BIRD DETECTION AND SPECIES CLASSIFICATION USING MASK R-CNN

RENUKA M AVERI[1], PROF. REKHA B S[2]

[1] *Student, Dept. of Information science and engineering RVCE, Bengaluru.*

[2] *Professor, Dept. of Information science and engineering RVCE, Bengaluru.*

Abstract– In this paper, presents a keenly fundamental, adaptable, and general development for an object occasion division. The methodology competently perceives birds in a picture while at the same time making an uncommon division cover for every occasion. In this paper, showing top outcomes in all of the three tracks of the COCO set-up of difficulties, including case division; bounding box Bird disclosure, without stunts, Mask R-CNN destroy the flow, single-model zones on each attempt, including the COCO challenge victors. Accept immediate and productive procedure will fill in as a strong check and help ease future examination in occasion level certification.

Keywords: Deep Learning, Mask RCNN.

1. Introduction

People visit bird refuges to see various birds and to appreciate the value of the phenomenal collections of shades and characteristics of the birds. People rarely have data about the various species and thus can only see the credits and the species name with enormous exertion without the proclivity in the field of ornithology. Bird watching is everything viewed as saw as a decent wearing improvement that an enormous segment of people does practice other than their standard lifestyle. The change seeing attestation and approach of birds by using the overall man-made thinking and AI raises the improvement of the proposed model. A computer duplicates the working of human frontal cortexes in data treatment and provides a dynamic depiction of data. This paper presents a perspective of the convolution neural connection model for seeing such birds. The Mask R-CNN based portrayal model that engineers the bird species given a bird picture as data has been routinely applied to keep away the visual imagery and image classification normally recommends assuming the liability of an image and assuming it to some particular class. Here presented the Mask R-CNN-based portrayal model that engineers the bird species given a bird picture as data. The convolution neural alliance model is ready to clear out the grouping features that are subject to measurement, shape, and stowing away from the images, making it suitable for sales.

2. Literature Survey

1. Pralhad Gavali, J.Saira Banu [1] The system shown comprises of four stages: (1) an input/upload image, (2) pre-processing, (3) deep learning, and (4) classification. On the Google Net framework, a deep convolutional neural network (DCNN) may now be used to classify bird species. E stands for evaluation. The application of deep learning techniques to solve the problem of bird species identification was the subject of this thesis.

2. Kazi Md Ragib, Raisa Taraman Shithi, Shihab Ali Haq, Md Hasan, Kazi Mohammed Sakib, Tanjila Farah [2] Proposed Detection, Convolutional Neural Network recently found outstanding achievement in the entire image classification field. Using pre-trained Convolutional Neural Networks, a significantly better example of an input image has recently been produced (CNN).

3. Seda Bayat, Gültekin [3] the bird species that were seen in the Idr Aras River Bird Sanctuary were described using convolutionary neural networks in this research. To investigate and interpret biological diversity, acoustic monitoring is conducted. For this reason, passive acoustic recorders are used. In general, the raw sound recordings obtained by these recorders are analyzed.

4. Satyam Raj, Saiaditya Garyali, Sanu Kumar [4] constructed a Deep Learning model to enable bird watchers distinguish 60 species of birds in their natural environment to provide an easy technique for them to identify birds in their natural surroundings. To extract information from bird pictures, this model was developed using the Convolutional Neural Network (CNN) technique. Creating the CNN architecture, as well as collecting and locating Bird Dataset.

5. Roeland T'Jampens1, Francisco Hernandez1, Florian Vandecasteele2, and Steven Verstockt [5] Birds are warm-blooded vertebrates that belong to the Aves class. They have a wide range of traits and appearances, and there are about 10,000 living species of birds on the planet. Humans find the bird watching to be a fascinating hobby in the natural world as well. This paper presents an automated model that automatically identifies the species of a bird provided as the test data set based on deep neural networks. Using the test image datasets, the creation model was successfully checked and the accuracy of the constructed model was found to be using the test image datasets and the accuracy of the model built was found to be 98.75% overall.

