Bird Species Identifier using Convolutional Neural Network

Kamlesh Borana
IT student, dept. Information Technology
VCET Mumbai, India

Umesh More
IT student, dept. Information Technology
VCET Mumbai, India

Rajdeep Sodha
IT student, dept. Information Technology
VCET Mumbai, India

Prof. Vaishali Shirsath
Assistant Professor, dept. Information Technology
VCET Mumbai, India

Abstract—**Due to climate changes, many species of flora and fauna are endangered. In order to protect them, we must first identify the species to which it belongs and the special care which needs to be taken care of for their survival. More than 10,000 species are part of the ecosystem. Identifying the species of the bird from an image is a challenging task as it requires techniques such as image processing and Convolutional Neural Network (CNN). CNN is a very challenging Research Area with lots of issues as a slight variation in the image can be perceived as a completely new image. In our approach, we are using the transfer learning approach for training our neural model.**

Keywords— Bird species identification, image processing, transfer learning, convolutional neural network.

I.   INTRODUCTION

Our living ecosystem consists of various types of species such as humans, animals, birds, etc. Our research focuses on identifying the species of the birds. By protecting these bird species its will create a huge positive impact on ecological balance, agricultural as well as forestry production. To protect these bird species, we firstly require accurate information about their species. For identification purposes we creating a neural model where the user can upload the image that image will be processed by the neural model and providing the output to the user the species of bird. Creating our own neural network model for the species identification task will require greater amount of data i.e. images of a bird with their annotation as well as its needs huge computing power to create a neural model from scratch but it will not provide assurance that it will perform the better result, so better option is to use the pre-trained model and perform the transfer learning on our dataset..

II.   PROBLEM STATEMENT

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 analysed 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.   MOTIVATION

As there are hardly any software available for Bird Species identification, so we decided to develop Bird Species identifier which identify the species of the bird from an uploaded image. This system helps in removing the knowledge barrier and smooth the species identification process. As there many software which provides the information of bird but none of them provide the identification feature which we will provide.

Large scale, accurate bird recognition is essential for avian biodiversity conservation. It helps us quantify the impact of land use and land management on bird species and is fundamental for bird watchers, conservation organizations, park rangers, ecology consultants, and ornithologists all over the world.

The main objective of designing this system is to construct a framework with the required tools to overcome the errors faced in the current system.

In this research, we try to develop a fully automatic, robust deep neural learning method that is able to overcome these issues. We evaluated our method on the biggest publicly available dataset which contains over 11,788 recordings of 200 different species.

We plan to build software that will identify the species of the bird accurately as well as provide the summarised information of the identified species.

IV.   LITERATURE SURVEY

To produce an Optimal Identifier there are numerous techniques, Bird DNA Barcoding, Bird Audio Analysis, Bird Species Identification Based on Image Features, etc.

In 2005, a student from Guelph University Paul Hebert which is a Canadian taxonomist have published a concept of DNA barcoding[1], its used DNA fragment of the mitochondrial gene cytochrome c oxidase(COI) for discriminating the species which help the scanners to identify it. DNA barcoding is a technique that is used when the Ornithologist is not able to acquire the whole picture and is only able to acquire the DNA of the bird through hair or blood of the bird, etc. But

these techniques can't be used by the normal users they are not able to operate the scientific instrument. Therefore, this technology only applies to scientific research instead of by the common public.

Fig [1] DNA Barcoding process

Recently, numerous projects are been developed to automate the bird classification process which involved use of technique such as audio data instead of images. This technique has certain advantage over image such as it doesn't require the line of sight and each species have unique calls which can be used for identification. But this technique is not reliable as a bird may not emit sound at all for longer duration and it's also does not help to count number of bird accurately. To overcome such challenges, number of research is been undergoing on techniques such as image-based techniques and computer vision. Several researchers have also purposed methods such as to use motion features of bird's curvature and wingbeat frequency. Atanbori et al [14] have done extended research on this method.

Cheng et al [2] purpose the system which uses discriminative features for classifying the bird species based on parts of birds that uses a support vector machine along with Normal Bayes classifier. Another researcher Marini et al [9] proposed an approach to eliminate background elements using a color segmentation and compute normalized color histograms to extract feature vector for classification.

