FRUITS QUALITY EVALUATION

A BTP Report

by

Mendem Mani Tej

Vanam Shankar Sreenu

Thalapally Pavan Kumar

Roll Nos.:

S20180010104

S20180010186

S20180010174

INDIAN INSTITUTE OF INFORMATION

TECHNOLOGY SRICITY

17-05-2021

1st Semester Report

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY SRICITY

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the BTP entitled "Fruits Quality Evaluation" in the partial fulfillment of the requirements for the award of the degree of B. Tech and submitted in the Indian Institute of Information Technology SriCity, is an authentic record of my own work carried out during the time period from January 2021 to May 2021 under the supervision of Prof. R. Shathanaa, Indian Institute of Information Technology SriCity, India.

The matter presented in this report has not been submitted by me for the award of any other degree of this or any other institute.

Signature of the student with date

Mendem Mani Tej - 17/05/21

Vanam Shankar Sreenu - 17/05/21

Thalapally Pavan Kumar - 17/05/21

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of BTP Supervisor with date

Dr. R. Shathanaa - 17/05/21

ABSTRACT

Grading of the fruits is one of the most important post-harvest tasks that the fruit processing agro-industries do. India is the second largest producer of fruits after China. Due to the lack of skilled workers, lot of the harvested fruits is wasted. Agro-industries employ expert laborers for grading and separating the fruits based on their quality. This is expensive and time consuming and classification and grading of fruits not done precisely. This motivated us to impose the automation system which will help the fruit industry.

Also, recognizing different quality of fruits is a repeated chore in supermarkets, where the cashier has to define each item type that will determine its cost. It may sometimes be inaccurate and people aren't satisfied by the classification done by the cashier. So, we want to develop a machine-based classification and grading which will be more accurate.

Our aim is to determine fruits quality and differentiate the fruits according to its quality. Quality of the fruit is determined on the basis of the features such as color, maturity, shape and size of the fruit. Choosing appropriate neural network for classification is a big task. With the help of DL we can come up with designing an automation system to differentiate the fruits according to its quality.

Contents

INTRODUCTION

One of the largest economic sectors is agriculture and it plays a vital role in the economic growth of India. In the agricultural and food industry the proper grading of fruits is very important to increase the profitability.

Fruits comprise an essential part of the human diet as they are the primary source of dietary nutrients. Although manual sorting and grading can be done for the quality assessments and freshness of fruit detection, the process will be slow and sometimes it will be full of errors. Large amount of time and money is wasted in the fields for checking the freshness of the fruits by humans.

Hence, classification of fruit freshness is important, for increasing the market share and establishing better quality standards. If the quality measures are mapped into an automated system by using suitable DL methodologies then the work will be faster and with minimum error.

LITERATURE SURVEY

Table - 1

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Anuja Bhargava and Atul Bansal (2019)	Fruits and Vegetables Quality Evaluation Using Computer Vision [1]	• Image Acquisition and Pre - Processing • Image Segmentation (thresholding) • Feature Extraction (color, texture, skin defects, size and shape) • Classification (ANN, CNN, SVM,.etc)	True positives (TP) - Predicted positive and are actually positive. False positives (FP) - Predicted positive and are actually negative. True negatives (TN) - Predicted negative and are actually negative. False negatives (FN) - Predicted negative and are actually positive. Accuracy, Calculated as, (TN+TP) / (TN+TP+FN+FP)	Uneven distribution of light on arch surface, powerful wavelength selection for different application, surface assessment, lengthy time exhausting of acquisition and processing for spectral image etc.

Table - 2

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Chandra Sekhar Nandi, Bipan Tudu, and Chiranjib Koley, Member, IEEE (2016)	A Machine Vision Technique for Grading of Harvested Mangoes based on Maturity and Quality [2]	Maturity (actual- days-to-rot) Estimation using SVM Quality evaluation by considering size, shape and surface defects. Grade evaluation using Fuzzy incremental learning algorithm based on maturity and quality level.	True positives (TP) False positives (FP) True negatives (TN) False negatives (FN) Confusion Matrix = [TP FN FP TN]	They have considered only one side of mango image. It failed to work if other side has some surface defects, marks etc.

