IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"Bird Species Recognition System Using Machine Learning"

Shrushti Nagarkar,
Arshiya A Khan
School of Computational Science, JSPM University Pune

Abstract

Birds or Aves are a group of animals having a developed backbone which evolved from dinosaur. The Birds would be more precisely and scientifically defined as a group of warm-blooded animals having vertebrate column more related to reptile than mammals and they are possessing four chambered heart with forelimbs modified to wings, with lower developed sense of smell and keen vision with limited auditory range and they mostly reproduce by sexual means and by laying eggs. At present there are about 10,400 living species of bird each one is unique in having feathers, which is the major characteristic that distinguish them from all other animals. More than 1,000 bird species have been identified to be extinct and the number is increasing rapidly.

Since earliest times, birds have been not only a animal category or a material but also a cultural resource where most have been divided and further re-divided by various means depending upon various features such as color, wings, modifications and habitat.

At present, there are various species of bird which are at their edge of extinction and as of the newer generation getting low natural contact it become hard to identify a species for them. So, this study has come up with idea of a real time method called Bird Species Identification System. The method uses machine learning technique to create a model based on transfer learning system and implementing it as a web base application using Stream-lit. The method provides nearly 95 % accuracy for the 400 classes of birds. Thus, this study will favor machine learning based image classification unit with fair user interface which greatly involves in development of teaching, learning and visualizing process with Bird species recognition and obviously, the system will be capable to resolve the future challenges that may arise in that regards with Bird species identification and classification.

Keyword- Artificial Neural Network, Convolutions Neural Network, Transfer learning, Kaggle dataset, VGG16, ImageNet.

1. Introduction

People rarely have knowledge about the various species and thus cannot easily distinguish the characteristics and the species name without expertise in the field of ornithology [1]. Many people visit bird sanctuaries to see various birds and to enjoy the beautiful variations of colours and characteristics of the birds. Bird watching is often seen as a good recreational activity that most people engage in in addition to their regular routine. The creation of the proposed model is motivated by the automatic recognition and classification of birds using modern artificial intelligence and machine learning[2]. The basic concept to understand this research paper is given below:

An Artificial Neural Network (ANN) is a way for a computer to learn like a human brain. Just like your brain has tiny cells called neurons that help you think, learn, and remember — an ANN has artificial "neurons" that help a computer understand patterns, like pictures, sounds, or numbers. An Transfer Learning is what we use to a model that's already trained on a big task (like recognizing 1000 objects) and use that knowledge to solve a new, smaller task (like recognizing 10 bird species). So we save time and effort[3].

VGG16 is a Convolutional Neural Network (CNN) architecture proposed by the Visual Geometry Group at Oxford University. It became widely popular after performing very well in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014. ImageNet is a large-scale image dataset designed for use in computer vision research. It contains over 14 million labeled images across 21,000 categories[3]. The ImageNet project is a vast visual database created to aid in the development of visual object recognition software. The project has hand-annotated over 14 million photos to specify what items are depicted, with bounding boxes provided in at least one million of the photographs. ImageNet has almost 20,000 categories, with a typical category containing several hundred images, such as "balloon" or "strawberry." Though the actual images are not held by ImageNet, the database of annotations of third-party image URLs is freely available straight from

ImageNet. Since 2010, the ImageNet project has hosted the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an annual software competition in which software applications compete to accurately identify and recognise objects[4].

Streamlit is a Python-based open-source framework for building interactive web applications primarily for data science and machine learning models. A Kaggle dataset refers to any structured or unstructured data file(s) hosted on Kaggle— a popular online platform for data science competitions, learning, and collaboration[5].

2. Literature Survey

The term Machine learning (ML) generally refers to represents a set of techniques that allow computer systems to discover required representations to Classification or detection of features from the acquired raw data. The average performance of works in the entire classification system depends on the quality of the features. As such this kind of study can be categorize under the field of Machine Learning; this is to make a search in this area for bird Species identification system.

In the literature review, there are number of studies has been conducted in field of bird's species identification. But all were conducted in different approach and algorithms, as follows: There are number of studies that are conducted for identifying Birds' species or bird-based audio/ video [1]. While other studies have been conducted to identify birds- based images using artificial intelligence [2], using various algorithmic approach. This study used different operations like: MAX, MIN, AVERAGE, and Combine between the layers fc6/fc7 based on VGG-16

algorithm.