Fine-Grained Image Categorisation – Technique for discriminating fine-grained classes (such as animal species or plants and man-made objects) which can be divided into two main groups. The prior group of methods uses distinct visual cues from local parts which are obtained by the detection or using the segmentation method. The other group of methods focal point is on finding interclass label dependency with the help of a pre-defined hierarchical structure of labels or manually-annotated visual attributes. Performance is drastically improved with the help of convolutional neural networks (CNNs), but for the training of CNNs large dataset of images of high quality are required. Fine-grained classification of low-resolution images is a very challenging process. Peng et al [10] purposed the method of transforming the detailed texture of information in High-resolution images to Low-resolution images with the help of fine-tuning for boosting the accuracy of recognizing fine-grained objects in Low-resolution images. This technique has certain limitations such as it requires the High-resolution images for the training of a model which limits their generalized implementation.

A similar error is found in Wang's work [12]. Chevalier et al. [13] their CNN-based fine-grained classifier is designed with different resolution images, this model is adaptive to simple convolutional and fully-connected layers but it does not have super-resolution unique layers in their CNN classification network.

Convolutional Super-Resolution Layers – Yang et al. [14] grouped together various Super Resolution algorithms into four groups: edge-based methods, example-based methods, prediction model and image statistical methods. The state-of-the-art performance is achieved in Convolutional Neural Network as CNN has been adopted for super-resolution recently. Dong et al. [10] attempted to use convolutional neural networks for image super-resolution.

In their method, a deep mapping was done between low- and high-resolution images. An additional deconvolution layer was added to avoid general up-scaling of input to accelerating training and testing on CNN. Another researcher Kim et al. [15] uses a deep recursive layer to avoid additional weighting layers, which reduce the computing time in case of increasing network parameters. Another way is to learn the mapping between the Low-Resolution image and its residue with the LR and HR image for speeding the training in CNN in a deep network. Super resolution-specific Convolutional network have been good at improving the quality of the image.

Deep transfer learning – Deep learning real-world applications which are been developed are drastically increased in recent times. Deep learning is different from machine learning as it tries to learn high-level features from mass data. By using the unsupervised or semi-supervised feature learning, it automatically extracts the data features. Whereas a machine learning model requires the manual features design which is time-consuming and error-prone. Challenging part in deep learning is a data dependence. Deep learning requires a huge amount of data for training contrast to the machine learning model. As it needs to map the features in the data without human intervention. It's a linear pattern i.e bigger the scale of the model, amount of data it requires also increase. In deep learning, initial layers identify the high-level features in training data, whereas final layers contain the information to make the final decision
.

Transfer learning is an important development in machine learning and deep learning area. Its solve the data insufficiency problem for training purpose. Transfer learning purposes the transfer of knowledge from the source neural model to the user's model if the source model and user model have a similar domain. This has the positive impacts on many domains which have difficulty in gathering the required amount of data. The transfer learning model is illustrated in Fig 2.
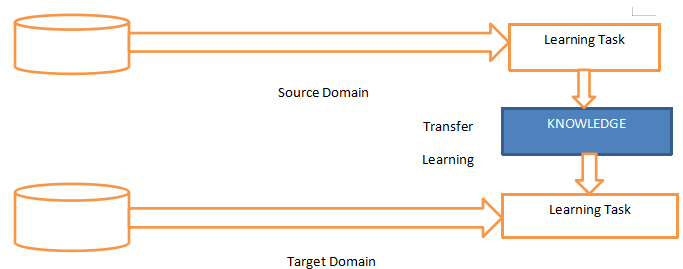

Fig [2] Transfer Learning

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NTASU - 2020 Conference Proceedings

Deep transfer learning is a method that helps to utilize the knowledge from similar model in the neural networks. Various deep learning method is been developed which can be classified and summarize in various categories such as instance-based deep transfer learning, network-based deep transfer learning, adversarial based deep transfer learning, and mapping-based deep transfer learning.

Instances-based transfer learning utilizes instances in the source domain by appropriate weight to transfer the required weight to the destination domain. Mapping-based transfer learning maps the instances from two domains into a new data space with better similarity. Network-based transfer learning reuses the partial network pre-trained in the source domain. Adversarial-based transfer learning uses adversarial technology to find transferable features that both suitable for two domains.

