PAPAYA RIPENESS PREDICTION USING MACHINE LEARNING

BY

MD IKTIDAR ISLAM

ID:183-15-11987

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

Supervised By

Md. Abbas Ali Khan
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
January 2023

APPROVAL

This Project titled **"PAPAYA RIPENESS PREDICTION USING MACHINE LEARNING"**, submitted by Md Iktidar Islam, ID No: 183-15-11987 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 24th January 2023.

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

Abdus Sattar
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Fatema Tuj Johra
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

24-01-23

Dr. Dewan Md Farid
Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We hereby declare that this project has been done by us under the supervision of Md. Abbas Ali Khan, Assistant Professor, Department of CSE Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Md. Abbas Ali Khan
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Md Iktidar Islam
ID:183-15-11987
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty Allah for His divine blessing making it possible to complete the final year project successfully.

I am really grateful and wish my profound indebtedness to **Md Abbas Ali Khan, Assistant Professor**, Department of Computer Science & Engineering, Daffodil International University, Dhaka for having Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to **Md Abbas Ali Khan**, Department of Computer Science & Engineering, for his kind help to finish my project and also to other faculty members and the staff of the CSE department of Daffodil International University. I would like to thank our entire course mates at Daffodil International University, who took part in this discussion while completing the coursework. Finally, I must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Fruits are a rich source of energy, minerals, and vitamins. Papaya is a perennial fruit of commercial importance due to its high nutritional properties. [13] The packaging of papaya fruit as per its maturity status is an essential task in the fruit industry. Classification of RIPE fruit can be used in many applications, whether industrial, agriculture or services, for example, it can help the manager in the hyper mall to determine the price according to the maturity(days). Manual labeling of papaya fruit based on human visual perception is time-consuming and inaccurate. [13]

In this paper, a machine learning-based approach is presented for classifying and identifying 7 different labels (days) with a dataset that contains 139 images, divided into 4 batches. Each batch contains 35 images. Use 28 images for training, 3 images for validation, and 4 images for testing. Few deep learning models were used that were extensively applied to image recognition. We used 80% of the images for training and 10 % of the images for validation and 10% for testing. One of our trained models achieved an accuracy of 100% on a held-out test set, demonstrating the feasibility of this approach.

Keywords: Fruit, Deep Learning, Classification, Prediction

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 Introduction

There are many ways in which the status of global food security can be improved for the world's increasing population [12]. Increased fruit production through adding to the area cropped is not sustainable, thus productivity per unit of the land area must be increased [12]. There's a necessity to forestall waste, and for fruit production, the temporal order of harvest is crucial to making sure that production meets the business ripeness specifications. Over- or under-ripe fruits have a lower or perhaps no retail worth and represent important financial gain loss and a waste of resources. For the consumers, too early harvest reduces the quality of fruits while a late harvest will result in a reduced period of time, poor look, and "off" flavors and odors. the first in-field assessment of fruit ripeness thus the and also the prediction of each harvest date and yield can therefore greatly cut back the waste within the provide chain and so facilitate towards up food security.

Bangladesh is a hugely populated country and its population is around 166.59 million (2019), where the remarkable population depends on agriculture. This agriculture has a remarkable contribution to GDP, which comparison 13.07% (2019) of the whole GDP [15]. When this agricultural product satisfies the country's demands it has exported to another country. To maintain this status, it is important to produce fresh fruits. In our process, images are through an expert system and the system determines the maturity of the ripe crops. Among numerous rural items, we have picked papaya for our work which is a popular and enormous cultivated fruit in Bangladesh.

Classification of ripe fruit can be used in many applications, whether industrial, agriculture, or services, for example, it can help the manager in the hyper mall to determine the price according to the ripeness(days). The manual grading of papaya fruit based on human visual perception may be inaccurate or wrong. Hence, an automatic system is needed to investigate the fruits and supply additional reliable info [3]. The identification of the maturity standing of fruits relates to intake taste and therefore the determination of storage time before consumption [4]. The objective of this paper is to suggest a novel non-destructive maturity status classification of papaya fruits.

In this paper, we proposed a classification model for the ripeness/maturity status classification of papaya fruits with the utilization of machine learning approaches. In this examination, we performed ripe/maturity(days) acknowledgment of papaya organic products with a machine learning-based framework which takes pictures and perceives the pictures of papaya organic products. In this system, we describe algorithms such as CNN, InceptionV3, VGG16, and, ResNet50 and their accuracy level. Based on the accuracy level we will use it to detect and recognize the ripe maturity(days) of the papaya fruits.

1.2 Motivation

There are numerous similarities between face recognition and fruit recognition, similar to colorful acts, illuminations, and occlusions. Nonetheless, there are some differences as well between both them. originally, compared with facial features (eyes, nose, mouth), the information contained in fruit is generally fairly simple. In general, the fruit point only includes the overall information (shape, color). Secondly, it is more likely to be brazened with heavy occlusion in the tasks of fruit discovery. Thirdly, there's no invariant standard for fruit discovery, and sufficient image accession and reflection are time-consuming tasks. Eventually, real-time is one of the most important indicators for fruit discovery. This is because the fruit discovery model is generally applied to automatic outfits, similar to picking robots, sorting robots, yield estimation robots, and so on. So, for the design of the fruit recognition model, the below-mentioned motives should be taken into consideration. Grounded on all that, we designed a fruit sensor that can descry fruits with different disguises, low resolution, and occlusion. Shows a typical illustration of papaya fruits appearing at different acts, sizes, distances, judgments, and occlusions. [14]

1.3 Rational of the Study

It is necessary to identify the difference between sufficiently ripe and unripe fruit; adequate ripening is required to ensure the taste and flavor of the fruit.

This paper is divided into the following sections. Section II introduces some recent research works related to fruit recognition and classification. Section III explains the proposed methodology. Section IV shows the experimental results and the last section compares the proposed algorithm with other existing algorithms and then presents the conclusion.