3. OBJECTIVES

The objective of the project is:

- To develop an automated model to detect and classify the bird species to a particular class of its species.
- To make efficient use of Mask R-CNN to detect Bird.
- Classification of Birds with high accuracy with the help of Mask R-CNN.

4. PROBLEM STATEMENT

The primary aim of the proposed work is to develop an automated model which has a capability of detecting and classifying the species of the bird where bird image is given as a test image from the dataset.

6. METHODOLOGY

The following is the process used to create this model:

6.1 IMAGE ACQUISITION.

Images of 20 different bird species(Crow, Purple Finch, Peacock, Myna, Barn Swallow, American Avocet, Dove, Duck, Flamingo, Skimmer, Rhea, Lark, an Ostrich, Parakeet, Penguin, Kiwi, Hawk, Wren, Swan, Sparrow) were gathered from internet sources, and they were used for deep neural network model training and testing.

6.2DATASET

A total of 265 bird species are represented in this data set. 36609 images were used for training, 1325 images were used for testing, and 1325 images were used for validation. All images are JPG files with a resolution of 224*224*3 colours. The train set is 70% and the test set is 30%. Each set contains 265 subdirectories. If utilizes Keras, the data structure is really convenient.

6.3 DATASET PREPROCESSING

The image datasets must be preprocessed before they can be used to train the model. The picture files are converted to image pixel arrays using CV2 libraries, after all of the images have been scaled to the same picture size ratio (512*512).

6.4SYSTEM DESIGN ARCHITECTURE

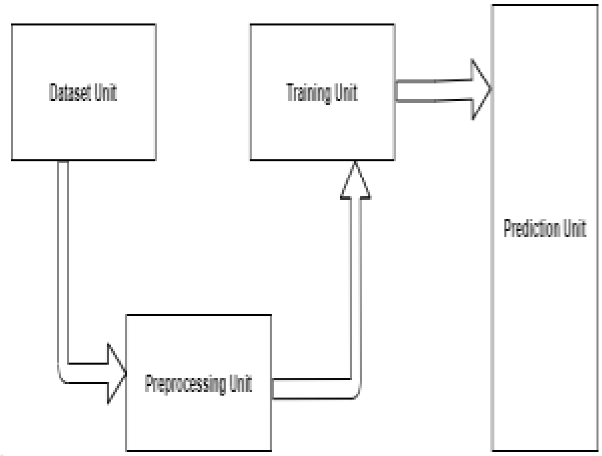

Figure 6.4: System design Architecture

1. Dataset Unit: It is used to import Birds species dataset to endeavour an examining the bird pictures.

2. Preprocessing Unit: It is used to pre-process image data and concentrate picture features of Bird species dataset.

3. Training Unit: Bird species dataset is used to a configuration using Mask R-CNN appraisal.

4. Prediction Unit: It predicts the Bird species at the last unforeseen development.

6.5 IMPLEMENTATION

Figure 6.5 illustrates the Mask R-CNN Architecture and implementation is explained below.

- Mask RCNN has two phases: the first scans the image and generates a proposal, while the second classifies the proposal
  and generates the bounding box and mask.

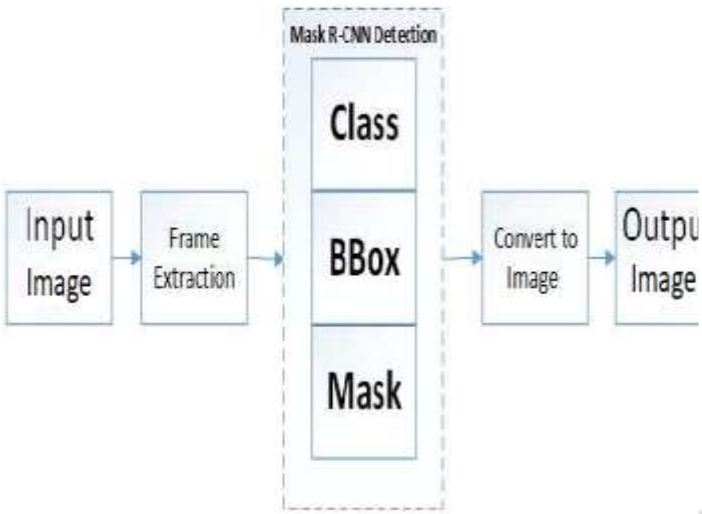

Figure 6.5: Mask R-CNN Architecture

- TensorFlow is used to implement the system, with RPN serving as the backbone network.
- We resize the photos to 512x512 pixels before training.
- Data augmentation utilizing horizontal flipping the images during the training period is conducted Mask RCNN's demand for more training data has been reduced.
- Bounding box regression, classification, and mask prediction are all activities that are carried out.

