and suitable for dealing with images captured. They have worked on many different ways to improve the system's performance by improving the dataset, by applying deeper CNN models and utilizing specific features of camera. For a fully automated wild life recognition system, we would go into transfer learning to deal with problem of imbalanced data. In the future, they will make a "hybrid" wild animal classification system whose automated module work as a recommendation system for the existing citizen Wildlife Spotter project.

Niemi et al. [10] they have collected the non-deep CNN for image classification, and showed that the model is acceptable for real-world application, especially, when the number of training data is limited. They presented and showed that data augmentation method enhances greatly the performance of the classifier, and the desirable state of performance as an image classifier can be achieved by applying it. Thus, they showed that the data augmentation is important for the classification performance. The main performance of the image classifier was obtained without using the parameters given by the radar, those parameters provide more and relevant knowledge to the system and they can turn a wrong classification into the correct. Data collection will be done at the test site which gives large data set.

Nyaga et al. [11] create a mobile application that can recognize Kenya bird species from an image and create a bird map of the observations. It is easy for a human to identify that a given image is a bird. This study starts by reviewing present methods of bird species identification. Machine learning supports image recognition and focus on transfer learning and convolutional neural networks that are used for image processing. This study displayed that advances in computer vision has been able to create deep neural networks that can effectively and efficiently perform fine-grained categorization. Then the study gives a detailed solution to solve the problem of bird species identification using transfer learning and came up with a mobile application for bird species recognition. The model performed well and was able to tell the bird species in an image provided by a user and also allowed to save the results into the database and later on retrieval.

Huang et al. [12] designed an automatic model which was made to classify the 27 endemic birds of Taiwan by skipped CNN model. The purpose behind skip connection was to give an uninterrupted gradient flow from the first layer to last layer, so it can solve the disappearing gradient problem. They do comparison of performance of various models such as CNN with skip connections, CNN without skip connections, and SVM. CNN with skip connection performed the other two algorithms. The proposed model was able to identify the uploaded image of a bird as bird with 100% accuracy. But due to the minute visual similarities between and among the bird species, the model sometime lacks the interspecific comparisons among the bird species. The test dataset yielded 93.79% of sensitivity and 96.11% of specificity.

Gavali et al. [13] tells that instead of identifying number of different categories, the problem of identifying a large number of classes in one category of bird was suspected. Classifying birds pose an extra challenge over categories, because of the large similarity between classes. Also, birds are flexible objects that can disfigure in many ways, and at the same time there is also a large variation within classes. So, the current study finds a method to identify the bird species using Deep learning algorithm on the dataset for classification of image. It has 200 categories or 11,788 photos. The generated system is connected with a user-friendly website where user will upload photo and it will give the desired output. The proposed model works on the principle of detection of a part and extracting CNN feature from multiple convolutional layers. These features are extracted and then given to the classifier for classification purpose.

III. PROPOSED SOLUTION

The proposed solution of the model is given below explains the working of the project, the user would be able to capture and upload the image to the system and can store the image in the database if that image is not available in the dataset. Then the image would be fed to the system and CNN would be applied. After that the features of the image such as face, expression, angle, beak etc. would be extracted and classifier would classify the image and predict with the help of the trained dataset.

Diagrams:

Work Flow Diagram:

The work flow of the model is like a user will upload the image as an input and it will store in database if it is not there. Then system will fetch the image and apply CNN on it and it will be compared with the trained model. After that features of the image will be extracted and the classifier will classify the extracted features and display the required result.