1.4 Research Objective

The objective of this study is to generate scientific knowledge based on a multidimensional understanding of interactions with ripening papaya fruit ripeness.

- To offer information about ripe papaya based on a multifaceted viewpoint.
- To investigate how ripe papaya ripes, the difference between day by day, until rotten.
- To provide such a system by which people can predict the shelf life of the fruits.

1.5 Report Layout

The report is divided into six sections. Each chapter addresses multiple aspects of the depression risk level. The various sections of each chapter are explained in more detail.

Chapter 1: Introduction

Chapter 1 serves as an important introduction to this research. This is related to predicting depression and gives an easy overview of the problem. This chapter discusses the research motivations, research objectives, relevant research questions, expected results and general management data, and economic consequences of results.

Chapter 2: Background & Literature Review

Chapter 2 contains a complete description of the research background. Based on the results of this research study, a machine learning system, classification information, and related activities were created. The chapter also provides a comparative analysis and scope of the problem and identified challenges.

Chapter 3: Research Methodology

Chapter 3 describes related work and presents the main general strategies associated with that work. Algorithmic details are presented for each method implemented, from the mathematical foundations to the current state of affairs. Also describes the approaches to data collection, data preprocessing, and element determination as well.

The principles of rating grouping, design specifications, and the results are also detailed in this chapter.

Chapter 4: Experimental Results & Discussion

Chapter 4 describes the experimental result and prediction.

Chapter 5: Future Scope and Conclusion

The future scope of this research activity is briefly outlined in Chapter 5 as the scope of this research study. This chapter concludes the entire research paper with a useful conclusion that briefly describes the main findings of the study.

CHAPTER 2
BACKGROUND & LITERATURE REVIEW

2.1 Related Works

Papaya is a high-nutritional fruit and is a rich source of vitamins A and C. In the year 2013, the production of papaya reached 1.25 × 107 metric tons worldwide [1]. The fruit classification and quality assessment by visual inspection cause errors due to external influences such as fatigue, vengeance, and bias [2]. Despite skilled operators, the classification of fruits within the fruit trade ends up in inconsistencies owing to variations in seeing. Hence, an automatic system is needed to investigate the fruits and supply additional reliable info [3]. The identification of the maturity standing of fruits relates to intake taste and therefore the determination of storage time before consumption [4]. The determination of those properties with the assistance of human operators may be inaccurate or wrong. Thus, rapid, intelligent, and non-destructive techniques are needed in this application domain [5].

Many researchers have reported their work on the ripeness/maturity prediction of various fruits. A fuzzy model [6] is projected for the classification of the ripeness/maturity level of pineapple fruit. Achieved 93.11% of accuracy with Particle swarm optimization, used for optimizing the fuzzy model.

Two automated algorithms, the mean color intensity algorithm, and the area feature algorithm examined the ripeness/maturity classification of banana fruit into three levels [7].

A fruit sorting system was designed to classify the mango fruits into four classes in step with their ripeness/maturity level [8]. Another model [9] supported multispectral imaging was projected for quality attribute estimation and ripeness/maturity standing classification of strawberries. Here two strategies were used, i.e. principal part analysis with backpropagation neural network (PCA-BPNN) and therefore the support vector machine (SVM) it had been reportable that the accuracy of 100 percent was obtained by SVM for ripeness/matureness stage classification on a dataset of 280 pictures. The non-destructive, non-invasive detector system is projected for durian fruit ripeness maturity prediction [10]. The tactic used the idea of wireless communication with the rician k-factor for maturity prediction and achieved 92.7% of accuracy.

The ripeness/maturity classification model for plum fruit [11] is proposed based on the image processing technique. The external quality features like color, texture, and size were used for evaluating the maturity of plum fruit. The correlation between image analysis and chemical analysis was 99.66%, the analysis implies the color features (RGB indices) were the dominant feature over the texture and size features.

Ripeness classification of fruit leads to inconsistencies because of variations in visual perception. In short, A fuzzy model is proposed for the classification of the ripeness/maturity level of pineapple fruit. The ripeness/maturity classification of banana fruit into three levels is examined by two automated algorithms. SVM for the ripeness stage classification of mango fruit.

Some work has been done on papaya fruit. Such as diagnosis of papaya fruit disease etc. However, there was no work done yet on predicting the ripeness of the papaya fruit. There is a research gap here. We have done this work to fill this gap.

2.2 Scope of the problem

I used a variety of machine learning algorithms and model training and test datasets. I am trying to identify ripeness by identifying associations between attributes in my dataset.

The use of automated fruit recognition proves to be a non-time-consuming procedure. Better maturity analysis algorithms allow the system to recommend maturity of ripeness to consumers, saving time and money, and fulfilling the desired taste of fruit.

2.3 Challenges

We tried to collect image data of ripe papaya from the beginning of this thesis.

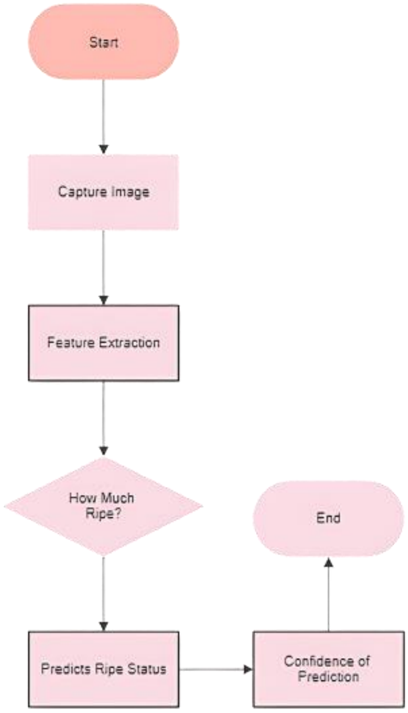
As a result, the accuracy of this sort of data always provides low accuracy. That’s why we have to apply hyperparameter tuning for the model.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Proposed Model

In this section, we will present an outworking architecture for our proposed model. Using an image of papaya fruit and sending it to the system that recognizes the fruit's ripe label(days). To recognize the papaya ripe label(days). At first, we will load all the training and testing images into our python IDE.