Table - 3

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Nashwa El-Bendary, Esraa El Hariri, Abdul Ella Hassanien, Amir Badra (2015)	Using machine learning techniques for evaluating tomato ripeness [3]	• Resize images. • Background of each image is removed with some morphological operations. • Each image is converted from RGB to HSV color space. • PCA algorithm is applied as a features extraction technique in order to generate a feature vector. • Finally, for classification, algorithms like SVMs are implemented for classification of ripeness stages.	ROC Curves. True positive rate is - (TP/(TP + FN)) False positive rate is - (FP/(TN + FP)) ROC AUC is just the area under the curve, the higher its numerical value the better.	Lack of Larger dataset lead to less accuracy as accuracy of SVMs increases by increasing the number of images per training class.

Table - 4

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Santi Kumari Behera, Amiya Kumar Rath, Abhijeet Maha patra, Prabira Kumar Sethy (2020)	Identification, classification & grading of fruits using machine learning & computer intelligence [4]	1. Fruit image as Input 2. Segmentation (thresholding, k-means clustering.) 3. Feature Extraction (Color, shape, size.) 4. Classification (ANN,SVM,CNN.)	True positives (TP) False positives (FP) True negatives (TN) False negatives (FN) Classification Accuracy (CA) is Calculated as, TN+TP/ TN+TP+FN+FP	The lighting conditions made some difference in the recognition and grading. Only considered the one-side image of fruit.

Table - 5

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Aniket Harsh, Kishan Kumar Jha, Shashwat Srivastava,Abhinav Raj, Raghav (June 2020)	FRUIT FRESHNESS DETECTION USING CNN APPROACH [5]	A.Preprocess class labels. B.Compile model with stochastic gradient descent optimizer and categorical-cross entropy C.Fit and train data. D.Evaluate model on test dataset and calculate the confusion matrix. E.Convolutional Neural Network (CNN) is employed	True positives (TP) False positives (FP) True negatives (TN) False negatives (FN) Confusion Matrix is [TP FN FP TN]	Quality detection of fruit samples with an unknown fruit Is inaccurate

Table - 6

AUTHOR AND YEAR	TITLE	METHODOLOGY	EVALUATION PARAMETERS	DRAWBACKS
Shuxiang Fan, Jiangbo Li, Yunhe Zhang, Xi Tian,Qingyan Wang, Xin He, Chi Zhang, Wenqian Huang. (2020)	On line detection of defective apples using computer vision system combined with deep learning methods [6]	• Input image is taken. • Convolution Neural Network. • Conv-layer, batch-normalization, max-pooling. • For Classification, softmax layer is used.	True positives (TP) False positives (FP) True negatives (TN) False negatives (FN) Classification Accuracy (CA) is Calculated as, TN+TP/ TN+TP+FN+FP	Only for one fruit apple.

