IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BIRD IDENTIFICATION USING MINIMAL SAMPLE

[1]Love Kumar Yadav, [2]P.padmanabhan
[1]Student, [2]Professor
[1]Galgotias University,
[2]Galgotias University

***Abstract*—There are thousands of categories of birds, it is very difficult to identify the birds by human and computer because different variation of birds to control and analysis. We need to identify Birds that weigh over 1.8 pounds (0.816 kg) because these birds strike can cause lots of damage to the aircraft. Bird strikes happen most often during LANDING, APPROACH, INITIAL ASCENT, TAKE OFF. In this research we have used Caltech-UCSD Birds-200-2011 Dataset Caltech-UCSD Birds-200-2011 (CUB-200-2011) is an extended version of the CUB-200 dataset. In this dataset there are 200 bird species categories, in this research project we create handled device (android application). In this handled device we use transfer learning to train this model. We try different model to get better accuracy like VGG16, VGG19, MobileNet, ResNet50 and Inception etc.**

***Index Terms:* bird species, Caltech-UCSD Birds-200-2011, Transfer learning, VGG16, VGG19, MobileNet, ResNet, Inception**

I. INTRODUCTION

Bird strikes by aircraft need to be addressed quickly for safety. Identification of the bird is critical to determine its size/weight and to determine if more testing for damage is required and what action to take. Currently, birds can be identified by DNA and feather identification, but the methods are time consuming (weeks) including shipping samples back and forth. A faster method for bird identification, preferably with a handheld device is required [2].

1. The accident between bird and aircraft cost millions of dollars per year and it is significant threads to plan safety and caused hundreds of human casualties.
2. The term is also used for bird deaths resulting from collisions with structures such as power lines, towers and wind turbines.
3. Bird strikes are a significant threat to flight safety, and have caused a number of accidents with human casualties. There are over 11,500 bird strikes per year.
4. The Geese have been ranked as the third most danger(hazardous) bird category. Most accidents occur when birds collide with the windscreen or is sucked into the engines of aircraft.
5. hese caused millions of dollars, sometimes whole engine needs to be change
6. Bird identification is really a difficult task because there is lots of different categories of birds who look similar.

Therefore, we built an android application which identify a bird in a certain distance to avoid bird strike and we can find out that which category bird belong to so that next time we avoid collision.

II. Related Work

Jaderberg et al. claim that they can achieve 78.3% with the Inception-V2 architecture and Krause et al. claim that they can get 80.4% accuracy with the Inception-V3 architecture [2]. IN this research we have used Caltech-UCSD Birds (200-2011) dataset. Caltech-UCSD Birds (200-2011) is an extended version of the CUB-200 dataset. In this dataset there are 200 Bird Species CATEGORIES. In this research project we create handled device (Android Application) to identify the images of the Birds. In this handled device we use transfer learning to train this model. We try different model to get better accuracy like VGG16, VGG19, MobileNet, ResNet50 and Inception etc. Through this model we get the 83.0% accuracy.

We basically work on different approach like: -

a) Landing
b) Initial Ascent
c) Take off of Aeroplan

We also work on classification of dataset such as it automatically detects-

a) Birds or Not Bird
b) If Bird then Check Weight; 1.8 Pounds

At airport we use these following things-

a) Speed- LR
b) Drones
c) 360* Cameras

There are number of ways to classify bird using audio data rather images. For this Purpose, Feature Extraction from audio data have some advantages like bird category have distinctive calls and no line of sight is need for detection. There are some disadvantages of this method, because some birds may not produce any sound at all for a particular time and we also not able to count the number of birds precisely.

IJCRT

To solve the problem of bird strike we must identify the number of birds approaching to airdrome so that we can delay the flights to avoid bird strike.

IV. Methodology

The very first step for training a model is to collect data. We used Caltech-UCSD Birds-200-2011 dataset [2] for training our model.
The Caltech-UCSD Birds-200-2011 dataset has 200 categories of birds which include 11,788 total images around 65-75 images per category of birds.
As transfer learning is the most used methodology in Deep learning. We used MobileNet V2 architecture for training our model. MobileNet V2 is most suitable for mobile and embedded devices-based vision applications where there is a lack of computational power. The MobileNet V2 is high speed, good performing and low maintenance deep learning architecture.