3.2 Data Collection

I have created a dataset by collecting some initially ripe papaya from a fruit garden. Then I took pictures of that ripe papaya every day till those papayas got rotten.

Those papayas were good in condition from the 1-5th day. On the 6th day, those took the color of rotten but the smell was good. And on the 7th day, all papaya got rotten completely by both looks and smell.

So, I have 7 classes of images. Names Day-1,2,3,4,5,6 and Rotten.

3.2.1 Dataset

We have used 139 images to predict the approaching model in 4 batches, each containing 35 images where 28 images are for training, 3 images for validations, and 4 images for testing the model. The data sets are collected from real-life captured images.

In the models, the input image data are used in 3 dimensions and that are height, width, and RGB colors. In our models, we fixed and first of all, images are converted as 256x256x3 pixels. So, the models mainly work RGB color of a papaya image, and then classification is completed.

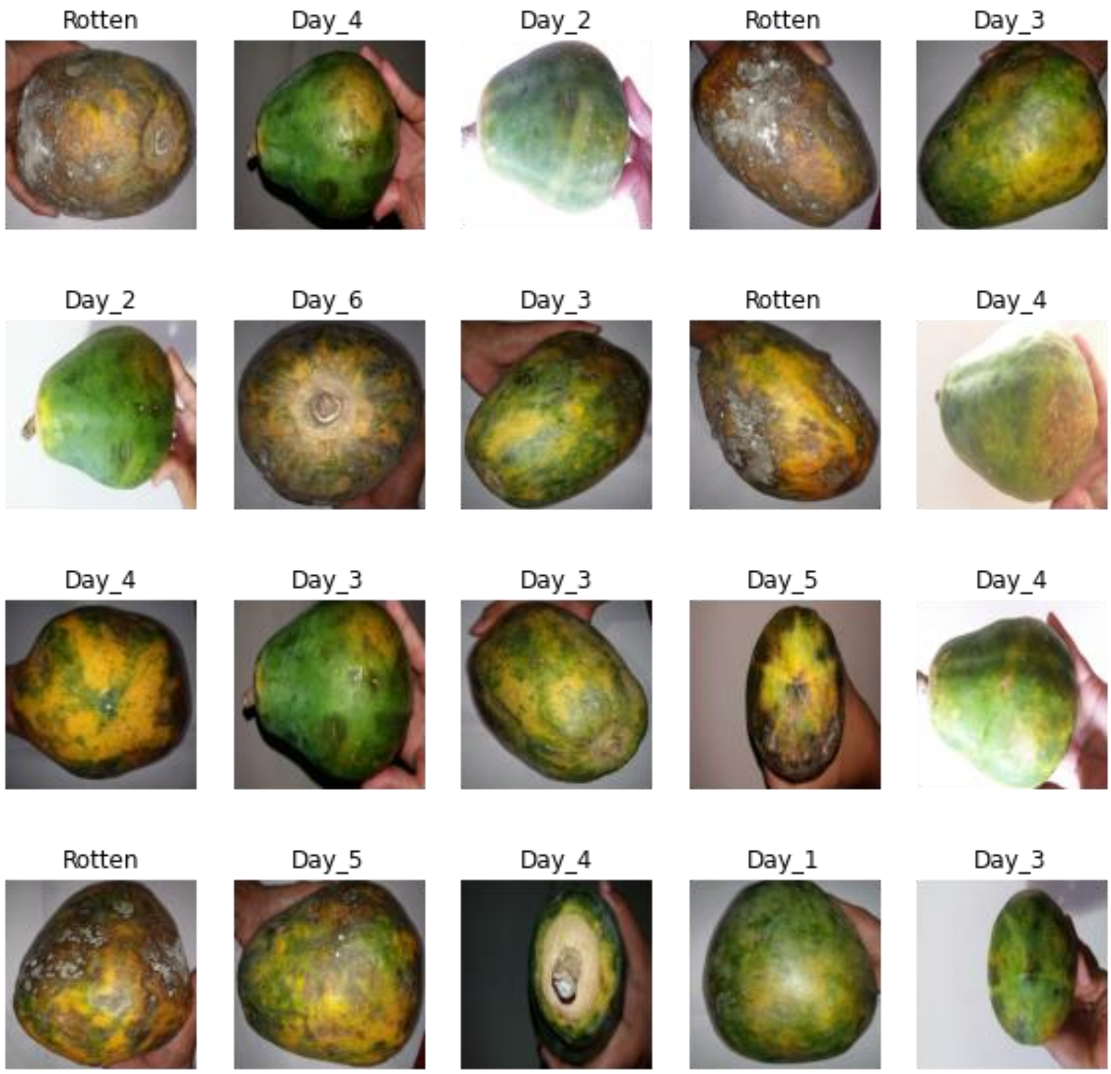


Fig 1: Images of Dataset

3.2.2 Data Set Description

In this work, nearly 140 images were collected from Ashulia, Dhaka. Each of these images contained one object of papaya fruit. All the objects were labeled manually as individual image datasets.

After the acquisition, we divided these images into seven different types, i.e. Day-1,2,3,4,5,6, and Rotten.

Accumulating diverse samples for the model training can improve the performance of the detector. However, collecting a sufficient number of samples is tedious. IOU value for seven types of sample division is a time-consuming task, and it is not convenient to train the detector with a new category. Therefore, many studies have been conducted on augmentation methods such as rotation, translation, scaling, adding Noise, and so on, and these methods have improved the performance to some extent as well. Through these methods, we found that the augmented samples were very close to the real environment and they improved the final performance as well. Motivated by this, we randomly extracted several sizes of negative patches from the original images. After that, we augmented our dataset.

3.2.3 Data Pre-processing

An arithmetic augmentation of horizontal and vertical is applied to each image. The fruit is segmented from the background by thresholding the images. Also, the image is resized to 256x256 pixels to increase the speed of the processing.

