MobileNet Model for Classifying Local Birds of Bangladesh from Image Content Using Convolutional Neural Network

Conference Paper · July 2019

DOI: 10.1109/ICCCNT45670.2019.8944403

3 authors:

Md Romyull Islam
Kennesaw State University

Nishat Tasnim
Daffodil International University

Shaon Bhatta Shuvo
Daffodil International University

MobileNet Model for Classifying Local Birds of Bangladesh from Image Content Using Convolutional Neural Network

Md. Romyull Islam
Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
islam15-5833@diu.edu.bd

Nishat Tasnim
Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
tasnim15-5709@diu.edu.bd

Shaon Bhatta Shuvo
Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
shaon.cse@diu.edu.bd

***Abstract*—To classify bird species is quite a challenging task due to complex interdependence on various factors. There have been numerous attempts at perfecting classification. The aim of our work is to classify bird species from image data with a computer vision classification system. In this paper, we put forward a MobileNet model, which gives an amazing accuracy of up to 100%. This is the first work relating to local bird species classification. The proposed model explores a systematic approach to classification. The outcomes prove the efficiency of the model.**

***Keywords*—species, classification, computer vision, MobileNet, local**

I. Introduction

Bangladesh is the most loved country of a great variety of birds. Their sound, color, appearance, and behavior are more or less developed depending on the species. We are used to with the song of a bird when we go to bed at night and when wake up in the morning. They are found singing, chirping and flying up and down from dawn to dusk.

The process of classification is very much important for day-to-day life and the reason we use it in anything we do. It makes things simple to discover and recognize. Differentiation of objects is what allows us to distribute them into classes. People from thousands of years back have utilized the classification process. To classify objects first the objects must be analyzed to recognize differences. The classes must be unique; an object should not fit into more than one class. Choosing the most appropriate class is also essential to the achievement of the classification system. Every classification is helpful to the success of any kind of collection of objects, without it, no one will realize where to discover anything. A well thoroughly considered classification system will be helpful to clients of the collection and will assure that they proceed to use and enjoy the collection.

In this paper, transfer learning [2] technique was used to retrain the MobileNet [4] model of TensorFlow [1] on the dataset of five bird species, locally known as bok, doel, machranga, pecha, and tiya. Here Section 2 describes the background study with TensorFlow [1]. The main theme, the structure of MobileNet [4] was described in section 3. The scenario of the building of the classification model was provided in section 4. Experimentation on five different birds' datasets was referred to in section 5. Relevant work was described in section 6. Section 7 terminates with the summary.

II. RELATED WORKS

No published research has been done on classification of local bird species of Bangladesh followed by the use of MobileNet [4] algorithm.

Nitin R. Gavai, Yashashree A. Jakhade et al. from SCOE, Pune used the transfer learning [2] technique to retrain the MobileNet [4] model of TensorFlow [1] on the flower category datasets of Oxford-102 for Flower Classification in 2017. The classification accuracy of the model was between 85% and 99% [4].

Andrew G. Howard, Menglong Zhu et al. used MobileNets [4] and CNN [3] for Mobile Vision Applications like object detection, face attributes, fine-grained recognition, large scale geolocalization etc. in 2017 [9].

Yiting Li, Haisong Huang et al. worked on Surface Defect Detection from the defect images on the sealing surface of the container in 2018. They used MobileNet-SSD [10], MobileNet [4] and VGG [11]. The classification precision of the model was 96.73% on MobileNet-SSD, 92.33% on MobileNet and 93.91% on VGG [12].

Sangita S.Londhe and Dr. Sudhir S. Kanade worked for Bird Species Identification based on the Support Vector Machine [13] model in 2015. They used the audio pattern as a dataset. SVM kernel function processed data accuracy in percentage. The precision was 59.591% for Linear, 84.591% for Quadratic, 84.119% for Radial Basis Function (RBF) and 79.402% for Polynomial [14].

Uma D. Nadimpalli, Randy R. Price et al. used the RNN [15] algorithm for Bird Recognition in 2006. All experiments were done on three different datasets of bird depending on image clarity. The accuracy of the datasets was 1:100%, 2:60% and 3:50% [16].