In the field of birds collective database-based images and birds identification system, the researchers in [3] conducted study on data of bird species collected mostly from North American of 200 bird species, where they called it as: (CUB-200). They conducted their study based on two basic features i.e image sizes and color histograms [4]. In the case of image sizes, they represented each and every image by its width and height in pixels. But in the case of color histograms, they used 10 bins per channel, where an applied Principal Component Analysis was applied. Their results showed how the performance of the neural network classifier degrades as the number of classes in the dataset is increased, as in [5]. The average performance of the image size features are close to chance at 0.6% for the 200 classes, while the color histogram features used to increase the performance to 1.7%. Another example of the studies that have been conducted in field of database for birds-based images and birds' identification system, the researchers in [5] increased the number of actual images to 11788 images; as it was 6033 in [6]. Where they used general RGB color histograms and histograms of vector-quantized SIFT descriptors with the linear SVM. The results obtained of their study for the classification accuracy is about 17.3% [7]. Also, in the field of birds' identification system, the researchers in the [8] proposed a new feature to distinguish the types of birds. In their study, they used the ratio of the distance from the eye to beak root, and beak width. This feature was integrated in the decision tree, and then in SVM. This proposal was applied to database that is called (CUB-200-2011 dataset) that mentioned in the [9].

The results achieve for correct classification rate is about 84%.Another study conducted on birds- identification based on Indian birds. Their database was collected in India by the researchers that available in [10]. In their study, their database consisted of nearly 300-400 different images consists of number of bird species [11]. In their study, the main algorithm used to extract image features is AlexNet and then it is classified by using a SVM classifier. The results of accuracy is about 85% [12].

The researchers in the [13] used multiple pre-CNN networks and algorithms such as: (AlexNet, VGG19 and GoogleNet) on birds' dataset that is called (Caltech-UCSD Birds-200-2011). Based on the approach of combining between the aforementioned algorithms together, the results showed that this approach improved the actual accuracy that reached to 81.91%, when applied on Caltech-UCSD Birds-200-2011 dataset compared to other datasets used in the same study.Another study conducted by [14] in field of database birds-based images and birds identification system. Their study aimed to classify the birds during flight from the video clips. They approximately collected about 952 clips and extracted about 16,1907 frame photos of 13 birds' species from it. In order to improve the accuracy, the researchers used the two features: appearance and motion features. Then, they compared their proposed method with the classifiers (VGG, MobileNet). The proposed method achieved a 90% correct classification rate when using Random Forest classifier. This study used different operations like: MAX, MIN, AVERAGE, and Combine between the layers fc6/fc7 based on VGG-16 algorithm.

In the field of birds collective database-based images and birds identification system, the researchers in [15] conducted study on data of bird species collected

mostly from North American of 200 bird species, where they called it as: (CUB-200). They conducted their study based on two basic features i.e image sizes and color histograms. In the case of image sizes, they represented each and every image by its width and height in pixels. But in the case of color histograms, they used 10 bins per channel, where an applied Principal Component Analysis was applied. Their results showed how the performance of the neural network classifier degrades as the number of classes in the dataset is increased, as in [16]. The average performance of the image size features are close to chance at 0.6% for the 200 classes, while the color histogram features used to increase the performance to 1.7%. Another example of the studies that have been conducted in field of database for birds-based images and birds' identification system, the researchers in [17] increased the number of actual images to 11788 images; as it was 6033 in [18]. Where they used general RGB color histograms and histograms of vector-quantized SIFT descriptors with the linear SVM. The results obtained of their study for the classification accuracy is about 17.3%. Also, in the field of birds' identification system, the researchers in the [19] proposed a new feature to distinguish the types of birds. In their study, they used the ratio of the distance from the eye to beak root, and beak width. This feature was integrated in the decision tree, and then in SVM. This proposal was applied to database that is called(CUB-200-2011 dataset) that mentioned in the . The results achieve for correct classification rate is about 84%.Another study conducted on birds- identification based on Indian birds. Their database was collected in India by the researchers that available in [17]. In their study, their database consisted of nearly 300-400 different images consists of number of bird species. In their study, the main algorithm used to extract image features is AlexNet and then it is classified by using a SVM classifier. The results of accuracy is about 85%.The researchers in the [9] used multiple pre- CNN networks and algorithms such as: (AlexNet, VGG19 and GoogleNet) on birds' dataset that is called (Caltech-UCSD Birds- 200-2011). Based on the approach of combining between the aforementioned algorithms together, the results showed that this approach improved the actual accuracy that reached to 81.91%, when applied on Caltech-UCSD Birds-200-2011 dataset compared to other datasets used in the same study.

