Plastic Egg Detection Using Machine Learning Approach

BY

Md. Palash Miah
ID: 181-15-10518

Md. Hasan Jamil Sany
ID: 181-15-10519

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering.

Supervised By
Mr. Riazur Rahman
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 06, 2022

APPROVAL

This Project titled "**Plastic Egg Detection Using Machine Learning Approach**", submitted by Palash Miah, ID: 181-15-10518 and Md. Hasan Jamil Sany, ID: 181-15-10519 to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 6th January, 2022.
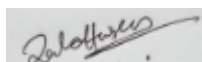
BOARD OF EXAMINERS

Chairman

_____

Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

_____

Zahid Hasan (ZH)
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

_____

Mohammad Monirul Islam (MMI)
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____ **External Examiner**
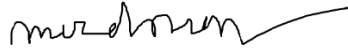
Dr. Dewan Md. Farid
Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We hereby declare that, this thesis has been done by us under the supervision of **Mr. Riazur Rahman, Sr. Lecturer, Department of CSE** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.
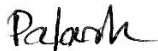
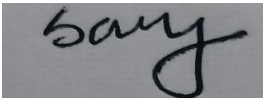**Supervised by:**

Mr. Riazur Rahman
Sr. Lecturer
Department of CSE
Daffodil International University

Submitted by:

Md. Palash Miah
ID: 181-15-10518
Department of CSE
Daffodil International University

Md. Hasan Jamil Sany
ID: 181-15-10519
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year thesis successfully.

We really grateful and wish our profound our indebtedness to **Mr. Riazur Rahman, Sr. Lecturer,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning" to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

We would like to express our heartiest gratitude to **Dr. Touhid Bhuiyan, Professor and Head,** Department of CSE, for his kind help to finish our thesis and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

In Bangladesh egg is almost like a daily meal for many peoples. And egg has so much cholesterol and vitamin in it. As a result, the demand of egg is so much higher in our country. So, there are many types of egg supplies seen in the past. But recently, it is seen that there are the supplies of plastic eggs. Which is pretty tough to recognize by physically whether it is genuine or fake. Nowadays, there are so many report in the newspaper, journal about trading plastic eggs in the market and peoples are unconsciously buy them as well. Those eggs are so much harmful for human health. Therefore, we have thought about to do something for this problem. We have implemented a machine learning approach to identify eggs in the market. For that, we have collected samples of real and plastic eggs and made large datasets by capturing and collecting the images. Using those datasets, we have created some Convolutional Neural Network (CNN) models for check their accuracy in identification of an egg that is real or plastic. Among the models, InceptionV3 was the most accurate which is 99.7%.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRUDUCTION

1.1 Introduction

Bangladesh's weather is perfect for poultry farming. The weather is friendly. And egg is a big part of the poultry farming. The annual investment in poultry in 2017 was 300 billion BDT. An international poultry Show and seminar in Bangladesh shows that the annual production of egg in Bangladesh is over 7.34 billion eggs [1]. In our country Bangladesh egg is a regular food. People love to eat egg in their morning breakfast to their other afternoon snack. Sometime it has been a crucial food for the poor families. Because of its pricing and easy to cook. Eggs are a better source of protein for low income people. Egg is lower than fish and meat in terms of pricing. Its 21st century on the calendar, and it has been so easy to make clone of anything around us. Egg is not apart from that. Those fake egg are made by people for extra profit and make the production level higher. Because the demand is so much higher for the egg between our peoples.

There were rumors that chines are making and exporting plastic or fake egg around the world. And the Asian countries are main target of that. A video got viral in Indian social media about the plastic/fake eggs [2]. Produced fake eggs looks mostly like the real one. It's tough for a normal human being to detect the difference by their own naked eye. In a normal Egg there is 6gm of protein in the white portion. Its more than the half of its protein. But the whole fake egg is filled with chemical, additives and auxiliary agents. So we understand there is no vitamin or nutrition left in the egg. By the result the fake egg is standing so much harmful for a human health. Those fake eggs effect a human body in many ways, among theme Cancer, Liver disease and Metabolic disorder are common. So generally the plastic eggs and their nutrition's are doesn't seems like a blessing for the human health [3].