3.2.4 Splitting Dataset

In order to use any machine learning technique, the dataset must be divided into three parts: one for model training, and testing and the other for model validation check. This is known as data partitioning. Before utilizing any machine learning approaches, this must be done. I divided my data set into 4 image batches. Each batch has 35 images. I divided a batch of 35 images into 80% for training and 10% each for testing and validation.

3.3 Training Model

Our objective is to present a comparative study with a predictive model. So, we have used four machine-learning models in our work. Our four machine-learning models are CNN, Inception V3, VGG16, and ResNet50.

3.3.1 Convolutional Neural Network

Convolutional Neural Network, A subfield of deep learning that deals with images on all scales of computer vision. It enables the computer, through an automatic process, to process and understand the content of a large number of images. Computer vision's primary architecture is the conventional neural network that is a derivative of feed-forward neural networks. Its applications are very different, such as image classification, object detection, transfer of neural styles, and identification of the face. A series of convolutional and pooling layers is a convolutional neural network that allows the main features to be extracted from the images that respond best to the final objective. The convolutional layer defends on convolutional operation between feature images. [15]

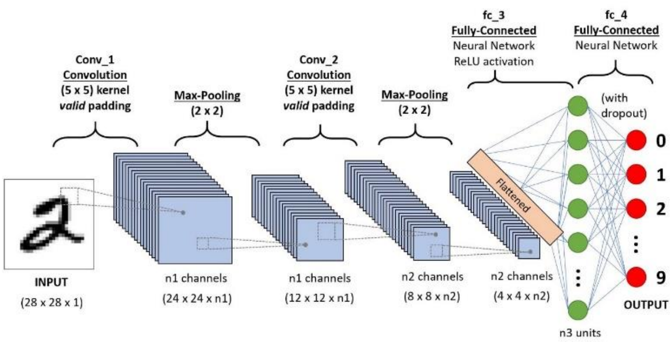


Fig 2: Convolutional Neural Network

3.3.2 Inception V3

The Inception V3 is a deep learning model based on Convolutional Neural Networks, which is used for image classification. The inception V3 is just the advanced and optimized interpretation of the Inception V1 model. The InceptionV3 model used several ways for optimizing the network for better model adaption.

- It has higher efficiency
- It has a deeper network compared to the Inception V1 and V2 models, but its speed isn't compromised.
- It is computationally less expensive.
- It uses auxiliary Classifiers as regularizes.

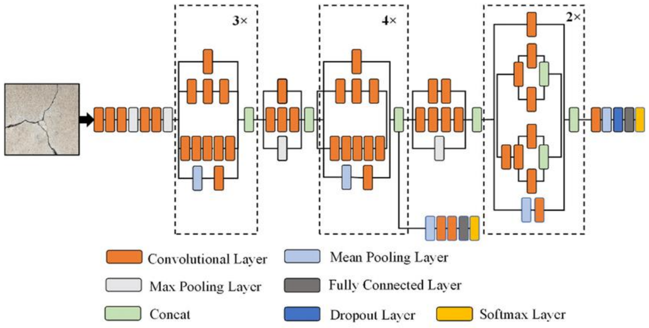


Fig 3: Inception V3 Architecture

3.3.3 VGG16

VGG16 is a type of CNN (Convolutional Neural Network) that's considered to be one of the most stylish computer vision models to date. The generators of this model estimated the networks and increased the depth using an armature with veritably small (3×3) complication pollutants, which showed a significant enhancement on the previous- art configurations.

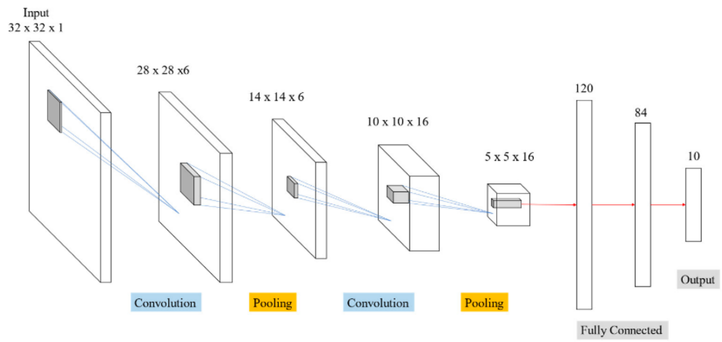


Fig 4: VGG16 Architecture

3.3.4 ResNet50

ResNet stands for Residual Network and is a specific type of convolutional neural network (CNN). The original ResNet armature was ResNet- 34, which comprised 34 weighted layers. It handed a new way to add further convolutional layers to a CNN, without running into the evaporating grade problem, using the conception of shortcut connection known as “skips over”. The regular network was based on the VGG neural networks (VGG- 16 and VGG- 19) — each convolutional network had 3×3 filters. Still, a ResNet has smaller pollutants and is less complex than a VGGNet. The ResNet armature follows two introductory design rules. First, the number of pollutants in each filter is the same depending on the size of the affair point chart. Second, if the point chart’s size is halved, it has double the number of pollutants to maintain the time complexity of each filter.

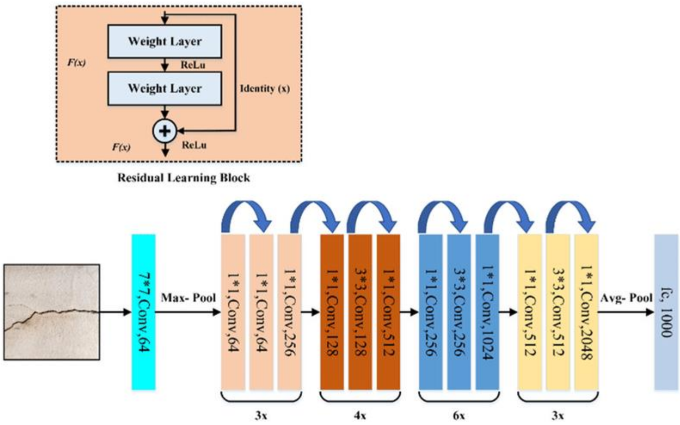


Fig 5: ResNet50 Architecture

3.4 Model building

To build our models, everything we used is described below,

- **Kernel** - In a Convolutional neural network,
  - The kernel is nothing but a filter that is used to extract the features from the images.
  - The kernel is a matrix that moves over the input data, performs the dot product with the sub-region of input data, and gets the output as the matrix of dot products.
  - Kernel moves on the input data by the stride value. If the stride value is 2, then the kernel moves by 2 columns of pixels in the input matrix.