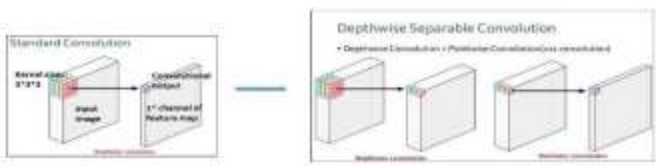


Fig. 1. Standard and Depthwise Separable Convolution [1]

Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5× Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
5× Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

Fig. 2. Mobile Net Architecture [1]

Implementation of MobileNet V2:

Create the base model from the MobileNet V2 model developed at Google, and pre-trained on the ImageNet dataset, a large dataset of 1.4M images and 1000 classes of web images.
First, pick which intermediate layer of MobileNet V2 will be used for feature extraction. A common practice is to use the output of the very last layer before the flatten operation, the so-called "bottleneck layer". The reasoning here is that the following fully-connected layers will be too specialized to the task the network was trained on, and thus the features learned by these layers won't be very useful for a new task. The bottleneck features, however, retain much generality. Let's instantiate an MobileNet V2 model pre-loaded with weights trained on ImageNet. By specifying the include top=False argument, we load a network that doesn't include the classification layers at the top, which is ideal for feature extraction [1]

```
[ ] IMG_SHAPE = (IMAGE_SIZE, IMAGE_SIZE, 3)

    # Create the base model from the pre-trained model MobileNet V2
    base_model = tf.keras.applications.MobileNetV2(input_shape=IMG_SHAPE,
                                                   include_top=False,
                                                   weights='imagenet')
```

Fig. 3. Using ImageNet

After training our model with the help of processed dataset we saved the model to a HDF5 file, for further processing we used TensorFlow Lite which is an open source deep learning framework for on-device inference by Google. And we converted our saved model to tflite (compressed flat buffer model) with the help of with the TensorFlow Lite Converter. We developed an Android app using TensorFlow Lite which can classify the category of bird using camera of mobile device in Real-time. Our android application takes less than 2 seconds to detect a bird and show the result with confidence score and index of bird category.

Fig. 4. Android App Demo

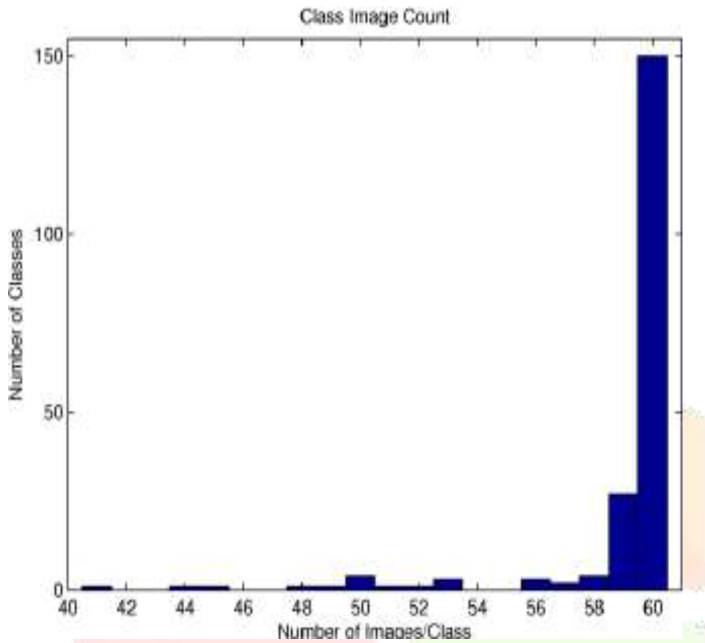
V. Experimental Results

Caltech-UCSD Birds-200-2011 (CUB-200-2011) is an extended version of the CUB-200 dataset, with roughly double the number of images per class and new part location annotations.
Caltech-UCSD Birds-200-2011 dataset has 200 classes of birds and 11,788 images.