V. CURRENT SYSTEM

There are many species identifier systems available in the market but they consist of a lot many errors with less accuracy and are even costly. Current Systems available in the market are limited to local bird species and are dependent on the environmental factors but the system that we are developing can classify up to 200 bird species and it can be scaled up. Thus this system is a better approach for species identification with higher accuracy.

VI. PROPOSED SYSTEM

Avian ecosystem are recognized as useful biodiversity indicators. They are very sensitive and responsive to even little change in the ecosystems, as population's level changes are visible as well as quantifiable. Due to the wide variation in the species appearance, it is quite difficult to non-professionals to identify the species of the bird. It is also an exhausting process to annotate all the bird images with expertise human knowledge by appearance only. Therefore, there is a need to develop an automatic classifier for bird species, which will be a great convenience for many practical applications. For researchers who are working on the field can quickly capture the image of the bird and identify the species to which it belongs, eliminating the tedious process of illustration books and the use of other tools. This will also arouse interest in birds that could also benefit the protection of birds. Classifying the species of the bird is also an interesting problem in fine-grained categorization, which is also known as subcategory recognition, which is also a subfield of object recognition. The identifier in this paper proposed an end to end deep learning-based approach using transfer learning for the process of learning micro as well as macro-level features from bird ROIs. A pre-trained Mask-RCNN is been used to get the Bird ROIs from the images which will be fed to the neural model which is been developed using transfer learning technique and fine-tune using the available dataset.

VII. IMPLEMENTATION

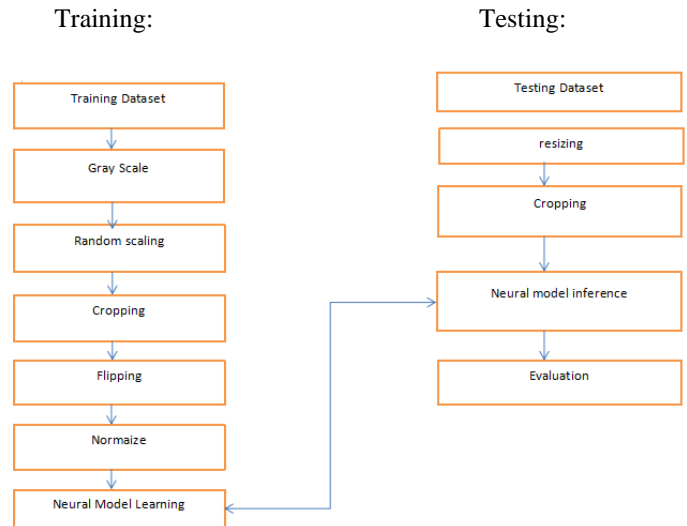Training: Testing:

Fig [3] 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, pre-trained 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)

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
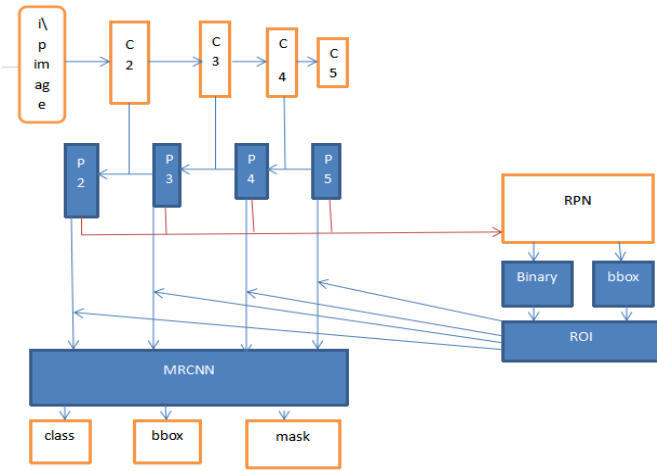**NTASU - 2020 Conference Proceedings**

Fig [4] Working of Mask R-CNN

Working of Mask R-CNN is divided into two stages. The first stage generates the guesses about the region where there is a higher probability of an object in the image. The second stage predicts the class of the object to which it belongs. Both stages are connected with the help of a pathway also known as the backbone which is FPN style deep neural network.