So, for that we have created a computer based model by that we can detect the real and fake egg. For that we have captured and collected a large dataset. Where a took real and plastic eggs separately. As it's tough to recognize by the naked eye, so we made the mode for detect it by the computer. To either it's a real egg or plastic one.
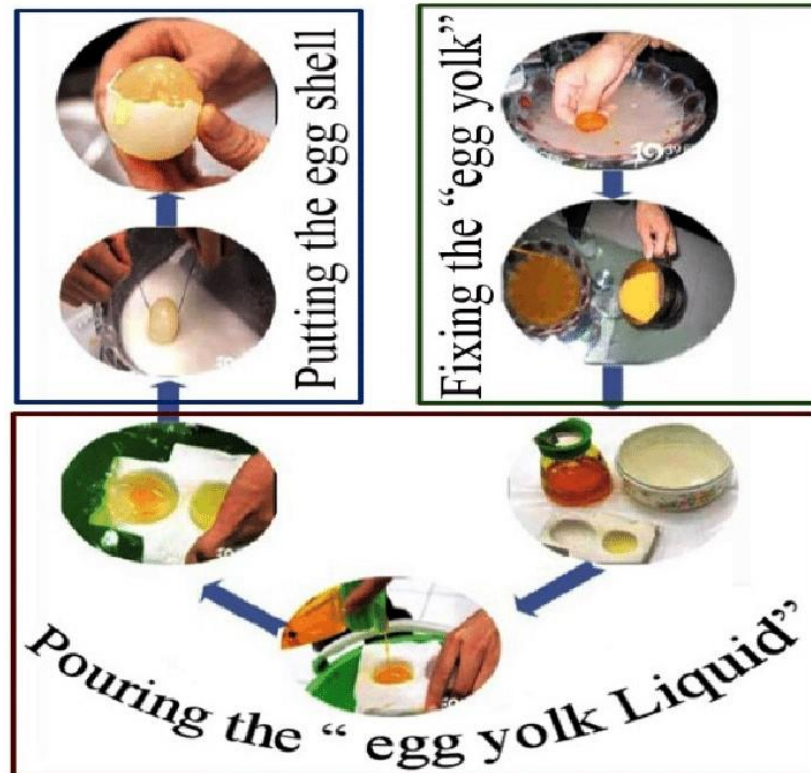
Figure 1.1.1 Procedure of preparing artificial eggs. ResearchGate [3]

1.2 Motivation

We know that the plastic/fake egg can be so much harmful for the human health. And as it is almost looks like the real one, it's tough for normal human being to detect that with his own naked eye. The sellers are taking advantage of that and mixing the fake eggs between the real ones. So, after taking all of those things in the consideration we choose to make a model like this to detect the fake one by using the computer area vision. By using this we can aware people from buying those unhygienic eggs which is bad for a human health and prevent them from buying those fake/plastic eggs. So, that was are motivation behind making this AI system.

1.3 Definition

As the eggs are common food in our country because of its easy to cook and its low price. So, we can understand the demand is so much high. It's a common food for the low income peoples. Taking it as an advantage some dishonest sellers are selling plastic/fake eggs

mixing with the real ones. They do it for extra income. As the plastic eggs looks fully similar to the real ones, people are buying those and eating them without even knowing it. Those eggs are filled with lots of chemicals and unhygienic products. So, it's so much harmful for our health. by eating plastic eggs, human's metabolism factor can be effected. Plastic eggs can also effect brain nerve and liver. For its harmful chemical's cancer can also effect in human body. Liver disease and cancer is common among of them [4].

1.4 Research Questions

Question which are focused in our thesis paper are given bellow:

i. Situation of egg production and importing in Bangladesh market?
ii. How can be classify between original and plastic egg?
iii. Why pre-trained models (VGG16, VV19 and InceptionV3) are proposed over Sequential model?
iv. How much the pre-trained models are performed in this project?

1.5 Research Method

At first we are going to collect of plastic and real both eggs. After that for identify we will sort or preprocess the data set. When the preprocessing is done, we will create a computer vision model from Machine Learning which suit our dataset. After choosing the right model we will start to train all the data we have in our dataset. When the training will be finished, we need to test some images from the dataset by the model to see either it can detect the real and plastic eggs or not.

1.6 Research Objective

Objectives behind our research:

i. To Make people aware to stay away from buying plastic eggs.
ii. To create a model which can identify real and plastic egg.
iii. To let people, know about the bad sides of eating plastic eggs.
iv. Help people to know about the plastic eggs characteristics.

1.7 Research Structure

- ✓ Chapter 1: Introduction, Motivation, Problem Definition, Research Questions, Research Methodology, Research Objectives.

- ✓ Chapter 2: Background Introduction, Related Works, Bangladesh Perspective.