6.5.1 BACKBONE MODEL

Mask R-CNN extracts features from pictures using ResNet 101 architecture. The first step in extracting features from an image is to use ResNet 101 architecture. These characteristics are used as a source of data for the subsequent layer.

ResNet 50 is 50 layers deep Convolutional neural network. ResNet 101 Provides better average accuracies and confusion matrix than ResNet 50. As ResNet 101 is a 101 layers deep Convolutional neural network and has 5 Convolution layers. Whereas the ResNet 50 has 3 Convolution layers.

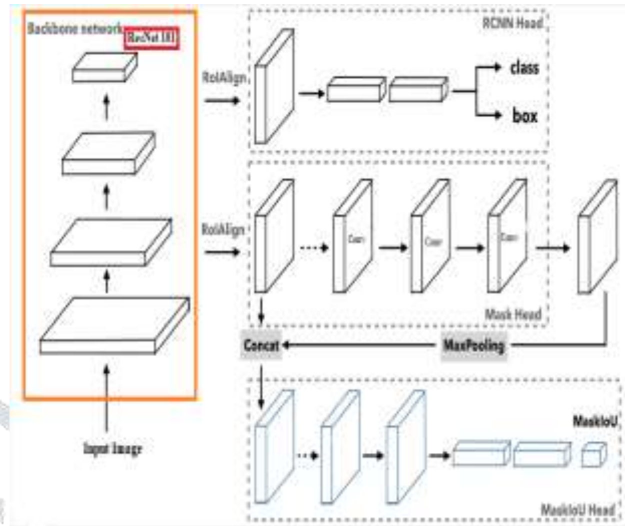

Figure 6.5.1: ResNet 101 Layers

6.5.2 REGION PROPOSAL NETWORK

Apply a region proposal network to feature maps created in the previous phase. This successfully predicts whether the object will be discovered in that location. Obtain the regions or feature maps that the model predicts will include some object at this step.

6.5.3 THE REGION OF INTEREST (ROI)

All of the regions are changed to the same shape using a pooling layer. The class label and bounding boxes are anticipated after running these regions through a fully connected network. Up to this point, the techniques are nearly identical to how Faster R-CNN works. Mask R-CNN also creates the segmentation mask.

6.5.4 SEGMENTATION MASK

This is the final stage of Mask R-CNN, where the masks for all of the objects in the image are predicted. Using the Mask R-CNN model's pre-trained weights from the COCO dataset. The Mask R-CNN model is then used to perform instance segmentation. This returns the segmentation mask for each object-containing section.

6.6 THE FLOW DIAGRAM

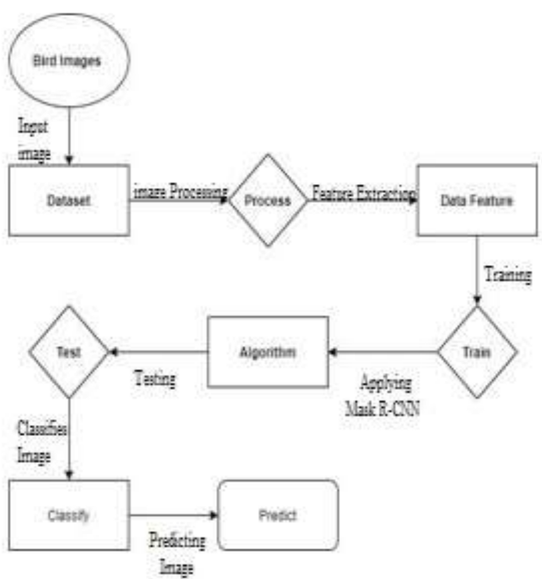

Figure 6.6: The Proposed system block diagram

The Birds species image dataset is processed. It is used to pre-process the image data and extract image features of the Bird species dataset. The Bird species dataset is used to train using the Mask R-CNN algorithm.