To create a neural model for the purpose of identifying the species of the bird instead of creating the neural model from scratch which is computing demanding as well as costly. We are using a transfer learning approach to learn both micro and macro-level features extracted from bird images for classification. We have used ImageNet pre-trained weights to initialize our Deepnet model for training. ImageNet contains 1.2 million images belonging to 1000 classes. Training using pre-trained ImageNet weights help us to learn fine-grained as well as global level features beforehand and learn the deep net more specific & discriminative features for each bird species which leads to increase the accuracy of our model.
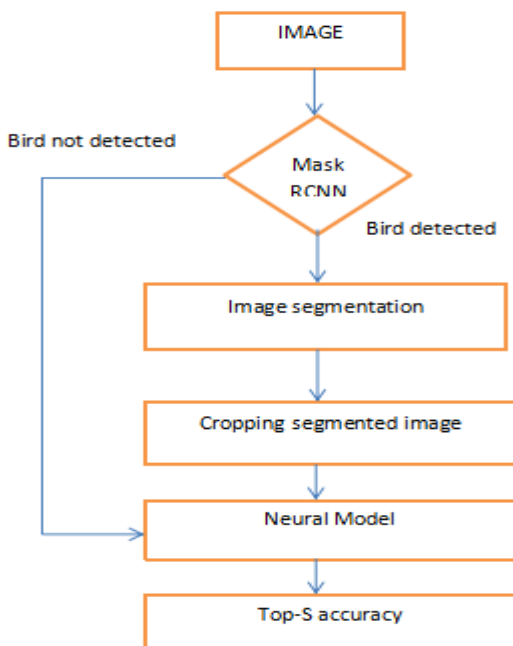

Fig [5] Bird species detection phase

Bird Detection: if the bird is detected by the Mask R-CNN algorithm in the image, then the rest of the image is removed except the segmented part. Then the cropped image is evaluated by the neural network. After evaluation, the prediction vectors are compared and the top 5 predicted species are displayed with their probability percentage.

No Bird Detection: if bird is not found in the image then the whole image is passed as an argument to the model then the model is evaluated and if the predicted values are less than 20% then it will prompt the user to input the valid image otherwise it will display the top 5 predicted species with their probability percentage.

Algorithm:
Step 1:
Users will be prompted to enter the image of the bird whose species need to be identified.

Step 2:
After the image is uploaded, the uploaded image is resized and gray scale.

Step 3:
A preprocessed image is passed to the Mask R-CNN algorithm to detect the Bird.

Step 4:
If the Bird is detected then other than segmented part of the image is removed, so as to reduce the burden on the neural model.

Step 5:
The cropped image is passed as an argument to the neural model for the inference purpose.

Step 6:
Top 5 accuracy is inferred from the model and a graph is formed representing the probability of each top 5 species.

Step 7:
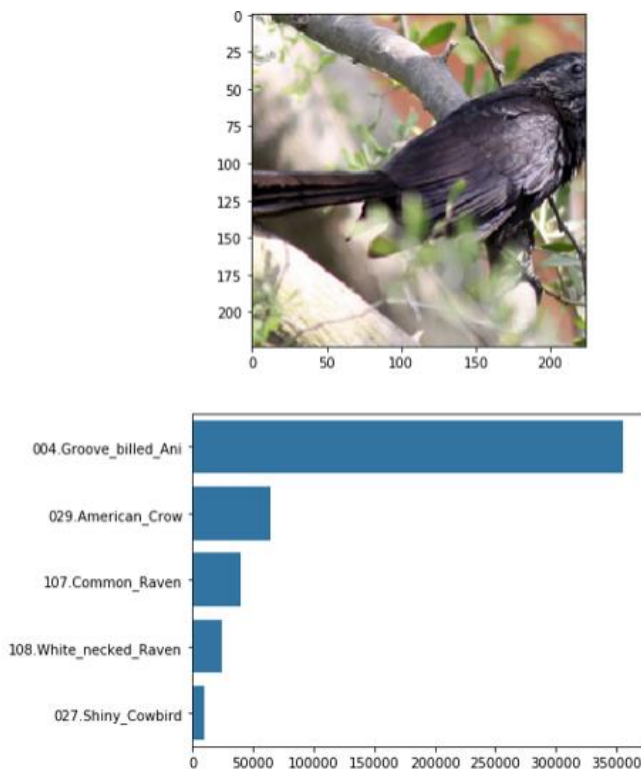Finally, the graph is displayed to the user.