During winter, many migrated birds come to this area. This is an essential task for the people to differentiate between the local birds and the migrated birds to ensure their safety and find an easy way to identify them. As a result, they can preserve the habitat and food. Because of the lack of proper habitat and food, now they cannot stay here longer and are gradually being extinct. So people need to work more with these birds.

III. BACKGROUND

Google formed TensorFlow [1], the second machine-learning structure which is used to build, train and design deep learning models. We may utilize the TensorFlow library to do mathematical calculations, which in itself does not seem all too extraordinary, but with data flow graphs, these calculations are done. In the graphs, edges show data and nodes show numerical processes. Where neural networks operate with multidimensional data arrays or with tensors, the "TensorFlow" term is originated from such processes.

Transfer learning [2] techniques reliant on the machine learning algorithms being utilized to learn the functions, and may frequently be dealt with expansions of the algorithms. The aim of transfer learning strategy is to develop the learning rate with a goal by influencing information gathering from the basic task. The transfer might enhance learning through three ordinary measures. First is the basic performance obtainable in the goal task using only the transferred learning, earlier than any additional learning is completed, in comparison with the basic performance of an unfamiliar operator. Next is the quantity of time it takes to analyze the goal task given the transferred learning in comparison with the quantity of time to analyze it from scratch. The last extent of performance is obtainable to the goal in comparison with the last measure with no transfer.

The convolutional network is a particular type of deep neural network, which is referred to as CNN [3] or ConvNet. It is also known as multi-layer perceptrons. The models are named "feed-forward" as data flows right via the model. There are not any feedback links wherein outcomes of the model are fed back into itself. In the field of computer vision, convolutional neural networks have been one of the significant innovations. They have worked significantly more than conventional computer vision and have generated state-of-the-art outcomes. These neural networks have turned out to be fruitful in many real-life case studies and operations.

IV. STRUCTURE OF MOBILENET

We explained the basic structure layer of MobileNets [4] that is constructed on depthwise distinct filters in this part.

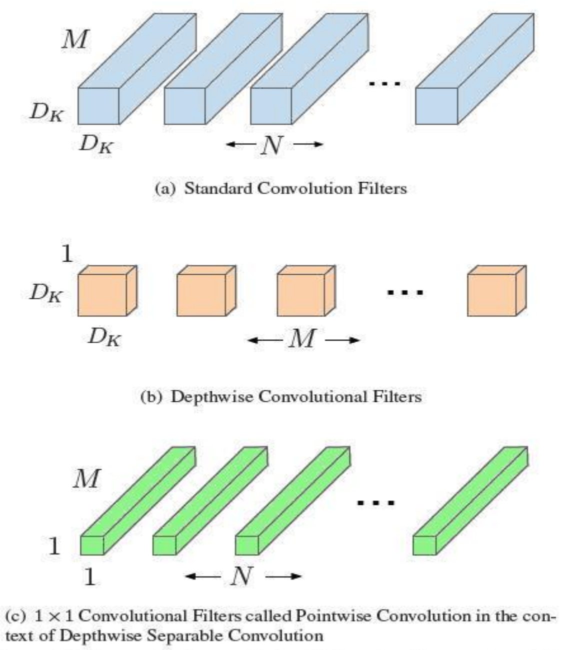


Fig. 1. The typically convoluted filters in (a) are exchanged by two layers: depth wise intricacy in (b) and pointwise intricacy (c) to construct a depth wise distinct filter [5].

In MobileNets the depthwise convolution is utilized which applies a different filter to every response channel. The MobileNet model is built on depthwise distinct convolutions, which occurs to be a shape of factorized intricacies, which factorizes a typical intricacy into a depthwise intricacy, and a 1x1 intricacy is labeled as pointwise intricacy. Afterward, the pointwise intricacy applies a 1x1 intricacy to summate the outcomes to the depthwise intricacy. In a typical intricacy, each is clarified as well as summate inputs to form a distinct outcome arranged in a single step.

The depthwise distinct intricacy splits this into dual layers, one for clarifying and the other for connecting. This partition has a vast impact of reduction with computational time and size of the model. Fig. 1 indicates various acts of how a regular intricacy 1(a) is factorized into a depthwise intricacy 1 (b) and a 1x1 pointwise intricacy 1(c).

V. CONSTRUCTION OF MODEL

Experimental setup for bird classification model using MobileNet [4] on TensorFlow [1] framework is described in this part.