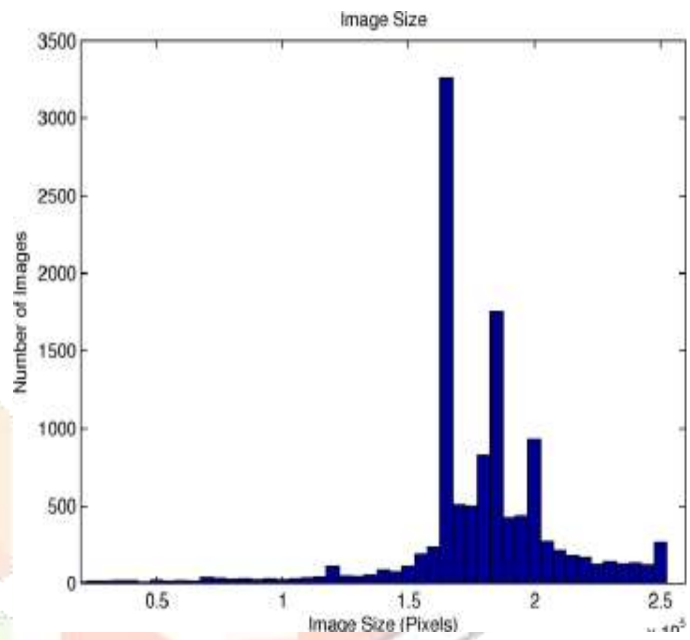
Dataset Statistics:

(a) Distribution of the number of images per class (most classes have 60 images).

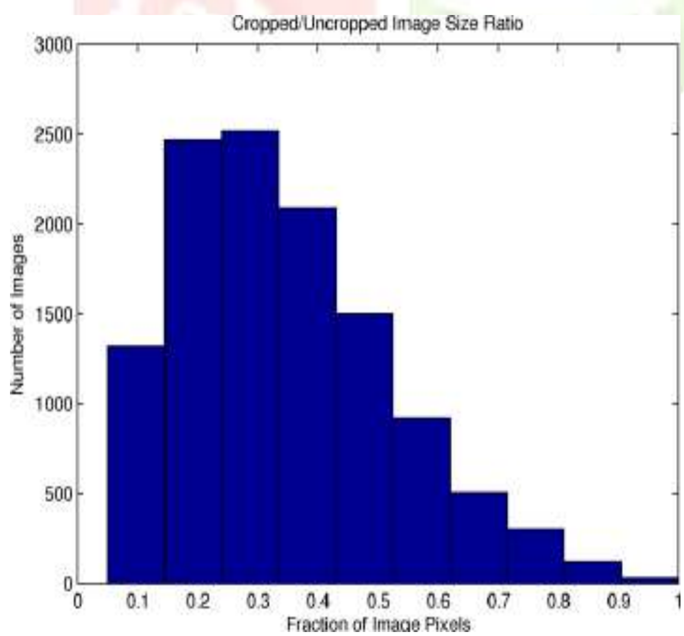
(b) Distribution of the size of each image in pixels (most images are roughly 500X500).



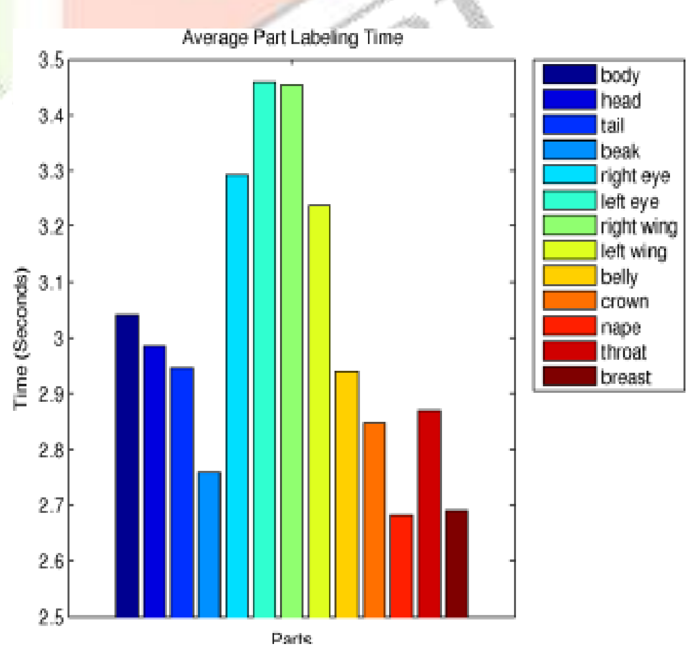
(a) Class Image Count



(b) Image Sizes



(c) Cropped/Uncropped Image Size Ratio



(d) Average Part Labeling Time

(c) The average amount of time it took MTurkers to label each part.
(d) Distribution of the ratio of the area of the bird's bounding box to the area of the entire image.