- ✓ Chapter 3: Research methodology introduction, Data collection, the data augmentation stage, the pre-processing stage, The Feature-Extraction Stage, Architecture of the Model, Working Methodology.

- ✓ Chapter 4: Result and analysis – Experimental Result, Model Accuracy, Output Report.

- ✓ Chapter 5: Conclusion and the future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This research is about plastic egg Identification system which we made based in Bangladesh. For that we have created a Machine Learning algorithm which will help people to identify between plastic and real eggs.

2.2 Related Works

We used CNN Algorithm for this identification. CNN is Convolutional Neural Network which is a deep learning algorithm. CNN actually take image as input and use it to identify what we want to identify. As we are detecting real or plastic egg by image that's why we used CNN algorithm as our computer vision model.

This paper [5] is about food detection and reorganization system. They gather food image samples, applied a computer vision algorithm on them to tell about the nutrition the food got in it. They used CNN algorithm for that.

This paper [6] is about fruit classification which can be used by any shop to generate price by identifying the fruit. They have done that with CNN algorithm. For that they used image datasets.

This paper [7] is about the modern agriculture in chine. Where they made a vegetable reorganization system. By using that they can recognize vegetable easily. That can make work efficiency higher and also cut low the amount labor force.

This paper [8] is about identify between cat and dog by using image classification. They gathered 10000 images to train the algorithm. They used CNN algorithm here. Because CNN is known for an image reorganization computer vision model.

This paper [9] is about detecting pneumonia via neural network using X-ray images. They use datasets of Normal peoples and people who already detected with pneumonia chest X-rays to train the CNN algorithm.

This paper [10] is on detecting breast cancer by using Deep neural network. They actually working here with Computer Aided Diagnosis (CAD) system. That is quite popular nowadays.

On this paper [11] they worked on recognizing digits from handwriting by using CNN algorithms. They took hand written 1-9 digits from various peoples to train the computer vision algorithm.

The paper [12] is about recognizing emotion by human's facial expression. For that the used pictures of human face with various expressions. They use those as datasets and trained CNN algorithm with those images.

2.3 Bangladesh Perspective

Egg is regular food for in Bangladesh. As it is lower in cost son it's easy for low-income people to buy and eat. So, the demand is so much high we can tell by those facts. Sellers are taking advantage of that and mixing the plastic egg with the real ones by importing them. There are some videos also got viral on internet. It shows that, how the plastic eggs are making [13].

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The Convolutional Neural Network (CNN) provide the most productive technique for image recognition. It is a renowned multilayer neural network and Its neurons are small configured whose are work as input/output. By using CNN, we have work with a dataset of two different classes images of original and plastic egg and the accuracy we have got is 98%. To do that, our proposed models and the architectural design that is go through some basic several stage.
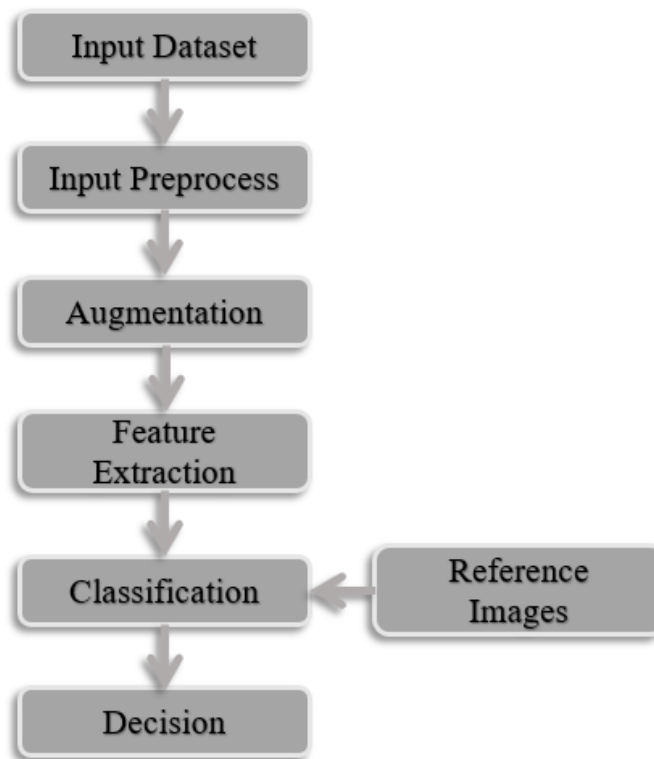
Figure 3.1.1 Working Flow Diagram

In Fig 3.1.1, describes the basic working flow of our research. At first, we have created a dataset of image of original and plastic egg. Then we have preprocessed and augmented the dataset to make as an input. After that we have created our proposed models and fit the

dataset into it. Then we have trained and validate our models. And finally, we have identified and analyzed the output by using different type classification technique which we will be discuss later.