- **Pooling** - Pooling in convolutional neural networks is,
  - A technique for generalizing features extracted by convolutional filters and helping the network recognize features independent of their location in the image.
  - Pooling is required to down sample the detection of features in feature maps.
  - Pooling layers provide an approach to down-sampling feature maps by summarizing the presence of features in patches of the feature map. Two common pooling methods are average pooling and max pooling which summarize the average presence of a feature and the most activated presence of a feature respectively.
  - Two types of Pooling. Max Pooling & Average Pooling
    - In max pooling, the filter simply selects the maximum pixel value in the receptive field. For example, if you have 4 pixels in the field with values 3, 9, 0, and 6, you select 9.
    - Average pooling works by calculating the average value of the pixel values in the receptive field. Given 4 pixels with the values 3,9,0, and 6, the average pooling layer would produce an output of 4.5. Rounding to full numbers gives us 5.

- **Flatten layer** - Flattening is used to convert all the resultant 2-Dimensional arrays from pooled feature maps into a single long continuous linear vector. The flattened matrix is fed as input to the fully connected layer to classify the image.
- **Dense layers** - In any neural network, a dense layer is a layer that is deeply connected with its preceding layer which means the neurons of the layer are connected to every neuron of its preceding layer. This layer is the most commonly used layer in artificial neural network networks.

3.5 Our Models

How we built our models is described below,

1. **CNN –**

   We have taken Six kernels and max pooling to build our model. Each kernel size is 3*3 and max-pooling size 2*2 is done after each kernel.

- Initially, our image is 256*256. When this image goes through the first kernel (3*3), the image size will reduce to (254*254) as the equation is {(image size-kernel size) +1}. Now the first max-pooling (2*2) layers will run, and it will reduce the image size to (127*127) as the equation is (image size/max-pooling size.
- Similarly, after the second kernel layers & second max-pooling, the image size will be (125*125) and (62*62) respectively.
- In the third kernel layers & third max-pooling, the image size will be (60*60) and (30*30) respectively.
- In the fourth kernel layers & fourth max-pooling, the image size will be (28*28) and (14*14) respectively.
- In the fifth kernel layers & fifth max-pooling, the image size will be (12*12) and (6*6) respectively.
- Lastly, the sixth kernel layers & sixth max-pooling, the image size will be (4*4) and (2*2) respectively.

At last, we flatten our results and compile the dense layers.

Table-1: CNN Summary

Learnable Layer	Calculate Parameters	No. of parameters
Conv2d_12	{(256-3) +1}	896
Max_pooling2d_12	254/2	0
Conv2d_13	{(127-3) +1}	18496
Max_pooling2d_13	125/2	0
Conv2d_14	{(62-3) +1}	36928
Max_pooling2d_14	60/2	0
Conv2d_15	{(30-3) +1}	36928
Max_pooling2d_15	28/2	0
Conv2d_16	{(14-3) +1}	36928
Max_pooling2d_16	12/2	0
Conv2d_17	{(6-3) +1}	36928
Max_pooling2d_17	{(4/2) +1}	0
Dense_4		16448
Dense_5		455

Summary of our model –

- Total params: 184,007
- Trainable params: 184,007
- Non-trainable params: 0

2. Inception V3 –

To build inception V3, we used the inception v3 library. We import the library and give it a shape of 256x256x3.

Summary of our model –

- Total params: 24,383,299
- Trainable params: 2,580,515
- Non-trainable params: 21,802,784

3. **VGG16**

First, we set our image size 256x256. Then we used 3 filters 64, 128, and 256 respectively. After each filter, we set a 2x2 max-pooling. Last, set the dense layer to 64 and activation was 'SoftMax'.

Table-2: VGG16 Summary

Learnable Layer	Calculate Parameters	No. of parameters
conv2d_768 (Conv2D)	(None, 256, 256, 64)	1792
conv2d_769(Conv2D)	(None, 256, 256, 64)	36928
max_pooling2d_92 (MaxPooling2D)	(None, 128, 128, 64)	0
conv2d_770 (Conv2D)	(None, 128, 128, 128)	73856
conv2d_771 (Conv2D)	(None, 128, 128, 128)	147584
max_pooling2d_93(MaxPoolingD)	(None, 64, 64, 128)	0
conv2d_772 (Conv2D)	(None, 64, 64, 256)	295168
conv2d_773(Conv2D)	(None, 64, 64, 256)	590080
conv2d_774 (Conv2D)	(None, 64, 64, 256)	590080
max_pooling2d_94(MaxPooling2D)	(None, 32, 32, 256)	0
max_pooling2d_95(MaxPooling2D)	(None, 16, 16, 256)	0
max_pooling2d_96(MaxPooling2D)	(None, 8, 8, 256)	0
flatten_15 (Flatten)	(None, 16384)	0
dense_36 (Dense)	(None, 64)	1048640
dense_37(Dense)	(None, 64)	4160
dense_38 (Dense)	(None, 7)	455

Summary of our model –

- Total params: 2,788,743
- Trainable params: 2,788,743
- Non-trainable params: 0

4. **ResNet50**

To build ResNet50, we used the ResNet50 library.