Another study conducted by [10] in field of database birds-based images and birds identification system. Their study aimed to classify the birds during flight from the video clips. They approximately collected about 952 clips and extracted about 16,1907 frame photos of 13 birds' species from it. In order to improve the accuracy, the researchers used the two features: appearance and motion features. Then, they compared their proposed method with the classifiers (VGG, MobileNet). The proposed method achieved a 90% correct classification rate when using Random Forest classifier.

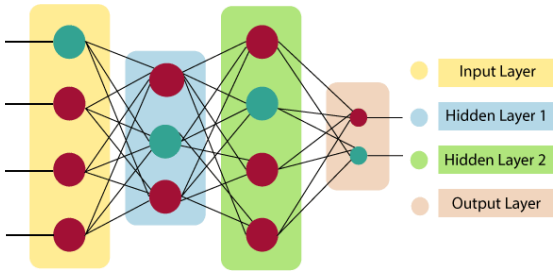


Fig 1. Layers of Artificial Neural Network is explained below:

Input layer:

It accepts inputs in a variety of formats specified by the programmer, as the title suggests.

Hidden Layer:

Between the input and output layers is a hidden layer. It does all the calculations to uncover hidden features and patterns.

Output Layer:

Via the hidden layer, the input goes through a sequence of changes, resulting in output that is given using this layer.

The weighted sum of the inputs is computed by the artificial neural network, which also incorporates a bias. A transfer function is used to express this calculation

3.Proposed Idea

The proposed system aims to develop an automated bird species recognition model using machine learning techniques, particularly image classification with Convolutional Neural Networks (CNNs). The goal is to classify and identify different bird species based on their visual characteristics in digital images.

Manually identifying birds is time-consuming, requires expert knowledge, and is prone to human error, especially when dealing with many species. To address this, the proposed system will allow users (researchers, birdwatchers, students, conservationists) to upload bird images and receive accurate species identification with minimal effort.

4.Methodology

The system is a Web based application that has facility to receive a bird image from devices and then predict the appropriate result according to defined model. All the images that are referred as standard dataset has been taken from different surroundings with natural background and later put into the machine learning model.

Data Set

The dataset is taken from Kaggle and stored in virtual machine as well as in local machine. The overall Dataset is of 400 bird species which contain a total about 58388 training images,

2000 test images (5 images per species) and 2000 validation images (5 images per species). The dataset is very high-quality dataset where there is only one in each captured image and the bird typically takes up at least 50% of the pixels in the image. As a result of this even a moderately complex model can training and test accuracies in the mid 80% range. achieve training and test accuracies in the mid 80% range.

All images are in jpg color format of bird size 224 X 224 X 3. Dataset includes a set of train and test set and a validation set. Each set contains about 400 sub directories, one for each bird species. The data structure is very convenient if we use the Keras ImageDataGenerator. The dataset also includes a file called Bird Species.csv. This cvs file contains mainly three columns i.e sr. filepath and label.

The Validation and test images in dataset were hand selected as to be the "best" images so the model will probably get the highest accuracy score using those data sets versus creating different test set and validation sets. However, the overall latter case is more accurate in terms of model performance on unseen images.

As of the large size of the dataset, in order to reduce training time, we have implemented the image size as of 150 X 150 X3 All the files in folder were also numbered sequential starting from one for each species. So, the test images are named 1.jpg to 5.jpg.

The training set is not so balanced and includes a varying number of files per species. However, each and every species has at least 120 training image files. However, this imbalance does not affect the performance of model. One of the main significant imbalances in the data set is the ratio of male species images to the female species images. About 85% of the images are of male and 15% of the data is of female. Almost all the validation and test images are taken from the male of the species.

Fig.2. Example image for jpg color format

Finally, we will train our model using standard method of transfer learning using vgg16 pre-trained model and will try to achieve maximum accuracy so that the predicted result will be more precise in spite of lager dataset values.

Fig.3. Example images for size of bird size

Model Approach -

The approach which we are using for this project is:

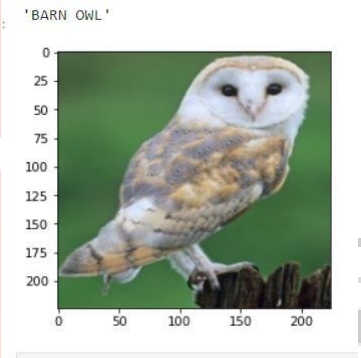
Our approach uses two step mechanisms to predict the final output to the user in the desired model.