The classification model is divided into these five phases: dataset, image preprocessing, training process, testing, and verification. In Fig. 2, our full working process is shown through a flowchart.

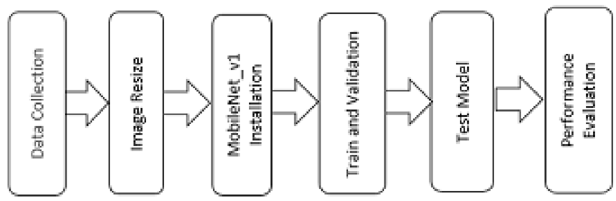
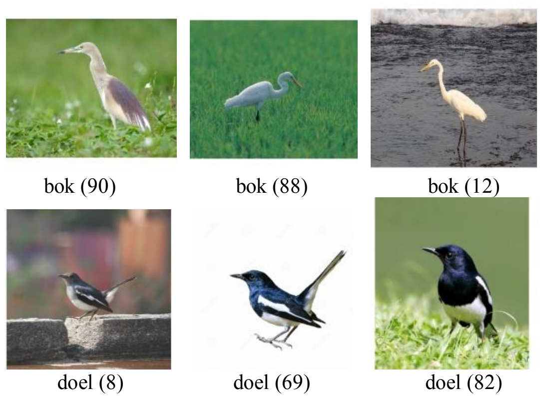


Fig. 2. Flowchart of the proposed model.

A. Dataset

We can identify approximately 10,000 divergent species of bird. Each of the species is non-identical from the others. The following images depict a portion of our dataset. This dataset holds five classes of birds, which was made by us; each label of species contains 100 jpg images.



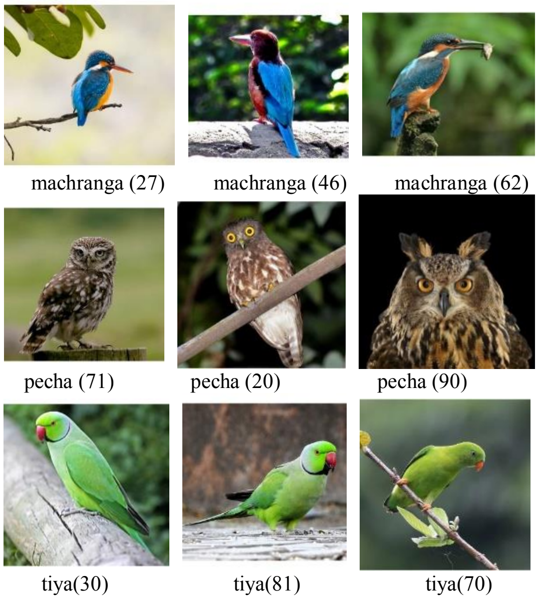


Fig. 3. Examples of our created dataset.

B. *Image-Preprocessing*

We have labeled and resized the image for training and testing in the image-preprocessing stage.

C. *Training Process*

Each image of bird dataset is utilized several times through the training procedure. Computing the layers behind the layer just before the last output layer that accomplishes the grouping for every image takes extensive time. As the inferior layers of the network are not being transformed, their results can be reserved and utilized again.

In Fig. 4, the main graph of MobileNet [4] model is demonstrated.

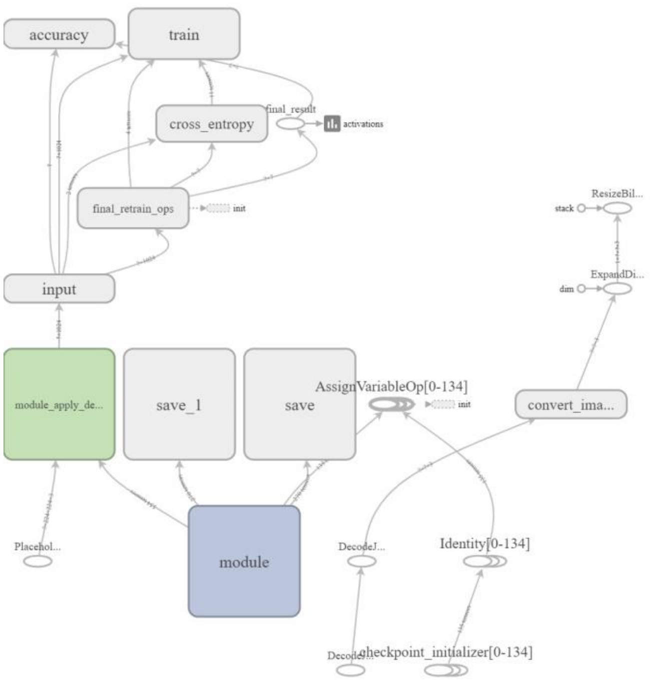
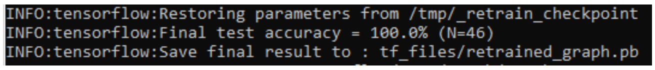