Summary of our model –

- Total params: 23,602,055
- Trainable params: 14,343
- Non-trainable params: 23,587,71

3.6 Evaluating Model

In this research, we have performed the Convolutional neural network algorithm (CNN), Inception V3, Visual Geometry Group (vgg16), and Residual Network (ResNet50) model using the Keras API module. Firstly, we have collected data sets regarding the ripeness of papaya and per-process them by minimizing shape and transforming. This research has performed ripe papaya image recognition and prediction for fruits. We trained the image dataset and validate the dataset using the same training and validation image dataset. The trained, test and validation ratio for this model is 80:10:10.

Later we create a CNN, inception v3, vgg16, and ResNet50 model using the Keras module. For this model, Keras conv2d is a constitutional 2D layer and this layer creates a convolution kernel that is wind with layers input which helps to produce a tensor of outputs.

Fig 6: Image of papaya

Our models using Keras API are able to classify the image layer and find the pattern of the area using neural networks. We did not get good accuracy for the CNN model with 20-100 epochs first, then we set the epochs to 200 and got very good accuracy. Then building other models, we ran 200 epochs on all models to maintain compatibility.

The average epoch time for the CNN model to train was 14 seconds. The InceptionV3 took 19 seconds for each epoch. The most 183 seconds of training time was taken the by vgg16 model. And ResNet50 took 27 seconds on average for each training set. Later we test the test dataset and where CNN obtained an average of 87.5% accuracy. The

Inception V3 model obtained an average of 75% accuracy. The vgg16 model obtained an average of 57% accuracy. The ResNet50 model obtained 100% accuracy.

The average test time for the CNN model to train was 3 seconds. The InceptionV3 took 2 seconds for testing. 7 seconds of test time was taken the by vgg16 model. And ResNet50 took 3 seconds of test time.

Table-3: Models Evaluation

Model	**Training Time /Epochs**	**Testing Time**	**Accuracy**
CNN	14s	3s	87.5%
InceptionV3	19s	2s	75%
Vgg16	183s	7s	57%
ResNet50	27s	3s	100%

From the above data, we can see that the accuracy of these two models Vgg16 and InceptionV3 is comparatively low and the training time is high. On the other hand, the training time of the CNN model is less and the accuracy is good. Again, the training time of the Resnet50 model is slightly higher than the inceptionV3 model, but the accuracy of the model is much better.

So, we will exclude the vgg16 and inceptionV3 models for further deep analysis. Let's look at more analysis of the CNN and ResNet50 models, such as classification reports and prediction in chapter 4.

3.6.1 Cross Validation

We employ k-fold cross-validation to provide an unbiased evaluation of the model's performance when there is just a small quantity of data available. In k-fold cross-validation, the data is split into ten pieces of equal size. Ten models are built, each with one of the training subsets left out and used as the test set. The term "leave-one-out" is used when k equals the sample size. There are two sections to the model evaluation:

- Classification Evaluation
- Error Detection

3.4.1.1 Classification Evaluation

The number of accurate and inaccurate predictions generated by the classification model in relation to the actual outcomes (target value) of the data is shown in a confusion matrix. The matrix has dimensions (NxN), where N is the number of expected values(classes).

The performance of such models is usually evaluated using matrix data. In the table below, a 2x2 confusion matrix for two classes is shown (Positive and Negative).

Table-4: Confusion Matrix

	Predicted Class		
Actual Class		1	0
	1	TP	FN
	0	FP	TN

Accuracy, specificity, and precession were determined using the Ramana et AI formula, which may be expressed as follows-

- Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$
- Precision = $\frac{TP}{TP+FP}$
- Sensitivity/Recall = $\frac{TP}{TP/FN}$
- F1-Score = $\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Accuracy}}$

All of the measurements are calculated using four values: true positive, false positive, true negative, and false negative. These values are discussed more below.

● FN = false negatives: the number of cases projected negative but found to be positive.

● TP = true positives: the number of positive cases that were expected.

● FP = false positives: the number of cases predicted positively but turned out to be negative.

● TN = true negatives: the number of expected negative cases that are truly negative.

3.4.1.2 Error Detection

We have utilized Mean Absolute Error, Root Mean Squared Error and Relative Absolute Error among other criteria for assessing and contrasting categorization models once they have been built.

3.4.1.2.1 Mean Absolute Error

Only models with the same error units can be compared since the mean absolute error (MAE) is assessed in the same units as the original data. Although it is significantly smaller, its magnitude is typically comparable to RMSE.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |p_i - a_i|}{n}$$

Here, a = actual target, p = predicted target

3.4.1.2.2 Root Mean Squared Error

The root mean square error (RMSE) is a popular approach for determining a model's error rate. It can only be contrasted with models that have equivalent error units, though.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} |p_i - a_i|}{n}}$$

3.4.1.2.3 Relative Absolute Error

The relative absolute error (RAE), like the RSE, may be used to compare models whose mistakes are quantified in various units.

$$\text{RAE} = \frac{\sum_{i=1}^{n} |p_i - a_i|}{\sum_{i=1}^{n} |\bar{a} - a_i|}$$

CHAPTER 4

EXPERIMENTAL RESULTS & DISCUSSION

4.1 Classification Report

The algorithm's accuracy and scores were evaluated after Machine Learning Model Creation was successfully implemented in order to determine how the algorithm is working and how successful at predicting.

Finally, the experimental findings constitute an analytical section in which each potential score for each algorithmic application and technique may be assessed.

Table-5: Performance analysis of CNN

	Precision	Recall	f1-score	Support
DAY_1	75.09	74.46	75.11	284
DAY_2	78.34	81.32	79.50	313
DAY_3	77.04	77.37	76.00	293
DAY_4	88.32	90.67	88.43	320
DAY_5	83.89	81.33	82.11	306
DAY_6	85.22	77.37	84.71	286
ROTTEN	87.42	82.80	83.23	304
ACCURACY	87.50	87.50	87.50	0.8750
MACRO AVG	82.18	80.76	81.30	2106
WEIGHTED AVG	88.64	87.50	86.64	2106

The above table shows the classification report of our CNN model.