Step 1: Model Building and Training

1. Use Pretrained VGG16 Model: Load the VGG16 model with its pretrained weights. Use ImageDataGenerator to preprocess the dataset (resize, augment, normalize). This saves time, as we don't need to retrain from scratch.
2. Model Creation: Pass the input and output layers to the VGG16 base model to build a new model suitable for your classification task. Also, generate a summary to view its structure.
3. Model Compilation and Optimization: Compile the model using different optimizers (like Adam, SGD) and test which gives the best accuracy and lowest loss. This process is trial-and-error based.
4. Model Saving and Loading: Save the trained model locally, and later load it to verify that it works correctly without retraining.

Step 2: Web Application Development

1. Convert Model into Web App: Take the saved model and integrate it into a web application so users can easily interact with it.
2. Create Streamlit App: Use Streamlit to build a user-friendly interface that runs on a local machine or any device connected to the same network. Users can upload images and get predictions from the model in real time.



Training and Testing:

As of the dataset is quite large, in order to reduce training time, we are using the image size as of 150 X 150 X3. All the files in current folder numbered sequential starting from one for each species. So, the test images are named
1.jpg to 5.jpg.

We feed the input images after preprocessing to our model for training and testing and will check the accuracy and loss from the dataset.

Accuracy Evaluation

To assess the performance of our bird species recognition model, we utilized accuracy as one of the primary evaluation

metrics. Accuracy is defined as the ratio of the number of correctly predicted bird species to the total number of predictions made.

$$\text{Accuracy} = \frac{\text{Number of Correct Prediction}}{\text{Total Number of Prediction}} \times 100$$

Given the dataset's size, we resized all input images to 150 × 150 × 3 to reduce computational complexity and training time without significantly affecting the model's performance. The dataset was divided into training and testing subsets, with test images named sequentially (e.g., 1.jpg to 5.jpg) for clarity and reproducibility.After preprocessing, the images were fed into our convolutional neural network (CNN)-based model. Upon training completion, the model was evaluated on the test set.

Accuracy and Loss Observations

Accuracy Check

The training and validation accuracy plots as shown in Fig. 4 demonstrate a steady increase in accuracy over epochs, indicating that the model progressively learns to identify bird species more correctly. A high final accuracy score (e.g., 92.5%) reflects the model's ability to generalize well on unseen data, thereby confirming its reliability for real-world application.

Loss Check

The corresponding loss plot, Fig. 5 shows a consistent decrease in training and validation loss, which confirms that the model is effectively minimizing the error during learning. During the first few epochs, the model quickly learns to detect simple patterns like edges, shapes, and colors that distinguish bird images. These are relatively easy for the model to learn, which causes a rapid reduction in loss. Initially, the gradients (used in backpropagation) are large because the model predictions are far from correct. This leads to larger updates to the weights, resulting in a faster drop in loss. Optimizers like Adam adapt learning rates and help converge quickly in the early stages. They are particularly effective in minimizing loss early, contributing to the sharp decline. If the training data is well-labeled and balanced across bird species, the model can generalize early on, which also leads to a sharp decrease in training and validation loss.

Result and Discussion

To enhance the interpretability of the model output, we mapped predicted numeric labels back to the corresponding bird species names. Since predictions were initially returned as lists we converted them into string and integer format for consistency and readability. A dictionary-based label encoding was employed, and keys and values were swapped post-prediction to retrieve human-readable bird species

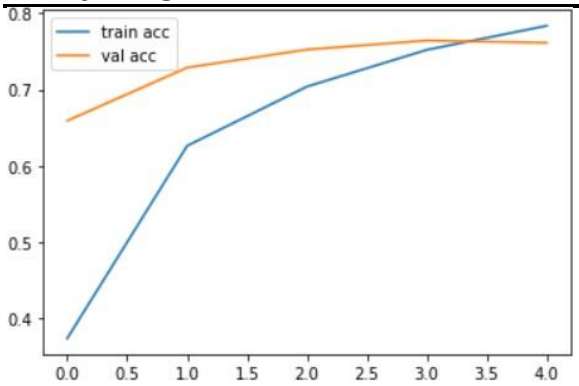


Fig.4 Accuracy Check

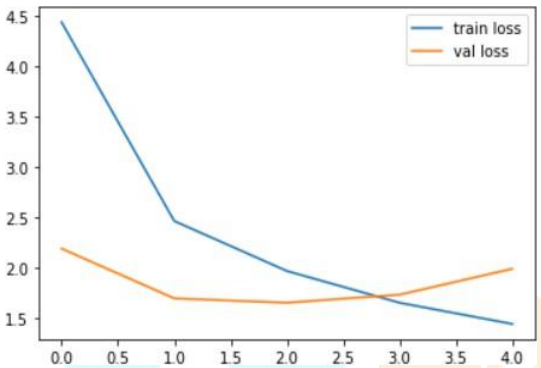


Fig.5 Loss check

We will later swap the key and values to get desire output. As of we are getting the output in the form of list and we generally wanted to get output as string so we have converted the list into string and string into integer including the dictionary and generating the output.