3.2 Data Collection

The dataset contains images of original eggs and plastic eggs which are captured physically, where some of plastic egg images collected from internet and rest of the images are captured from artificial plastic eggs which we have ordered from renowned online shopping market named Daraz. That's how, we have successfully collected around 6,444 images. After collected all of them, we have made a dataset of 2 classes (original or fake) and placed them in separate folders. The dataset contains two type of egg images i.e. egg shell and boil egg images. Figure 3.2.1, represent the sample dataset of our work.

Figure 3.2.1 Sample Dataset of Egg Image

3.3 The Data Augmentation Stage

The data augmenting is a technique to modify or increase the amount of data. So, we have increased our data by doing slightly modification as well as augmentation. To do that, we have imported ImageDataGenerator class from Keras. we have applied horizontal flip as constant mode, height-width shift, transformations i.e. shear, zoom etc. And Finally, applied a random rotation about 45 degrees. After that, we have rescaled images pixel value (0-255 to 0-1) dividing 1 by 255.

3.4 The Pre-Processing Stage

The primary goal of using CNN in most image classification tasks is faster processing and reduce the computational complexity which is likely to increase if the input are images. In our dataset, the images we collected are various sizes. So, we have resized all the images into 224×224 pixel to better visualization as well as reduce heavy computation and to do faster processing.

3.5 The Feature-Extraction Stage

The features were extracted from different variants of pre-trained CNN models. The classification result obtained from CNN Sequential model as a benchmark model which is not a pre-trained model. Therefore, this stage deals with the description of Sequential Model. When we make our own model we will compare it with benchmark model that how efficient and accurate it is. Since, CNN can easily do the feature extraction, so we can do that with the multiple models with different combination of features.

3.6 Architecture of the Model

We use CNN (Convolution Neural Network) which is the most efficient framework in this field. To do that, Firstly, we have ensured the dataset which has around 6444 images and the size of dataset about 3.5GB depending of the how much images in the dataset. Then, we split data 75% for training and 25% for testing the models. Since the data are collected via physically so that there might be have some form of glitch, size variation etc. So, we reshape our data so our models will be performing better.

As we know, Sequential model is the base model of CNN. It has the several hidden layers like Conv2D, MaxPooling2D etc. And every parameter of the layers is not pre-trained and participate in the operation which is so complex to computation and increase processing time. So, we proposed to use some CNN pre-trained model which is offered by Keras. Those models are as follow:

    i.   VGG16

ii.   VGG19

iii.   InceptionV3 etc.

During configure the model, the input weights was provided as 'ImageNet' so that only the parameter which achieved by the ImageNet layers will be train and the rest of all the existing layers will remain disabled. For instance, VGG16 model will train only 25,089 parameters out of 14,739,777 parameters which is huge change in the model configuration. It's enough to reduce computational pressure and will increase performance of the model, where sequential model takes all the parameters of layers weather it's needed or not. So, we think our proposed model will do better in this prospect. Figure 3.6.1 describes the architecture of VGG16 model.
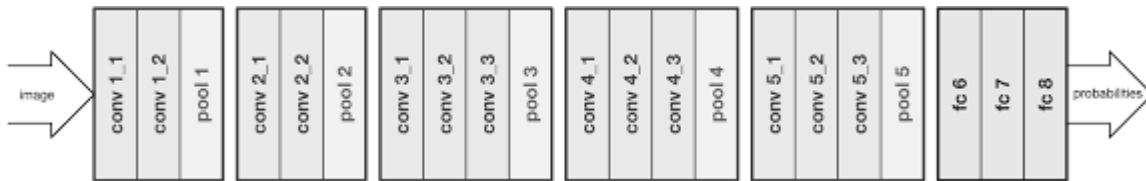