Table-6: Performance analysis of ResNet50

	Precision	Recall	f1-score	Support
DAY_1	**98.7**	**100.00**	**100.00**	**284**
DAY_2	**98.1**	**100.00**	**100.00**	**313**
DAY_3	**99.9**	**100.00**	**100.00**	**293**
DAY_4	**99.5**	**100.00**	**100.00**	**320**
DAY_5	**99.6**	**100.00**	**100.00**	**306**
DAY_6	**99.2**	**100.00**	**100.00**	**286**
ROTTEN	**99.8**	**100.00**	**100.00**	**304**
ACCURACY	**100**	**100**	**100**	**0.100**
MACRO AVG	**99.25**	**100**	**100**	**2405**
WEIGHTED AVG	**98.67**	**100**	**100**	**2405**

The above table shows the classification report of our ResNet50 model.

Table7: Performance analysis

Model Name	MAE	MSE	RMSE
CNN	**10.7**	**26.4**	**16.2**
ResNet50	**1.03**	**1.87**	**1.37**

The above table contains the error rate such as Mean absolute error, mean squared error and Root mean squared error.

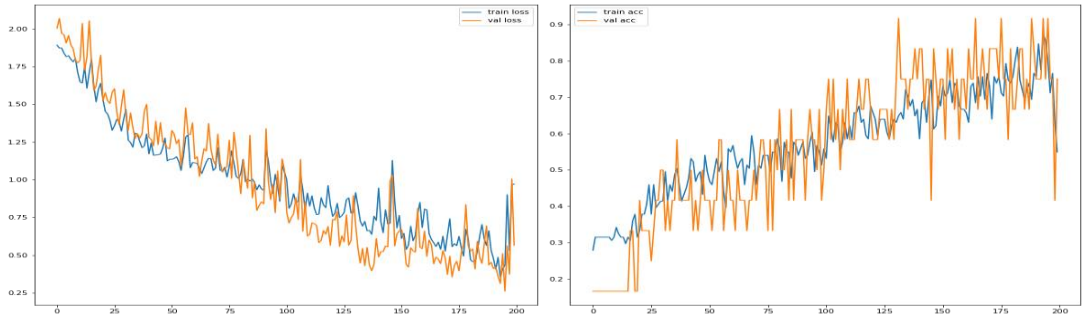


Fig 7: CNN Loss and Accuracy

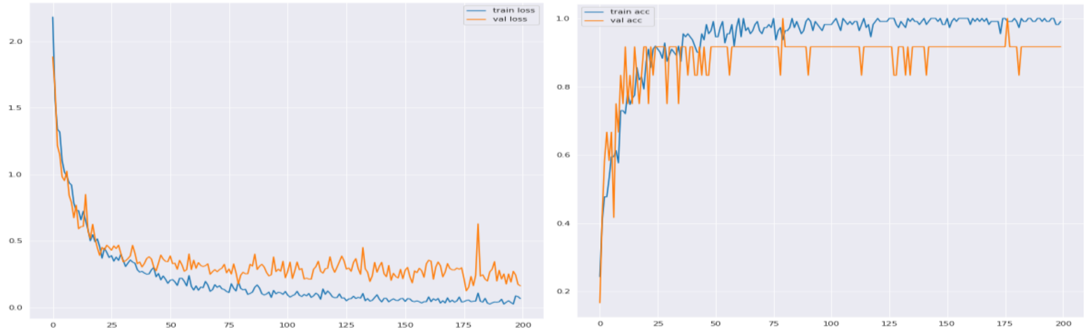


Fig 8: ResNet50 Loss and Accuracy

From the above four graphs, we can see the training accuracy, training loss, and validation accuracy, validation loss of both the CNN model and ResNet50 models.

If we look at training loss and accuracy, we can see that training loss decreases with increasing accuracy. On the other hand, looking at the validation loss and accuracy, it is seen that, as validation accuracy increases, validation loss decreases. We achieve an average of 87.5% accuracy from our CNN prediction model and 100% accuracy from ResNet50 prediction Model.

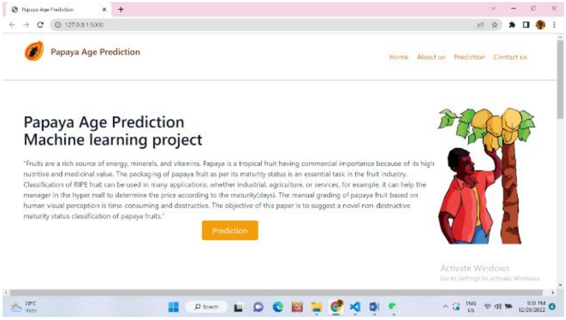
4.2 Predictions

Here are some of our predictions from both models. The system randomly Selects some photos and predicts the ripeness. The system also generates confidence in prediction based on how sure the system is about the prediction.