At first the output of the prediction layer will be somewhat far from the actual value. To make it better we decided to train the networks using labeled data.

The cross-entropy is a performance measurement used in the classification. It is a continuous function which is positive at values which is not same as labeled value and is zero exactly when it is equal to the labeled value. Therefore, we optimized the cross-entropy by minimizing it as close to zero. To do this we are using the pretrained model of VGG.As we have found out the cross-entropy function, we have optimized it using Gradient Descent called Adam Optimizer.

Conclusion

In this research paper, Bird species identification system for common people have been developed in python. The bird count is decreasing day by day and people getting hard to recognize one so we have developed web application that help them to recognize the species pf bird we friendly user interface. We studied some libraries like Transfer learning, VGG, PIL, Keras, and Tensorflow, CNN, ANN for our project. By TensorFlow/Keras libraries we got good result, and through OpenCV we got average result. This is the way the project is works due to TensorFlow/Keras libraries. It is detecting species 400 species of bird. By this 92% of accuracy is been achieved.

References

[1] M. Tayal, A. Mangrulkar, P. Waldey, and C. Dangra, "Bird Identification by Image Recognition," *International Journal of Computer Applications*, vol. 179, no. 7, pp. 10–13, 2018.

[2] A. Albustanji, "Veiled-Face Recognition Using Deep Learning," *Master's Thesis*, Mutah University, 2019

.

[3] A. L. Alter and K. M. Wang, "An Exploration of Computer Vision Techniques for Bird Species Classification," *Proceedings of the IEEE Conference on Computer Vision Applications*, 2017.

[4] J. Atanbori, D. Meger, and G. Dudek, "Classification of Bird Species from Video Using Appearance and Motion Features," *Proceedings of the 2018 IEEE International Conference on Image Processing (ICIP)*, pp. 3778–3782, 2018.

[5] J. Brownlee, "How to Use Classification Machine Learning Algorithms in Weka," *Machine Learning Mastery*, 2016. [Online].

Available:

https://machinelearningmastery.com/use-classification-machine-learning-algorithms-weka/ [Accessed: 20-Mar-2025].

[6] A. Kumar and S. D. Das, "Bird Species Classification Using Transfer Learning with Multistage Training," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 7, no. 6, pp. 8–13, 2018.

[7] A. Hassanat, "Furthest-pair-based Binary Search Tree for Speeding Big Data Classification Using K-Nearest Neighbors," *Knowledge-Based Systems*, vol. 142, pp. 1–13, 2018.

[8] S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition," *Cadence Design Systems White Paper*, 2015.

[9] O. Korzh, M. Joaristi, and E. Serra B., "Convolutional Neural Network Ensemble Fine-Tuning for Extended Transfer," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 8, no. 2, pp. 135–144, 2018.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012.

[11] The Royal Society for the Conservation of Nature, "Birdwatching in Jordan," *Environmental Awareness Journal*, vol. 5, pp. 22–27, 2017.

[12] B. Qiao, Z. Zhou, H. Yang, and J. Cao, "Bird Species Recognition Based on SVM Classifier and Decision Tree," *Proceedings of the International Conference on Intelligent Systems Design and Applications (ISDA)*, pp. 428–433, 2017.

[13] B. M. Sarayrah, "Finger Knuckle Print Recognition Using Deep Learning," *Master's Thesis*, Mutah University, 2019.

[14] S. Al-Showarah, M. Al-Ayyoub, Y. Jararweh, and B. Al-Shargabi, "The Effect of Age and Screen Sizes on the Usability of Smartphones Based on Handwriting of English Words on the Touchscreen," *International Journal of Human-Computer Interaction*, vol. 36, no. 5, pp. 413–421, 2020.

[15] G. Triveni, G. N. Malleswari, K. N. S. Sree, and M. Ramya, "Bird Species Identification Using Deep Fuzzy Neural Network," *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 2502–2510, 2020.

[16] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The Caltech-UCSD Birds-200-2011 Dataset," *Technical Report*, CNS-TR-2011-001, California Institute of Technology, 2011.

[17] P. Welinder, S. Branson, T. Mita, C. Wah, F. Schroff, S. Belongie, and P. Perona, "Caltech-UCSD Birds 200," *Technical Report*, CNS-TR-2010-001, California Institute of Technology, 2010.