Figure 3.6.1 Architecture of a VGG16 Model

3.7 Working Procedure

As the input data are images, so we have converted all the images into 1-dimensional array. Though, we will only train our defined layer, we have freeze early layers i.e., Conv2D, Maxpooling2D etc. in the network and define ImageNet weights as a layer. Then we had flattened all the layers and connected all the layers with a fully connected Dense layer. By computing all the parameters of dense layer will have produced a 1-dimensional array of output value by using the 'sigmoid' activation function to identify weather the array value defined as an original egg or plastic egg. Figure 3.7.1 shows the basic configuration of models where we defined a layer at the bottom and froze rest of all the layers.
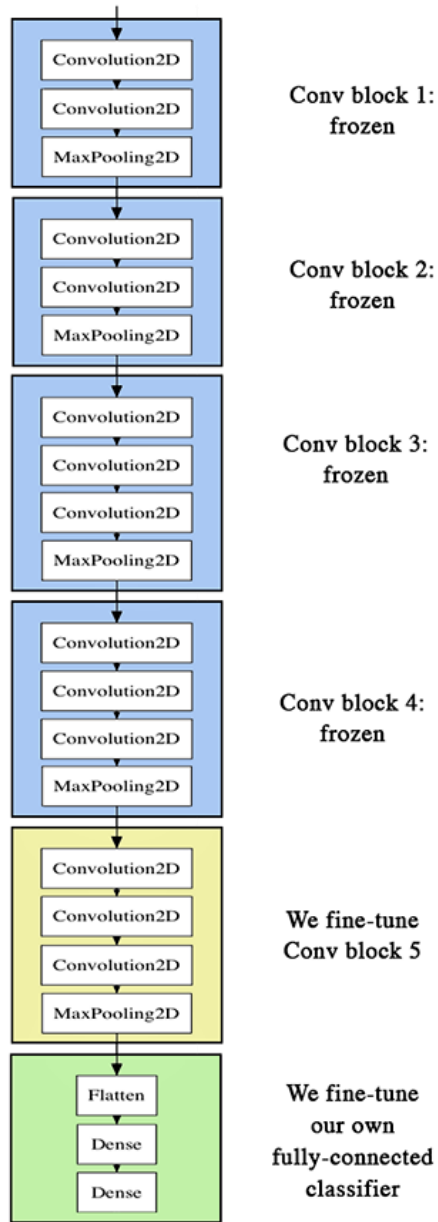
Figure 3.7.1 Basic Configuration of Layers [14]

The similar process will be repeat for all the model we selected to working with them. We have sketched a generic code to making the process for other models easier.

CHAPTER 4

Evaluation & Measure

4.1 Experimental Result

In this section, we have trained our pre-defined models with the desired dataset. Firstly, we have fitted the training and testing dataset into the VGG16 model. We have trained the model for 5 epochs step. After completing the training and testing process, we have generated a curve to show that how much performed our VGG16 model. For that, we have shown the comparison between 'accuracy' and 'loss' curve of VGG16 model.
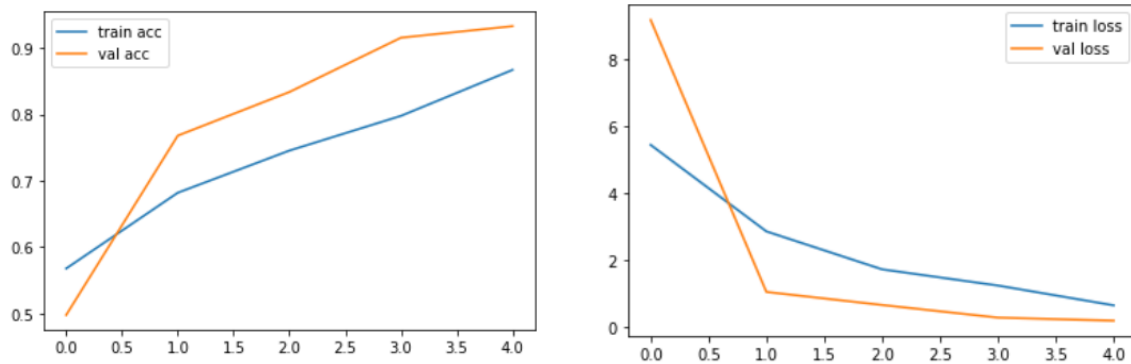