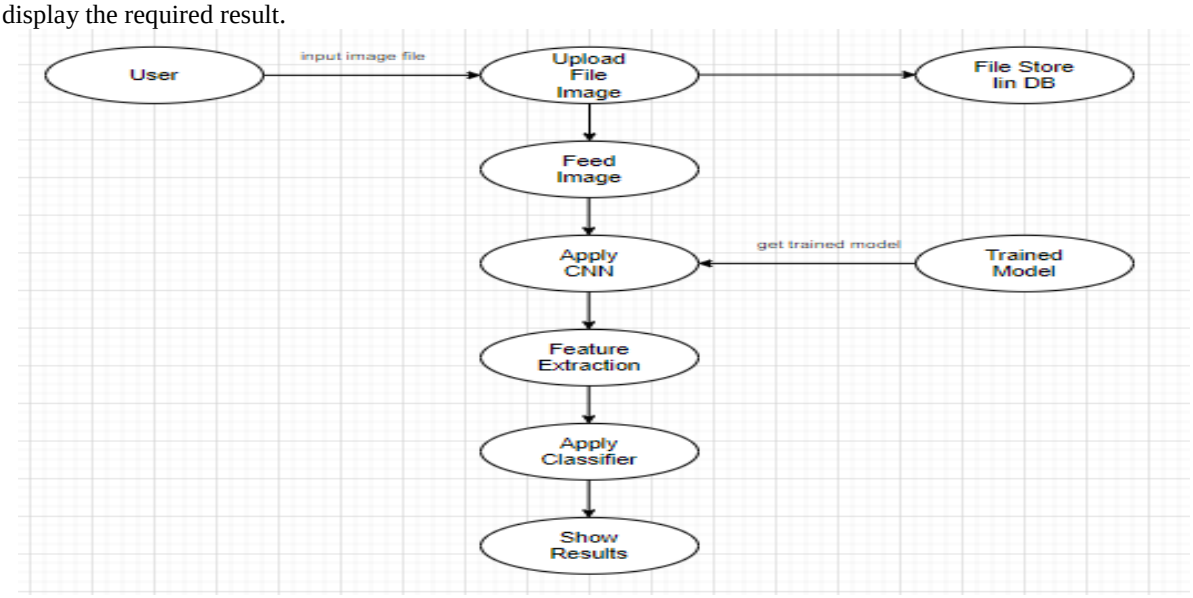

Figure 4: Work Flow Diagram

Use Case Diagram:

The use case below shows that the model would be accessed by 2 participants, the administrator and the user. The system administrator would be able to pre-process the dataset and train the model then he would generate a model which will help in comparison. Afterwards he can upload an image by either capturing it or by choosing it from there gallery and make the upgradations. Meanwhile the user would be able to register and login into the system and make inference by uploading an image by selecting it from the gallery or by capturing it, then he would be able to get the desirable results.

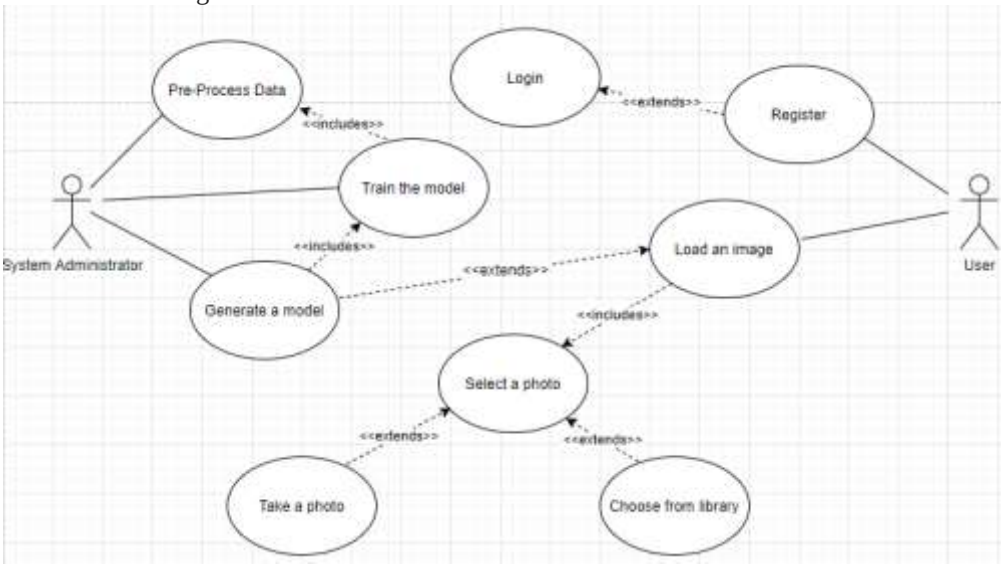

Figure 5: Use Case Diagram

IV. IMPLEMENTATION

The implementation of the model is done with the help of OpenCV:

OpenCV- Object detection is related to computer vision and image processing that detects semantic objects of a certain class such as humans, buildings, or cars in digital images and videos. OpenCV is an open source computer vision. It is an image processing library created by Intel and later supported by Willow Garage and now maintained by Itseez. It is available on Mac, Windows, Linux. It works in C, C++, Python. [3] Library of OpenCV is a collection of algorithms and C/C++ functions. OpenCV was used for computational efficiency and real time applications in many areas like vision, factory product inspection, medical imaging, stereo vision, security, user interface, camera calibration and robotics. The goal of OpenCV is to provide a simple use computer vision infrastructure that helps people build quite sophisticated vision applications quickly, vision research by providing open and optimized code for basic vision infrastructure, spread vision knowledge by providing a common infrastructure so that code would be more readily readable and transferable, advance-vision based commercial applications by making portable, performance optimized code available for free.

This is how the Image Processing in the model will be done-

1)Acquisition- It could be very simple. The main work involves-

i)Scaling

ii)Color conversion (RGB to gray or vice versa)

2)Image enhancement- It is the simplest amongst the all and most appealing in areas of Image Processing it is also used to extract some hidden details from an image.

3)Image restoration- It also deals with appealing of an image. It is based on mathematical or probabilistic model or image degradation.

4)Color image processing- It deals with pseudo color and full color image processing color models are applicable to digital image processing.

5)Wavelets and multi-resolution processing- It is foundation of representing images in various degrees.

6)Image compression- It involves in developing some functions to perform this operation. It mainly deals with the image size or resolution.

7)Morphological processing- It deals with tools for extracting image components that are useful in the representation and description of shape.

8)Segmentation procedure- It includes partitioning an image into its constituent parts or objects. Autonomous segmentation is most difficult task in image processing.

9)Representation and description- It follows output of segmentation stage, choosing a representation is only the part of solution for transforming raw data into processed data.

10)Object detection and recognition- It is process that assigns a label to an object based on its descriptor.