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NTASU - 2020 Conference Proceedings

Fig [] Inference output of Bird.

VIII. LIMITATION

This system is capable of identifying the bird species from an image. As fine-grained processing is difficult, it becomes difficult to process when more than one species of bird is present in the provided image. There are also other reasons such as small bird ROIs, lighting conditions not suitable, the similarity in bird body parts and camouflage conditions. Such conditions are difficult to overcome in neural networks. Other limitations are to learn those discriminative features (both micro features like color, gradients, textures, etc. and macro features like shape, color patch, etc.). Various lighting conditions (like a picture taken during daylight, dawn, dusk, evening, etc.) affected our model most as due to low background light, many micro-features of the bird-like color, texture, gradients, etc. are lost.

IX. CONCLUSION

In this paper, we have proposed a method to localize as well as classify the species of the bird from an image that is uploaded by the users by using techniques such as mask R-CNN, transfer learning, and Convolutional Neural Network. The transfer learning technique helps to reduce the need for huge computing power for processing as well as speed up the learning process by reutilizing the knowledge.

XI. REFERENCE

[1] Hebeft PDN,Stoeckle MY, Zemlak TS and Francis CM, "Identification of birds through DNA barcodes [J]", PLoS Biol, 2(10): 1657-l663, 2004.

[2] Wang, Hsien Chang, Y. S. Chen, and M. Y. Wu. "A user-augmented ob-ject query system using color and shape features for Taiwan wild birds photos", International Conference on Machine Learning and Cybernet-ics, ICMLC 2010, Qingdao, China, July 11-14, 2010,Proceedings 2010, pp.2516-2520..

[3] LI Jian, ZHANG Lei, YAN Baoping, "Research and Application of Bird Species Identification Algorithm Based on Image Features", 2014 Inter-national Symposium on Computer, Consumer and Control,2014.

[4] Baowen Qiao, Zuofeng Zhou, Hongtao Yang, Jianzhong Cao," Bird Species Recognition based on SVM Classifier and Decision Tree", 2016, vol. 22, no. 1, pp.9-13.

[5] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connec-tions on Learning", arXiv:1602.07261v2 [cs.CV] 23 Aug 2016..

[6] Kaiming He Georgia Gkioxari Piotr Dollar Ross Girshick, "Mask RCNN", arXiv:1703.06870v3 [cs.CV] 24 Jan 2018.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", In NIPS, 2015.

[8] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A.Fathi, I.Fischer, Z.Wojna, Y.Song, S.Guadarrama, etal, "Speed/accuracy trade-offs for modern convolutional object detectors", In CVPR, 2017..

[9] Marini, A.A. Marini, A. J. Turatti, A. S. Britto Jr., and A. L. Koerich,"Visual and acoustic identification of bird species", IEEE International Conference on Acoustics, Speech and Signal Processing, 2015, pp. 2309-2313

[10] C. Dong, C. C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, IEEE transactions on Pattern Analysis and Machine Intelligence 38 (2) (2016) 295–307

[11] X. Peng, J. Hoffman, X. Y. Stella, K. Saenko, Fine-to-coarse knowledgetransfer for low-res image classification, in: IEEE International Conference of Image Processing, 2016, pp. 3683–3687.

[12] Z. Wang, S. Chang, Y. Yang, D. Liu, T. S. Huang, Studying very low resolution recognition using deep networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4792–4800

[13] M. Chevalier, N. Thome, M. Cord, J. Fournier, G. Henaff, E. Dusch, LR-CNN for fine-grained classification with varying resolution, in: IEEE International Conference of Image Processing, 2015, pp. 3101–3105

[14] C.-Y. Yang, C. Ma, M.-H. Yang, Single-image super-resolution: a benchmark, in: European Conference on Computer Vision, 2014, pp. 372–386.

[15] J. Kim, J. Kwon Lee, K. Mu Lee, Deeply-recursive convolutional network for image super-resolution, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1637–1645

[16] Atanbori, John & Duan, Wenting & Murray, John & Appiah, Kofi Dickinson, Patrick. (2016). Automatic classification of flying bird species using computer vision techniques. Pattern Recognition Letters. 81.53-62.https://doi.org/10.1016/j.patrec.2015.08.015