Fig 4.1.1 Accuracy and Loss Curve of VGG16

Again, we have trained our pre-defined models with the desired dataset. But this time, we have fitted the training and testing dataset into the VGG19 model. We have trained the model for 5 epochs step. After completing the training and testing process, we have generated a curve to show that how much performed our VGG19 model. For that, we have shown the comparison between 'accuracy' and 'loss' curve of VGG19 model.
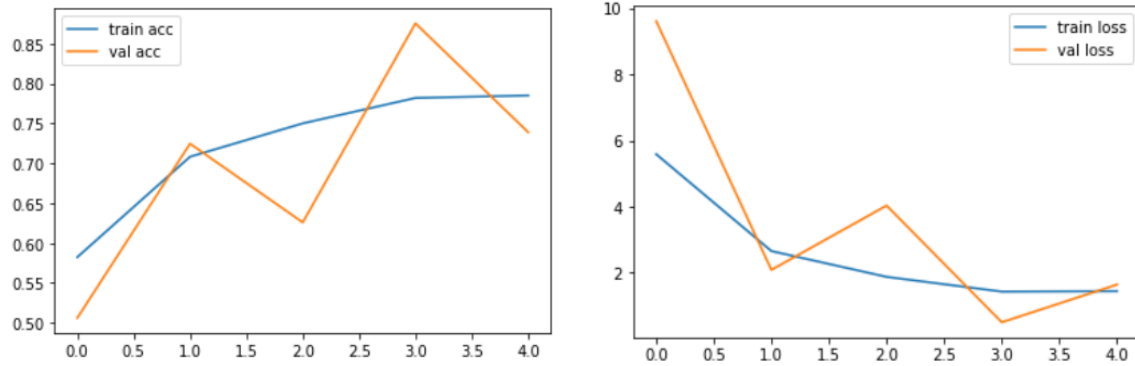
Figure 4.1.2 Accuracy and Loss Curve of VGG19

After again, we have trained our pre-defined models with the desired dataset with another model. And this time we have selected InceptionV3 model. So, we have fitted the training and testing dataset into the model. We have trained the model for 5 epochs step. After completing the training and testing process, we have generated a curve to show that how much performed our InceptionV3 model. For that, we have shown the comparison between 'accuracy' and 'loss' curve of InceptionV3 model.
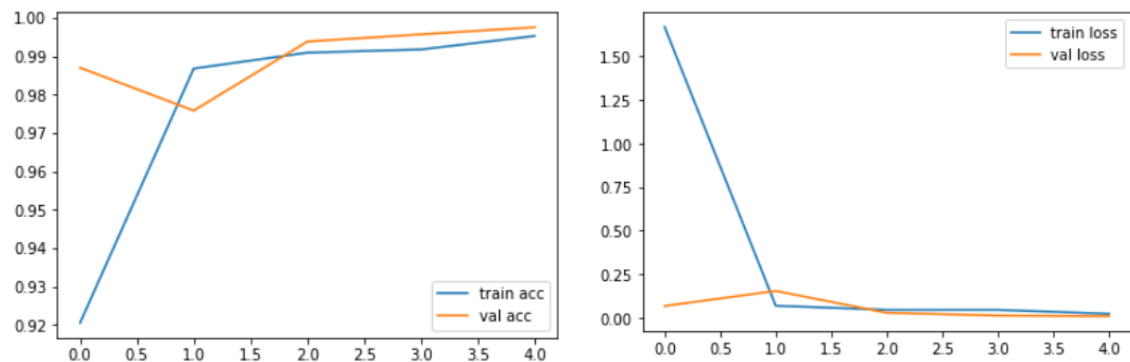

Figure 4.1.3 Accuracy and Loss Curve of InceptionV3

From above all the curves we can see that, VGG16 provided more consistency and efficiency over of all models during the training and validation period. And VGG19 provided an inconsistence curve with lower accuracy rate. But, the InceptionV3 provided higher accuracy rate over all the model. Therefore, InceptionV3 can perform better than all other models in the ImageNet classification.

4.2 Model Accuracy

In this project, we have trained and validated three different models, and each of the performed well with better accuracy. Among all of them, InceptionV3 model provide the higher accuracy rate with 99.7%. Here, Table 4.2.1 describes the average accuracy rate of the models.