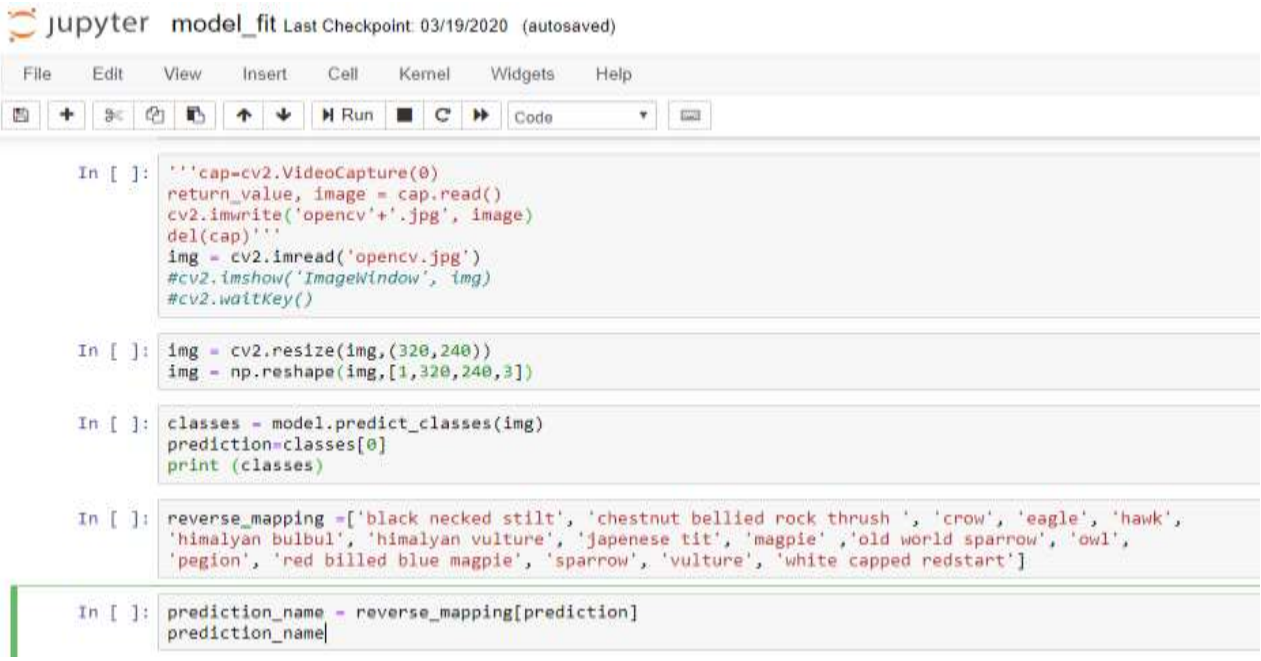

Figure 6: Image recognition

Algorithm-

1. Obtain from test image the location (A, B) and magnitude(R) of N strongest scatters.
2. Order (A, B, R) triples by descending R.
3. For each origin O from 1 to N do 4.
4. For each point M from O+1 to N do 5,6
5. $d_A = A_p - A_o$; $d_B = B_p - B_o$
6. For D_A from d_A-1 to d_A+1 do 7
7. For D_B from d_B-1 to d_B+1 do 8, 9, 10
8. Weighted_vote = $|D_A| + |D_B|$
9. Look up list of model data entries at D_A, D_B
10. For each entry C in the list do 11
11. If $|t_a = A_o - A_e| <$ translational_limit and $t_b = B_o - B_e| <$ translational_limit and $|1 - R_{eo}/R_o| <$ magnitude_limit and $|1 - R_{ep}/R_p| <$ translational_limit THEN increment accumulator array [Specie, A_z, t_a, t_b] by weighted_vote.
12. Query accumulator array for each Specie, A_z, t_a and t_b, summing the votes in a 3x3 neighborhood in translation subspace about t_a, t_b; record the maximum vote_sum and the corresponding Specie.
13. IF maximum vote_sum > threshold THEN result is Specie ELSE result is “Not found”.

Result

This study developed a software platform that uses deep learning for image processing to identify bird species from digital images uploaded or captured by an end-user on a smartphone in real time. [13] To develop such system a trained dataset is required to classify an image. Trained dataset consists of two parts trained result and test result. The dataset has to be retrained to achieve higher accuracy in identification. The trained dataset is created using 50000 steps, higher the number of steps higher its accuracy. The accuracy of trained dataset is 93%. The testing dataset has nearly 1000 images with an accuracy of 80%. Whenever a user will upload an input file, the image is temporarily stored in database. This input file is then passed to the system and is given to CNN where CNN is coupled with trained dataset. Various features such as head, body, color, beak, shape, entire image of bird is considered for classification to have maximum accuracy. Each feature is given through deep convocational network to extract features out. These features are then collected and forwarded to classifier. The input will be compared with the trained dataset to generate results.

Image is compared with the pre trained dataset images and the score sheet is generated. The score sheet is an output of top 5 match results by which the highest matching value of score sheet is the result of bird species.

Consider below figure as input given to the system for classification of bird. Let’s see the procedure.

Figure 7: Image of Elegant tern [13]

S.No.	Species	Score Obtained
1.	Elegant tern	0.00943
2.	Red Faced cormorant	0.00924
3.	Brant cormorant	0.0082
4.	Pelagic cormorant	0.0082
5.	White pelican	0.00808

Figure 8: Score Sheet Table

The above table shows the score sheet based on the result made by the system.

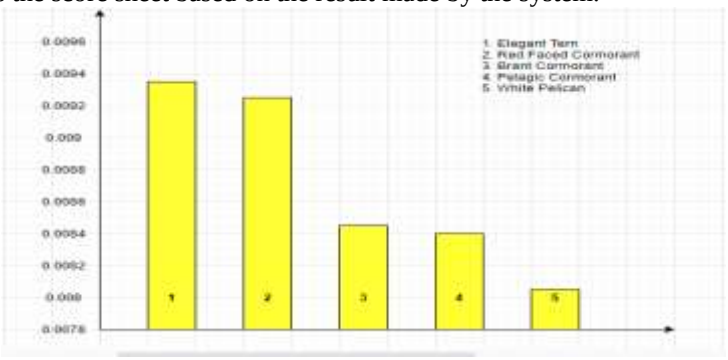

Figure 9: Graph for Score Sheet

After analysis of these result it has noticed that, the species those are having the highest score has been predicted as a required species. This result can be shown in the above graph.