In this MobileNet model we resized images to 224 by 224 and then used for training and validation.

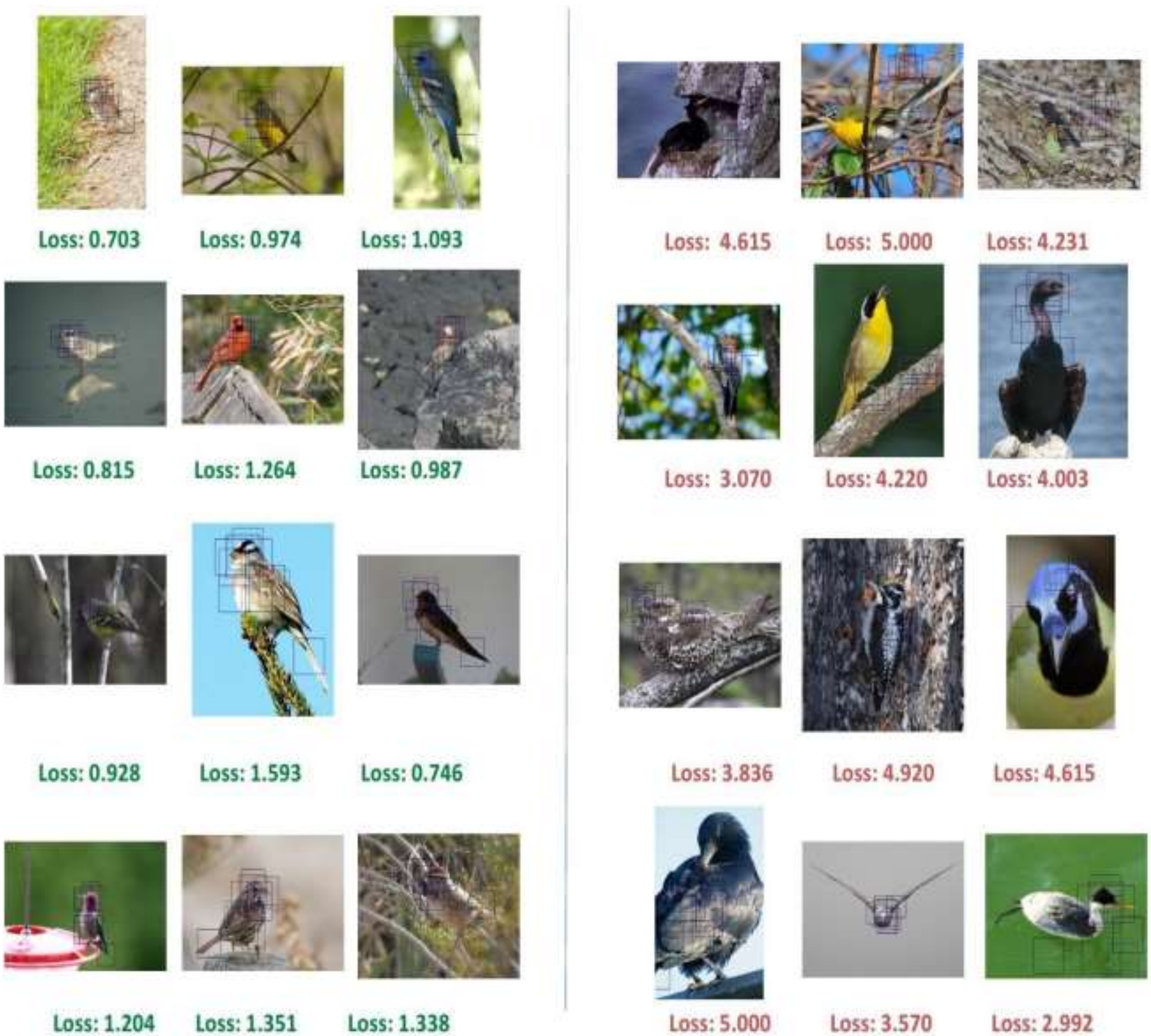


Fig. 6. Example Part Detection Results [3]

Example Part Detection Results, with good detection results on the left and bad detection results on the right. indicates that the predicted part locations are about as good as the average MTurk labeler.