METHODOLOGY

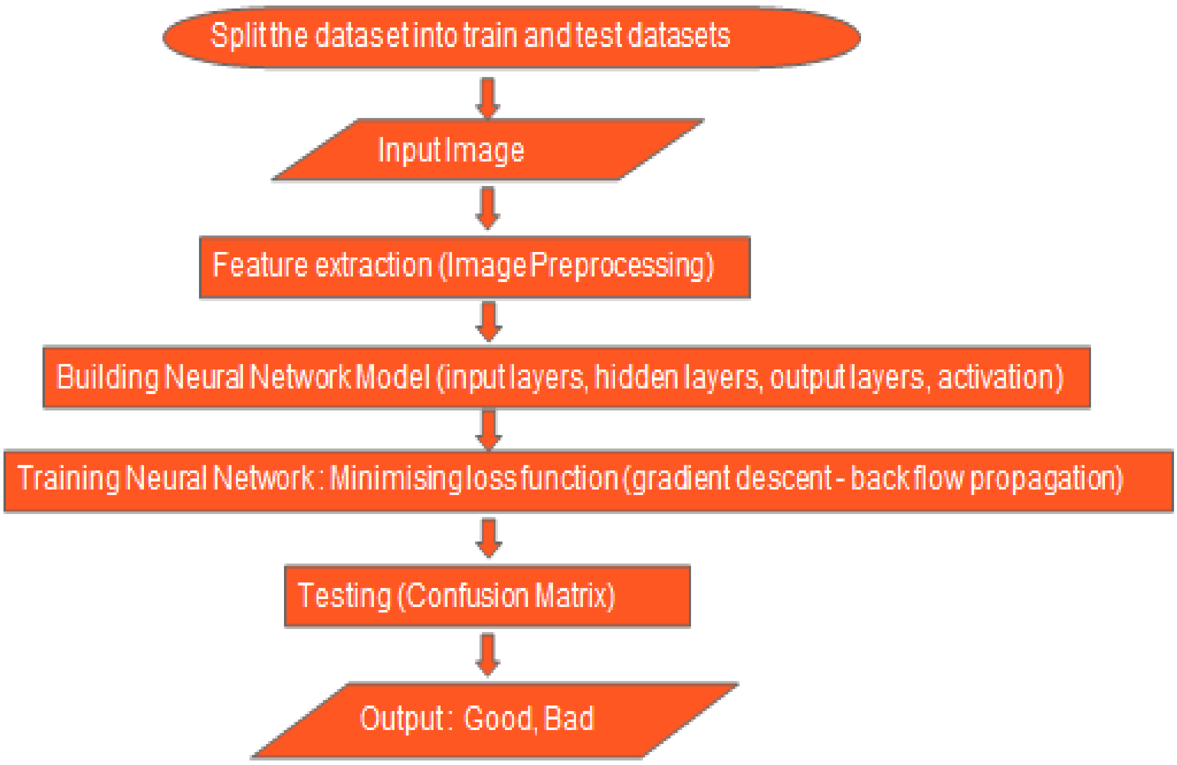


Figure - 1 - Methodology (Flow-diagram)

DNN Approach

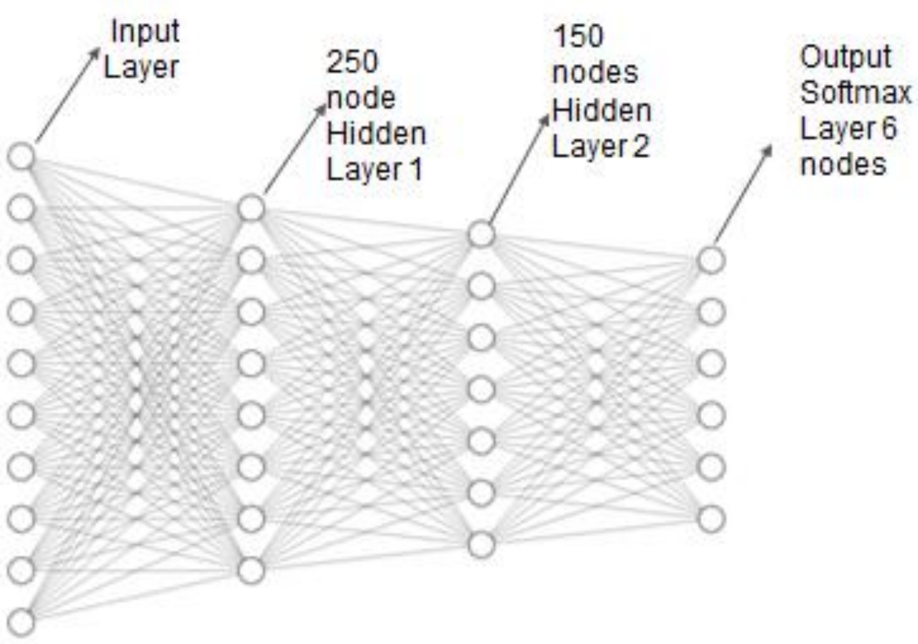


Figure - 2 - NN Architecture

NN Model

Table - 7

Hidden Layer 1 = 250 nodes
Hidden Layer 2 = 150 nodes
Output Layer = 6 nodes
Activation Function - Sigmoid = 1/(e^-x+1)
Loss Function - Cross Entropy Loss Function (L= -Σylog(x))
Learning rate = 0.1
Epochs = 1700
Optimization Technique Used - Gradient Descent