Table 4.2.1 Accuracy rate of the models

Model	Accuracy
VGG16	93.8%
VGG19	73.8%
InceptionV3	99.7%

4.3 Comparison

The key objective of this project was to show that how much perform our proposed models over benchmark model. The sequential model was selected as benchmark model with the accuracy rate 93.8%. Where proposed models VGG16, VGG19 and InceptionV3 show the accuracy respectively 93.8%, 73.8% and 99.7%. VGG16 perform good as like sequential model. VGG19 doesn't perform very well where accuracy score is less than the benchmark model. That means, this model cannot provide justified score over sequential model. And last, InceptionV3 score highest accuracy rate among all the models. Therefore, our proposed model performs well over benchmark model to identify the original and plastic egg. Figure 4.2.1 show the comparison between the benchmark and selected models.
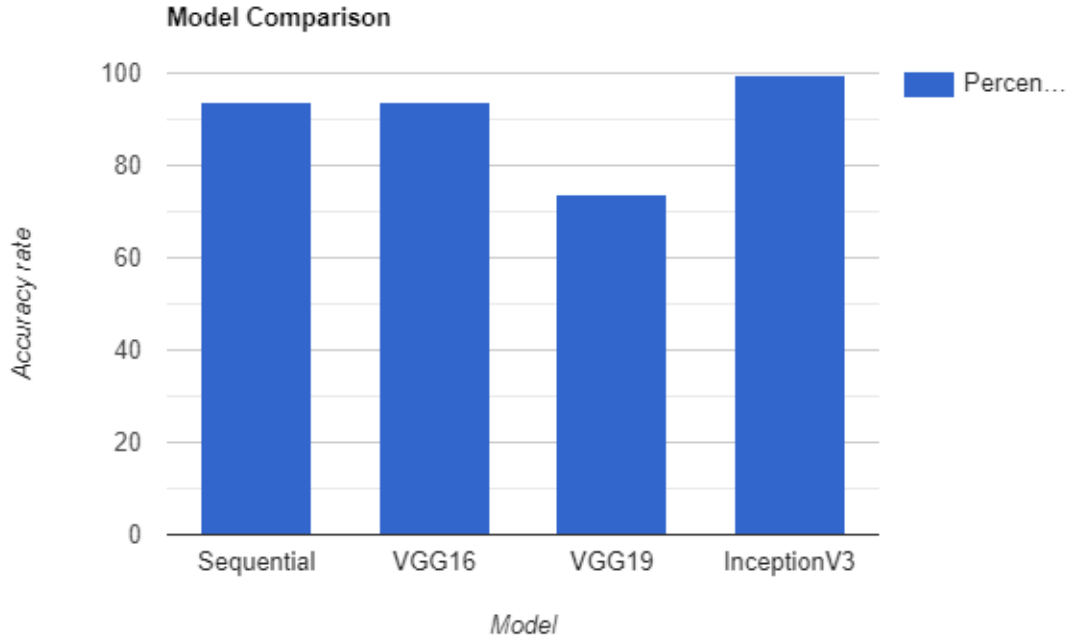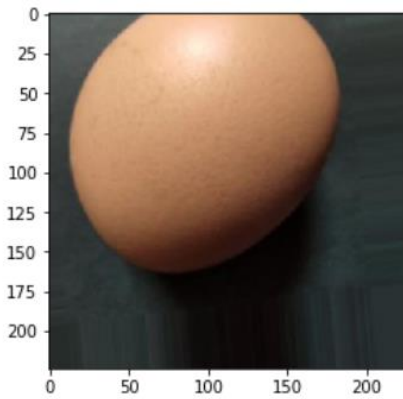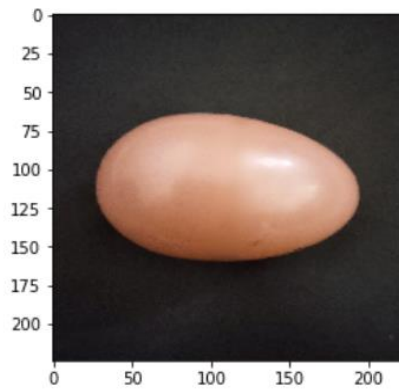
Figure 4.3.1 Comparison between the models

4.4 Output Result

The image is classified as Original Egg   The image classified as Plastic Egg

Figure 4.4.1 Sample output of images

We have selected some random images of original and plastic egg, then we have input the images as before which will covert images in form array. Then the array values are

processed through the output model. If the final calculation of the model provides output is the greater then '0.5' will be classified as original egg. And the output is less than or equal to '0' will be classified as plastic egg. Figure 4.4.1 describes the desired output of our project.
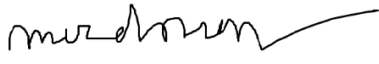
CHAPTER 5

CONCLUSION AND THE FUTURE WORK

In this project, our core objective was to identify the eggs available in the market are original or artificial/plastic egg which can be very dangerous for us. So that we implement a project to detect it where fraud traders are trying to sell the plastic eggs to the common buyer. To identify the problem, we have used a machine learning approach which is Convolutional Neural Network (CNN). In CNN algorithm, we have implemented the different models and compare that which model are perform well to identify the eggs. Among them, InceptionV3 model perform very well with the maximum accuracy. Therefore, this model can easily identify whether the eggs are genuine or plastic.