V. CONCLUSION

Neural network is used to classify bird species based on data set. [16] Multiple-width frequency delta data augmentation cannot be used to raise classification accuracy when compared to raw spectral data, but the accuracy is close to the state-of-the-art and has an advantage over raw spectral data when computational resources are limited. Then, the use of additional meta-data raise the rank of the species in the predictions of the models, but it does not seem to be enough to push it to the highest rank, which means that the model has to predict fewer species, but the actual top-1 accuracy does not seem to be affected. Through an analysis of the data set we also found that the relative number of training samples for each bird species is quite uneven, which seems to lead to a favoritism, from the model of bird species, and that some bird species are difficult to classify than others.

This study [15] will help the researchers to work on various fields such as image processing, fault detection in industrialized Industries, medical image segmentation. The biggest disadvantage of all these algorithms is that the accuracy of these algorithms is dependent on the quality of camera and view angle between camera and the target object. It is also noticed that at some angles the results were not accurate beyond a certain range of camera.

Future Scope

The future of image processing involves new intelligent, digital automated robots made by research scientists in various parts of the world. [15] It includes development in various image processing applications. Due to changes in image processing and other related technologies, there will be millions of robots in the world in a few, transforming the way of living. Researches in image processing and artificial intelligence will involve voice commands, anticipating the information requirements of governments, translating languages, recognizing and tracking people and things, diagnosing medical conditions, performing operation & surgery, reprogramming defects in human DNA, and automatic driving all formats of transportation.

And for Image based species recognition of birds [6] we can further enhance the system with cloud feature which can store large amount of data for comparison and in case of neural network it can provide high computing power for processing.

VI. REFERENCES

[1] Jiang, Y., Yang, C., Na, J., Li, G., Li, Y., & Zhong, J. (2017). A brief review of neural networks-based learning and control and their applications for robots. Complexity, 2017.

[2] Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. Neural computation, 29(9), 2352-2449.

[3] Mohamad, M., Saman, M. Y. M., Hitam, M. S., & Telipot, M. (2015). A Review on OpenCV. Terengganu: Universitas Malaysia Terengganu.

[4] Nadimpalli, U. D. (2005). Image processing techniques to identify predatory birds in aqua cultural settings.

[5] Christiansen, P., Steen, K. A., Jorgensen, R. N., & Karstoft, H. (2014). Automated detection and recognition of wildlife using thermal cameras. Sensors, 14(8), 13778-13793.

[6] Nadimpalli, U. D., Price, R. R., Hall, S. G., & Bomma, P. (2006). A comparison of image processing techniques for bird recognition. Biotechnology progress, 22(1), 9-13.

[7] Moreira, R. D. S., Ebecken, N. F. F., Alves, A. S., Livernet, F., & Campillo-Navetti, A. (2014). A survey on video detection and tracking of maritime vessels. International Journal of Recent Research and Applied Studies, 20(1).

[8] Shalika, A. U., & Seneviratne, L. (2016). Animal Classification System Based on Image Processing & Support Vector Machine. Journal of Computer and Communications, 4(1), 12-21.

[9] Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., ... & Phung, D. (2017, October). Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In 2017 IEEE international conference on data science and advanced Analytics (DSAA) (pp. 40-49). IEEE.

[10] Niemi, J., & Tanttu, J. T. (2018). Deep Learning Case Study for Automatic Bird Identification. Applied Sciences, 8(11), 2089.

[11] Nyaga, G. M. (2019). A Mobile-based image recognition system for identifying bird species in Kenya (Doctoral dissertation, Strathmore University).

[12] Huang, Y. P., & Basanta, H. (2019). Bird image retrieval and recognition using a deep learning platform. IEEE Access, 7, 66980-66989.

[13] Gavali, P., Mhetre, M. P. A., Patil, M. N. C., Bamane, M. N. S., & Buva, M. H. D. (2019) Bird Species Identification using Deep Learning.

[14] Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., ... & Phung, D. (2017, October). Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In 2017 IEEE international conference on data science and advanced Analytics (DSAA) (pp. 40-49). IEEE.

[15] Dewangan, S. K. (2016). Importance & Applications of Digital Image Processing. International Journal of Computer Science & Engineering Technology (IJCSET), 7(7), 316-320.

[16] Martinsson, J. (2017). Bird Species Identification using Convolutional Neural Networks (Master's thesis).

[17] Pandelea, A. E., Budescu, M., & Covatariu, G. (2015). Image Processing Using Artificial Neural Networks. Buletinul Institutului Politehnic din lasi. Sectia Constructii, Arhitectura, 61(4), 9.