Forward Propagation

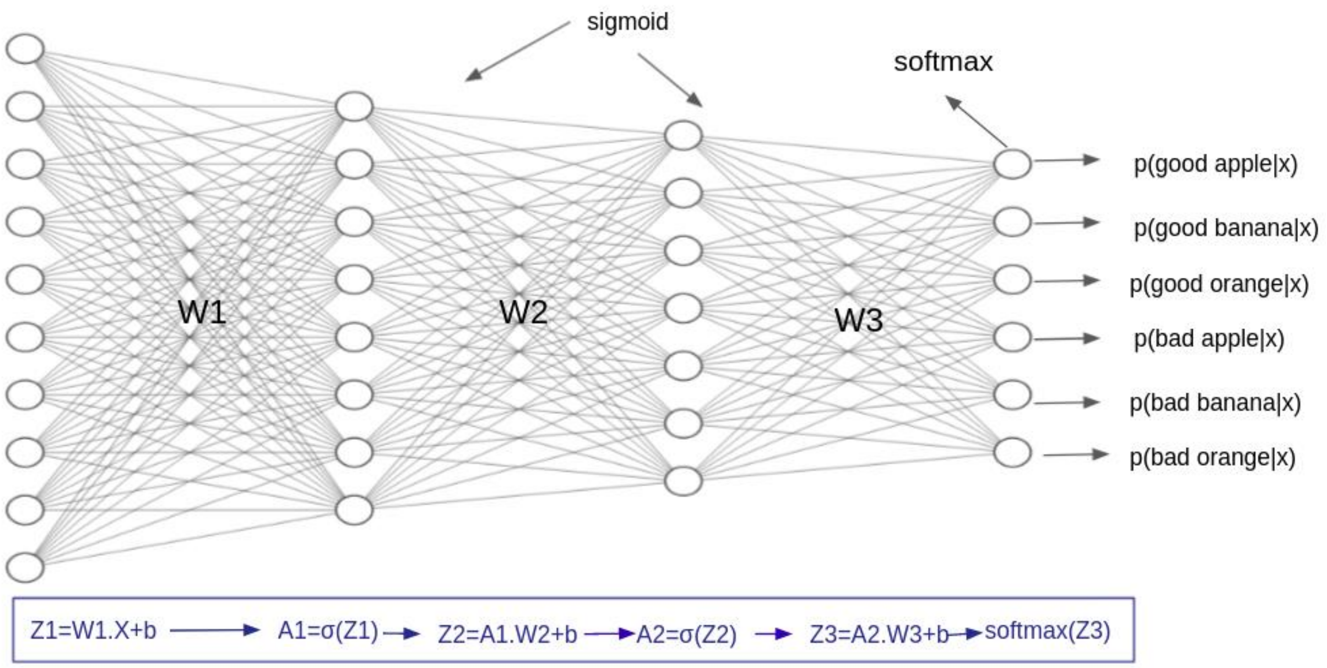


Figure - 3 - Forward Propagation

Back Propagation

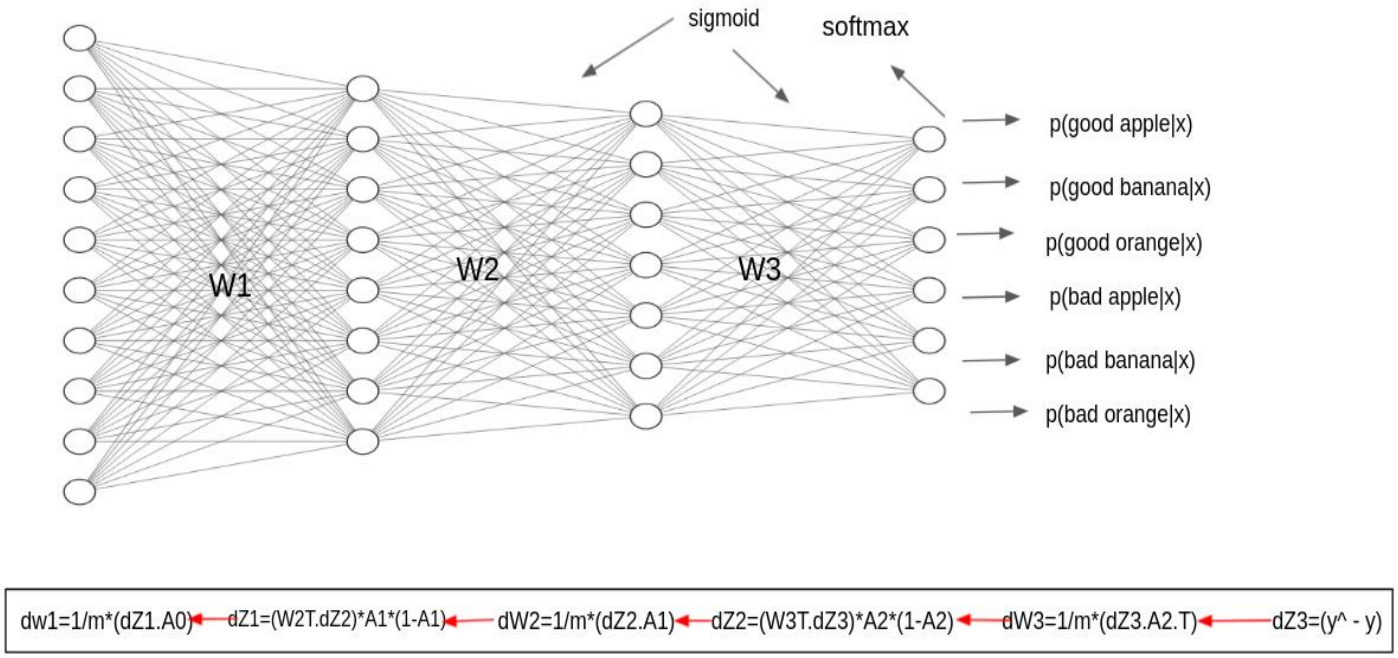


Figure - 4 - Backward Propagation

CNN Approach

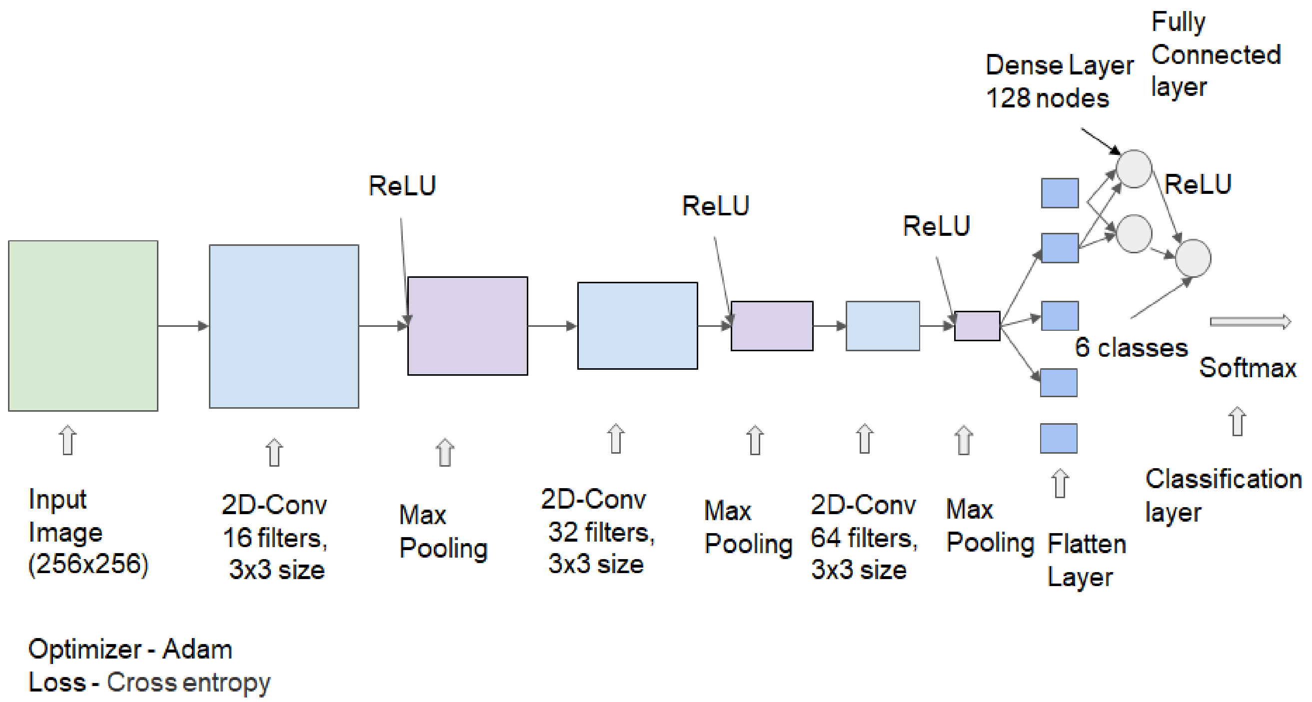


Figure - 5 - CNN Architecture

CNN Model

Table - 8

Layer 1
conv-2d(kernel): padding - ‘same’ (as input dimensions), stride - 1
Filters: 16 (3x3 size)
Activation - ReLU