As, we have implemented binary classification here that a single egg can be identified by original or plastic made, in future we will try to implement multiple object classification where multiple eggs are can be detected as original or fake from a single scenario. Besides, we will implement a platform-based app for this problem that will be easily identify the egg.

References

[1] "Poultry Farming in Bangladesh," 16 September 2021. [Online]. Available: https://en.wikipedia.org/wiki/Poultry_farming_in_Bangladesh.

[2] "Plastic Eggs in Circulation in Indian Markets," July 2020. [Online]. Available: https://thelogicalindian.com/fact-check/fake-eggs-indian-markets-plastic-china-23273.

[3] "Procedure of preparing artificial eggs," August 2021. [Online]. Available: https://www.researchgate.net/publication/281149909_Artificial_and_Fake_Eggs_Dance_of_Death.

[4] "Fake eggs side effects and harmful sides," August 2021. [Online]. Available: https://www.quora.com/Is-it-really-healthy-to-eat-artificial-eggs-nowadays.

[5] K. A. M. O. Hokuto Kagaya, "Food Detection and Recognition Using Convolutional Neural Network," ACM Multimedia Orlando Florida, 2014.

[6] F. A. N. A. T. J. N. P. S. Mehenag Khatun, "Fruits Classification using Convolutional Neural Network," *Global Research and Development Journal for Engineering,* vol. 5, pp. 1-6, July 2020.

[7] . L. ,. Z. ,. Y. Zhenbo Li, "Vegetable Recognition and Classification Based on Improved VGG Deep Learning Network Model," *International Journal of Computational Intelligence Systems,* vol. 13, pp. 559-564, June 2020.

[8] P. G. Tushar Jajodia, "Image Classification – Cat and Dog Images," *International Research Journal of Engineering and Technology (IRJET),* vol. 06, pp. 570-572, December 2019.

[9] S. J. P. S. R. W. S. D. S. N. D. P. Hunmant Magar, "Pneumonia Detection using X-Ray Images with Deep Learning," *International Journal of Computer Applications Technology and Research,* vol. 9, pp. 183-185, May 2020.

[10] S. P. R. C. M. P. Karthik Sekaran, "Breast Cancer Classification Using Deep Neural Networks," in *Knowledge Computing and Its Applications*, 2018, pp. 227-241.

[11] B. Vijayalaxmi R Rudraswamimath, "Handwritten Digit Recognition using CNN," *International Journal of Innovative Science and Research Technology,* vol. 4, pp. 182-187, June 2019.

[12] A. A. Apeksha Khopkar, "Facial Expression Recognition Using CNN with Keras," *Bioscience Biotechnology Research Communications,* vol. 14, pp. 47-50, January 2021.

[13] "Video Leaked on Internet about plastic egg making," [Online]. Available: https://www.youtube.com/watch?v=s1JWHOwc4qk. [Accessed October 2021].

[14] "Building powerful image classification models using very little data," June 2021. [Online]. Available: https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html.

Plagiarism Checked By	
Mr. Riazur Rahman, Sr. Lecturer, Department of CSE	*(signature)*

Plastic Egg Detection Using Machine Learning Approach

ORIGINALITY REPORT

20%	18%	9%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	10%
2	www.coursehero.com Internet Source	4%
3	vixra.org Internet Source	2%
4	Biao Liu, Xiang Guo, Fan Zhang, Mao Feng, Kehu Yang, Tonghai Wang, Longfei Zhang. "Automatic Classification System For Mine Car Loading Based On Convolutional Neural Network In Coal Mine Auxiliary Shaft", 2020 Chinese Control And Decision Conference (CCDC), 2020 Publication	1%
5	"MultiMedia Modeling", Springer Science and Business Media LLC, 2020 Publication	1%
6	www.ijirset.com Internet Source	1%
7	98.131.160.83	

Internet Source

<1 %

8   Ahmed Khalid Ismael, AbdulSattar
     Mohammed Khidhir. "Evaluation of Transfer
     Learning with CNN to classify the Jaw
     Tumors", IOP Conference Series: Materials
     Science and Engineering, 2020
     Publication

<1 %

9   Lecture Notes in Computer Science, 2015.
     Publication

<1 %

10   www.grin.com
     Internet Source

<1 %

11   liboasis.buse.ac.zw:8080
     Internet Source

<1 %

12   iugspace.iugaza.edu.ps
     Internet Source

<1 %

13   rspublication.com
     Internet Source

<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		