Fig. 4. Main graph of MobileNet [4] model.

D. *Verification and Testing Process*

After completing the training process of this model, we got 100% test accuracy. The image shows the exact scenario of the training process.



```
INFO:tensorflow:Restoring parameters from /tmp/_retrain_checkpoint
INFO:tensorflow:Final test accuracy = 100.0% (N=46)
INFO:tensorflow:Save final result to : tf_files/retrained_graph.pb
```

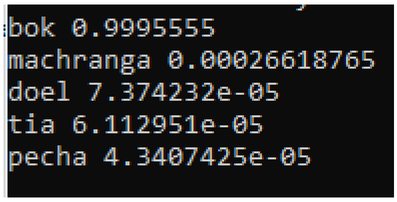
Fig. 5. Final test accuracy after training.

For the verification, we used "bird" dataset.

VI. EXPERIMENTS' RESULT

This analysis is established on python with TensorFlow [1] library and MobileNet [4] model.

After training this model entirely the images in the test set demonstrates the particular label.



```
bok 0.9995555
machranga 0.00026618765
doel 7.374232e-05
tia 6.112951e-05
pecha 4.3407425e-05
```

Fig. 6. Outcomes of MobileNets [4].

A sequence of outcome stages may be visible as it trains, everyone, displaying training accuracy [6], validation accuracy, and the cross-entropy [8]:

The training accuracy is the accuracy of a model on examples it was constructed on [6]. The validation set is utilized to evaluate the model's functioning [7]. Cross-entropy shows the distance between the outcome allocation and the original allocation [8].

The variations in accuracy established on cross-entropy [8] in our training dataset are demonstrated in Fig. 7 and Fig. 8 respectively. The orange line represents the training set and the blue line represents the validation set [7].

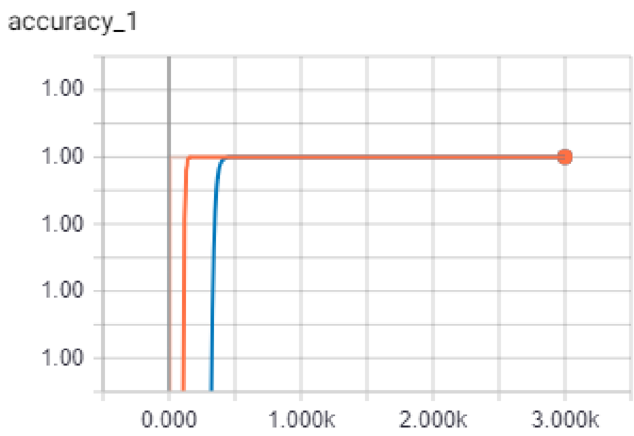


Fig. 7. Training accuracy [6].

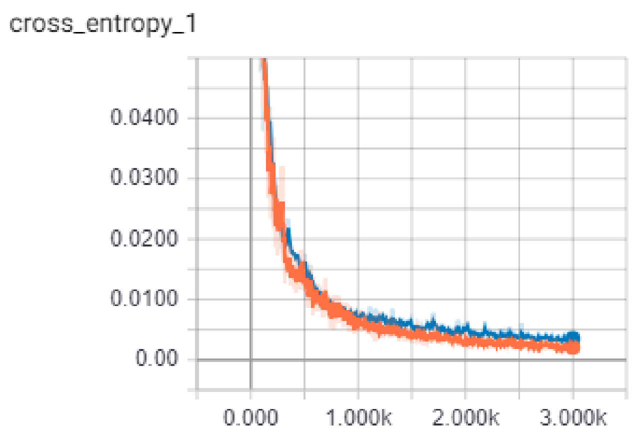


Fig. 8. Cross-validation graph.