Using only the full, uncropped bird images, assign each image to one of 200 bird classes. Since the images are uncropped, we anticipate that the problem cannot be solved with high accuracy without obtaining some degree of localization.

Fig. 7. 200 Categories of Birds [2]

On this dataset we use transfer learning and try with different model like VGG16, VGG19, ResNet50, and MobileNet etc.
As we do not have much data for individual species of bird so it is very difficult to get better validation accuracy after experimenting with different models, we found that with MobileNet V2 we had some better results for training accuracy and validation accuracy.

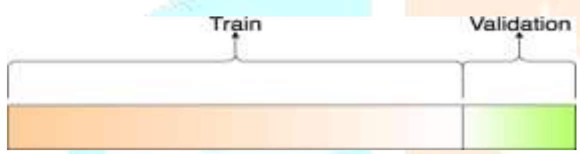


Fig. 8.Train and Validation Split [4]

Highest accuracy achieved 83% with MobileNet V2 after fine tuning our model.Learning Curves while freezing the base layers:

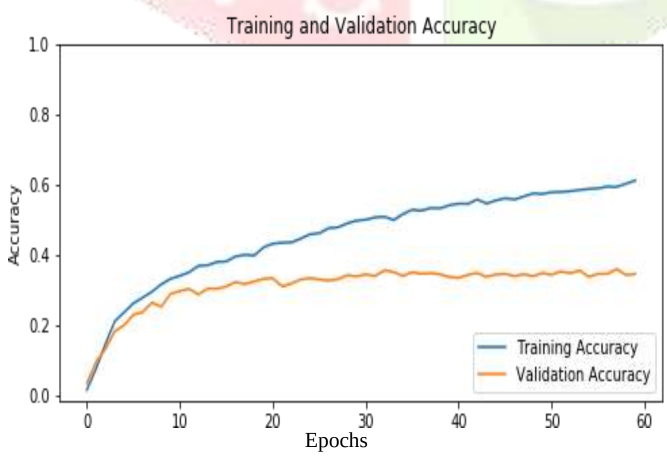


Fig. 9. Learning Curves

VI. Conclusion

Caltech-UCSD Birds-200-2011 dataset has large number of categories make it more interesting. With the help of Caltech-UCSD Birds-200-2011 we train a MobileNet V2 Model using transfer learning and save that model in a model.h5 file and convert it into tflite file and with the help of tflite file we develop an android application that can predict image category with the probability and with the help of this mobile application we can easily find out the category of the dead bird (bird strike with aircraft) so that we can easily get the information of the bird that which category birds are can for crash aircraft engine so that next time we can reduce the damage at the time of bird strike.

VII. Reference

[1] Z. Ge, A. Bewley, C. McCool, P. Cork e, B. Upcroft, and C. Sanderson, "Fine-grained classification via mixture of deep convolutional neural networks," in Proc. IEEE Winter Conf. Appl. Comput. Vis., 2016, pp. 1–6.

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

[3] T. Berg, J. Liu, S. W. Lee, M. L. Alexander, D. W. Jacobs, and P. N. Belhumeur, "Birdsnap: Large-scale fine-grained visual categorization of birds," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 2019–2026.

[4] A. Marini, J. Facon, and A. L. Koerich, "Bird species classification based on color features," in IEEE Int'l Conf Sys Man Cybernetics, Oct 2013, pp. 4336–4341.

[5] Branson S., Wah C., Babenko B., Schroff F., Welinder P., Perona P., Belongie S., "Visual Recognition with Humans in the Loop", European Conference on Computer Vision (ECCV), Heraklion, Crete, Sept., 2010.

[6] Welinder, P., Perona, P. Online crowdsourcing: rating annotators and obtaining cost-effective labels. Workshop on Advancing Computer Vision with Humans in the Loop at CVPR. 2010.

[7] Vedaldi, A., Gulshan, V., Varma, M., Zisserman, A.: Multiple kernels for object detection. In: ICCV. (2009)

[8] Holub, A., Perona, P., Burl, M.: Entropy-based active learning for object recognition. In: Workshop on Online Learning for Classification (OLC). (2008)

[9] P. Perona. Vision of a Visipedia. Proceedings of the IEEE, 98(8):1526 –1534, 2010

[10] M-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, Dec 2008