Fig 9: CNN Prediction

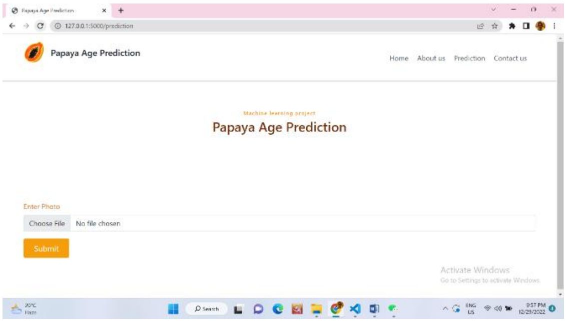
Fig 10: ResNet50 Prediction

4.3 Predictions on Our Website

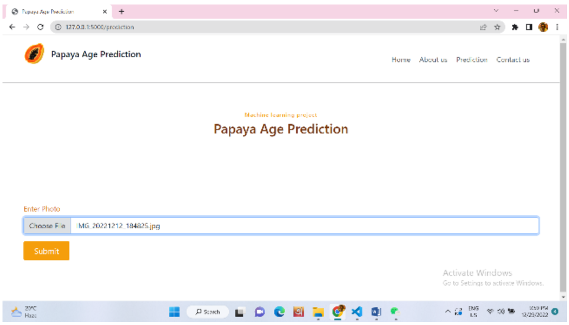


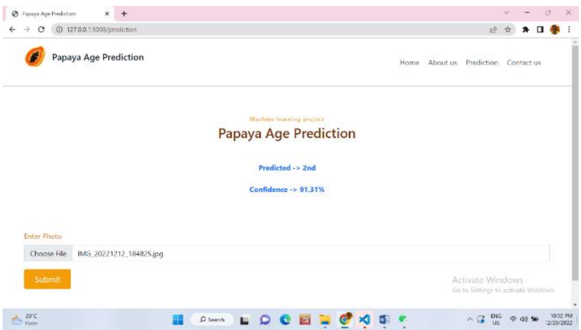
Home page of the Website

- Click on Prediction



- Chose a file





- Predicted on the Website

CHAPTER 5
FUTURE SCOPE AND CONCLUSION

5.1 Future Scope

We find proposed multi-task cascaded convolutional networks-based fruit detectors have good performance of timeliness and accuracy to meet the requirements for the visual system of harvesting robot from the experimental results. However, there is still a long distance for practical application and promotion of the harvesting robot. One of the most important tasks is to determine the order for all detected fruits. In other words, is to decide which object should be first considered for picking. Compared with picking manually, human visual attention can solve this kind of problem effectively. On the basis of this study, we will focus on the study and mimic human visual attention when viewing the scene through relevant studies such as visual saliency detection and semantic segmentation. In the future, we will also study the characteristics of fruit deeply and design a more reasonable and effective network model for fruit recognition tasks. Besides this, improving and optimizing the accuracy of the detector is also an important task for the future.

5.2 Conclusion

In this study, we exploited a multi-task cascaded convolutional networks-based detector for fruit detection. We chose papaya for our study and collected more than one hundred images and labeled them. Alongside this, we also added an appropriate number of supplementary images. Furthermore, we proposed a novel augmented method. The comparative experiment results demonstrated that this augmented method can improve the final result. The dataset for training was obtained from the dataset, which contains more than hundreds of images. Our results showed that the detector can conveniently adapt to other kinds of fruit as well. Finally, we tested the detector on 4 groups of images with different resolutions. Each group had thirty-five images. The average time cost of the detector was less than 14 seconds per thirty-five images, which is very close to the real-time response.

In the future with this work, the detection of unripe, ripe, and rotten papaya can be done by adding those images of papaya.

References

[1] Faostat F. Agriculture Organization of the United Nations Statistics Division 2014.

[2] D.G. Savakar, B.S. Anami Grading of bulk food grains and fruits using computer vision J Agric Eng Biotechnol, 3 (1) (2015), pp. 1-10

[3] J.L. Damez, S. Clerjon Meat quality assessment using biophysical methods related to meat structure Meat Sci, 80 (1) (2008), pp. 132-149

[4] L.S. Magwaza, U.L. Opara Analytical methods for determination of sugars and sweetness of horticultural products a review Sci Hortic, 184 (2015), pp. 179192

[5] Medina JD, Gutiérrez GV, García HS. Pawpaw: Post-harvest operation. Compendium on Post-harvest Operations.

[6] S. Marimuthu, S.M. Roomi Particle swarm optimized fuzzy model for the classification of banana ripeness IEEE Sens J, 17 (15) (2017), pp. 4903-4915

[7] D.S. Prabha, J.S. Kumar Assessment of banana fruit maturity by image processing technique J Food Sci Technol, 52 (3) (2015), pp. 1316-1327

[8] C.S. Nandi, B. Tudu, C. Koley A machine vision-based maturity prediction system for sorting of harvested mangoes IEEE Trans Instrum Meas, 63 (7) (2014), pp. 1722-1730

[9] C. Liu, W. Liu, X. Lu, F. Ma, W. Chen, J. Yang, *et al*. Application of multispectral imaging to determine quality attributes and ripeness stage in strawberry fruit PLoS ONE, 9 (2) (2014), pp.1 -8

[10] P. Leekul, S. Chivapreecha, C. Phongcharoenpanich, M. Krairiksh Rician kfactors-based sensor for fruit classification by maturity stage *IEEE* Sens J, 16 (17) (2016), pp. 6559-6565

[11] H. Kaur, B.K. Sawhney, S.K. Jawandha Evaluation of plum fruit maturity by image processing techniques J Food Sci Technol, 55 (8) (2018), pp. 3008-3015

[12] Bo Li, J. L. (2018). Advances in Non-Destructive Early Assessment of Fruit Ripeness towards Defining Optimal Time of Harvest and Yield Prediction. *plants*.

[13] Santi Kumari Behera, Amiya Rath, Prabira kumar Sethy(2020-21). Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *ScienceDirect*.

[14] LI ZHANG, G. G. (2019). Multi-Task Cascaded Convolutional Networks Based Intelligent Fruit Detection for Designing Automated Robot. *IEEE*.

[15] Islam, M. A., Islam, M. S., Hossen, M. S., Emon, M. U., Keya, M. S., & Habib, A. (2020). Machine Learning-based Image Classification of Papaya Disease Recognition. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)

PAPAYA RIPENESS PREDICTION USING MACHINE LEARNING

ORIGINALITY REPORT

29%	%	%	29%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	17%
2	Submitted to Motilal Nehru National Institute of Technology Student Paper	2%
3	Submitted to University of Warwick Student Paper	1%
4	Submitted to Somaiya Vidyavihar Student Paper	1%
5	Submitted to Coventry University Student Paper	1%
6	Submitted to University of Salford Student Paper	1%
7	Submitted to Greater Noida Institute of Technology Student Paper	1%
8	Submitted to Visvesvaraya Technological University, Belagavi Student Paper	1%