The details of Fig. 7 and Fig. 8 are demonstrated in the following table.

TABLE I. DETAILS OF THE TWO FIGURES

Dataset	Index	Efficiency
Dataset	The precision of the training set	100%
	The precision of the validation set	100%
	The cross-entropy of the training set	0.0025
	The cross-entropy of the validation set	0.003 to 0.005

The training and the validation accuracy reached 100% for our dataset.

VII. CONCLUSION

With the advancement of science, people started to search the technique of automatic bird species recognition by computer. The main goal of our research is to establish a new strategy to recognize bird species automatically by utilizing new techniques of Computer Vision and bettering the accuracy.

Scarcity of perfect and diversified images is the main limitation of our paper. Our bird recognition algorithm has been tested on only five bird species, bok, doel, machranga, pecha, and tiya. Other species may be utilized in future studies. Future research will also aim at testing other recognition algorithms.

REFERENCES

[1] DataCamp Community. (2019). TensorFlow Tutorial For Beginners. [online] Available at: https://www.datacamp.com/community/tutorials/tensorflow-tutorial?. [Accessed 20 Mar. 2019].

[2] 2019. [Online]. Available: https://www.igi-global.com/chapter/transfer-learning/36988. [Accessed: 19- Mar- 2019].

[3] "Convolutional Neural Networks in Python", DataCamp Community, 2019. [Online]. Available: https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python?. [Accessed: 22- Mar- 2019].

[4] N. R. Gavai, Y. A. Jakhade, S. A. Tribhuvan, and R. Bhattad, “MobileNets for flower classification using TensorFlow,” In 2017 International Conference on Big Data, IoT and Data Science (BID), pp. 154-158, December 2017.

[5] "MG2033/MobileNet", GitHub, 2019. [Online]. Available: https://github.com/MG2033/MobileNet/blob/master/figures/dws.png?fbclid=IwAR2arODz3uTgdVCAp34OOXPJBJJSAxIXmsg5XjwbZ-fnu9YyEqnRxqzEYH4. [Accessed: 21- Mar- 2019].

[6] T. matrix, P. Alves and S. Balen, "Train Accuracy vs Test Accuracy vs Confusion matrix", Data Science Stack Exchange, 2019. [Online]. Available: https://datascience.stackexchange.com/questions/28426/train-accuracy-vs-test-accuracy-vs-confusion-matrix. [Accessed: 18- Mar- 2019].

[7] W. accuracy, "What is the difference between the terms accuracy and validation accuracy", Stack Overflow, 2019. [Online]. Available: https://stackoverflow.com/questions/51344839/what-is-the-difference-between-the-terms-accuracy-and-validation-accuracy. [Accessed: 30- Mar- 2019].

[8] P. Dahal, "Classification and Loss Evaluation - Softmax and Cross Entropy Loss", DeepNotes, 2019. [Online]. Available: https://deepnotes.io/softmax-crossentropy. [Accessed: 13- Mar- 2019].

[9] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, …, and H. Adam, “Mobilenets: Efficient convolutional neural networks for mobile vision applications,” April 2017.

[10] M. SSD, "Mobilenet vs SSD", Stack Overflow, 2019. [Online]. Available: https://stackoverflow.com/questions/49789001/mobilenet-vs-ssd. [Accessed: 18- Mar- 2019].

[11] "CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more ….", Medium, 2019. [Online]. Available: https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5. [Accessed: 19- Mar- 2019].

[12] Y. Li, H. Huang, Q. Xie, L. Yao, and Q. Chen, “Research on a Surface Defect Detection Algorithm Based on MobileNet-SSD,” Applied Sciences, 8(9), 1678, p.1678, September 2018.

[13] "Support-vector machine", En.wikipedia.org, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Support_vector_machine. [Accessed: 17- Mar- 2019].

[14] S.S. Londhe, and D.S. Kanade, “Bird Species Identification Using Support Vector Machine,” 2015.

[15] "Recurrent neural network", En.wikipedia.org, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Recurrent_neural_network. [Accessed: 18- Mar- 2019].

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