It classifies the Bird species. It predicts the Bird species at the final step as shown in Figure 6.6.

7. THE FUTURE WORK

The Other machine learning approaches may be included in future work, and a thorough comparison of them will be provided. Furthermore, increasing the accuracy of the prediction model by including more software metrics in the learning process is a viable option.

8. RESULTS

Examining how often a Mask R-CNN model uses pre-trained weights to split the features in an image. To visualise the model's predictions, start with the predictions and then plot the results:

Figure 8: Snapshot of Masked bird.

The above output shows bird detection and the classification and bird mask detection of the bird species name (Barn Swallow) and accuracy (93.46%) as shown in Figure 8.

8.1 CONFUSION MATRIX

The confusion matrix visualizes a classifier's accuracy by comparing the actual and predicted classifications.

Figure 8.1 illustrates the confusion matrix of the project. Considering 7 Classes of bird species (BG, Crow, Purple Finch, Peacock, Myna, Barn Swallow, and American Avocet). The accuracy is calculated of the individual bird. I.e. Here Class A is Background class. Class B is of Crow. Total 12 actual images of crow have been taken in which 11 images are predicted as a crow and the 1 image is predicted as other. So the accuracy of Class B is 91.67%

The confusion matrix is as follows:

Figure 8.1: Snapshot of Confusion matrix

- A predicted class is linked to each row of the matrix.
- Each row of the matrix represents a different class.
- The table shows the total number of correct and incorrect classifications.
- The predicted column and expected row of class value are filled with the sum of correct predictions for that class.
- The predicted column for that specific class value and the expected row for that class value both are filled with the sums of incorrect predictions for that class.

8.2 GRAPH

The Mask R-CNN loss function combines classification, regression loss, localization, and segmentation mask. The confidence score of a true class is used to calculate classification loss values. Therefore, they indicate how close the model is to correctly predict the proper class. The Mask R-CNN class loss algorithm covers all object classes, but the RPN class loss algorithm just labels anchor boxes as foreground or background.

The Snapshots of weight loss Graphs are as follows:

Figure 8.2.a: Snapshot of rpn_class_loss

Figure 8.2.b: Snapshot of rpn_bbox_loss

Figure 8.2.c: Snapshot of Mask R-CNN_class_loss

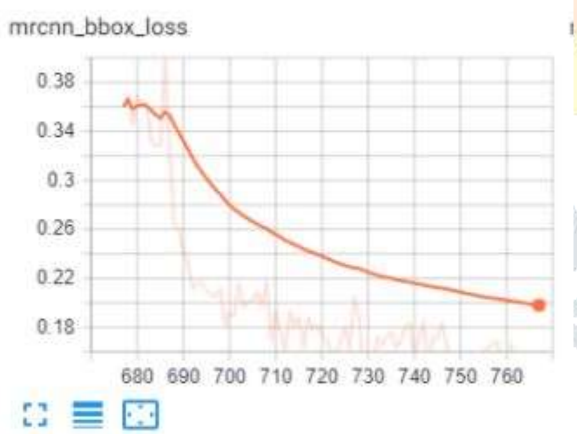

Figure 8.2.d: Snapshot of Mask R-CNN_bbox_loss

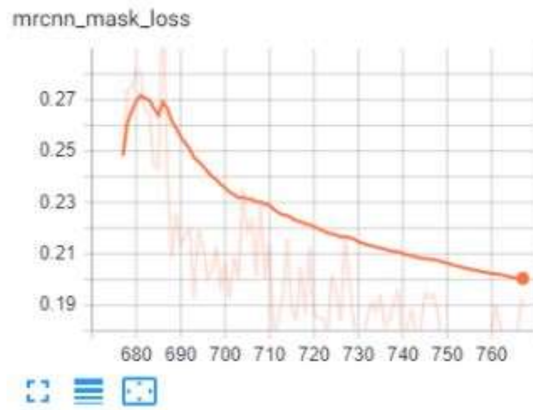

Figure 8.2.e: Snapshot of Mask R-CNN_mask_loss

- The bounding box loss values show the difference between the true and predicted box parameters, which are the (x, y) coordinates of the box location, its width, and its height. Figure 8.2.b and Fig 8.2.c illustrates the bounding box loss.
- Similar to the classification loss, the mask loss penalizes incorrect per-pixel binary classifications. For each of the regions of interest, it's calculated differently: Mask R-CNN encodes a binary mask per class for each of the RoIs, and the mask loss for each RoI is calculated only using the mask matching to its true class, preventing class predictions from influencing the mask loss. Figure 8.2.e illustrates the mask loss
- Since the genuine class's confidence score influence the classification loss values, the classification losses measure the model's confidence in predicting class labels, or how close it is to predicting the correct class. Figure 8.2.a and Figure 8.2.b illustrates the classification loss.

Loss weight:

rpn_class_loss	1.0
rpn_bbox_loss:	1.0
Mask R-CNN_class_loss	1.0
Mask R-CNN_bbox_loss	1.0
Mask R-CNN_mask_loss	1.0

Table 8.2: losses encountered while training

Starting with a base model with a default weight of 1 as shown in Table 8.2, visualize the model performance on various images and look at the number of objects detected, object classification accuracy, localization of identified objects, and mask localization is an ideal approach for tuning the loss weight of Mask R-CNN.

9. CONCLUSION

In this paper, developed an automated model by making use of train and test colored images of birds to identify/ classify the bird species to a particular class of its species using the Mask R-CNN algorithm.

The project is structured modeling-based and capable of delivering the intended outcomes. With some tweaks, it can be successfully implemented as a Real-Time system.

10. REFERENCES

[1] Pralhad Gavali, J.SairaBanu, "Bird Species Identification using Deep Learning on GPU platform". Conference: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020.

[2] Kazi Md Ragib, Raisa Taraman Shithi, Shihab Ali Haq, Md Hasan, Kazi Mohammed Sakib, Tanjila Farah," Automatic Bird Species Identification Using Deep Learning". 978-1-7281-6823-4/20/$31.00 c 2020 IEEE, 2020.

[3] SedaBayat, Gültekin," Identification of Aras Birds with Convolutional Neural Networks", Conference: 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2020.

[4] Satyam Raj, Saiaditya Garyali , Sanu Kumar, "Image based Bird Species Identification using Convolutional Neural Network". International Journal of Engineering Research & Technology (IJERT), 2020.

[8] Roeland T'Jampens1 , Francisco Hernandez1 , Florian Vandecasteele2 and Steven Verstockt," Automatic detection, tracking and counting of birds in marine video content". 978-1-4673-8910-5/16/$31.00 ©2016 IEEE.

[6]Jingyu Liu, Youngzhen Huang, JunranPeng, Jun Yao, Liang Wang (2018). Fast object detection at constrained energy. 2018 Digital object identifier 10.1109/TETC.2016.25775.38

[7]Rajaram Anantharaman, Matthew Velazquez, Yugyung Lee. (2018). Utilizing Mask RCNN for detection and segmentation of Oral Diseases. 2018 IEEE International Pattern Recognition Letters, 81, 53–62." (2016) Conference on Bioinformatics and Biomedicine (BIBM) (2018)

[8]KVV Subash, MRV Siddartha, Praneeth Akkala, Mr. M VenkataSrinu, NC Sri Harsha,

"Object Detection using Ryzetello drone with help of Mask RCNN" Proceedings of the second international conference on innovative mechanisms for industry applications(ICIMIA 2020).

[9]Xiaobo Yang, Aibinchen, Guoxiongzhou, Jianwuwang, Wenjiechen, Yuan Gao, Rundongjiang. (2020). "Instance Segmentation and classification method for plant leaf images based on ISC-MASK R-CNN and DCCNN" (2020)

[10] Zhijian Zhou, Meng Zhang, Jiefuchen, Xuqingwu. (2020). Detection and classification of multi-magnetic targets using Mask-RCNN (2020).