max pooling: no-padding, pool-size - 2 x 2, stride - 2

Layer 2
conv-2d(kernel): padding - ‘same’ (as input dimensions), stride - 1
Filters: 32 (3x3 size)
Activation - ReLU
max pooling: no-padding, pool-size - 2 x 2, stride - 2

Layer 3
conv-2d(kernel): padding - ‘same’(as input dimensions), stride - 1
Filters: 64 (3x3 size)
Activation - ReLU
max pooling: no-padding, pool-size - 2 x 2, stride - 2

Layer 4
Dense layer - 128 nodes

Layer 5
Classification Layer - Softmax - 6 nodes

Learning rate – 0.001
1 epoch = 500 steps
Epochs – 10 epochs (5000 iterations)

RESULTS

NN Results

Table - 9

Dataset Name	Total Training Images	Total Validation Images
Kaggle (Apple, Banana, Orange)	10901	2698

NN Train Results

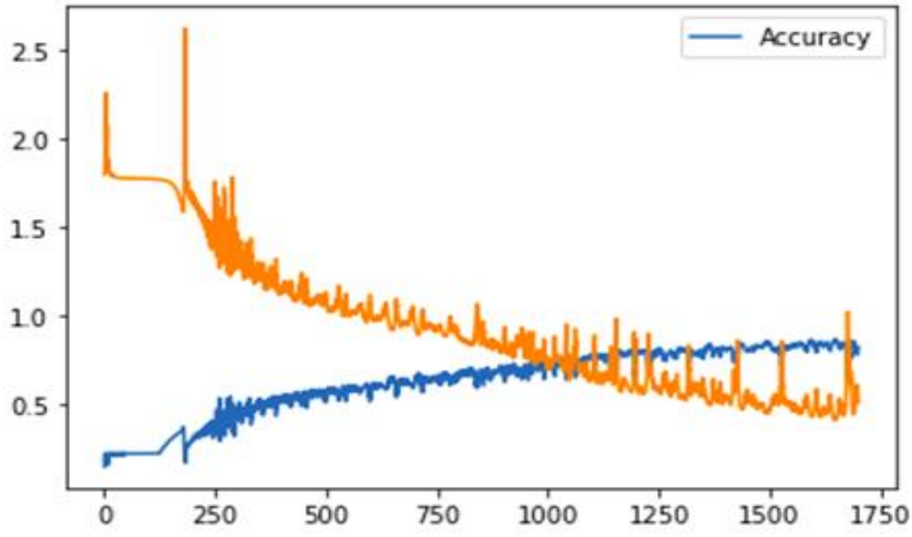
Table - 10

- Total no of correctly classified = 8904 / 10901
- Total no of wrongly classified = 1997 / 10901
- Accuracy = 81.68057976332447

NN Test Results

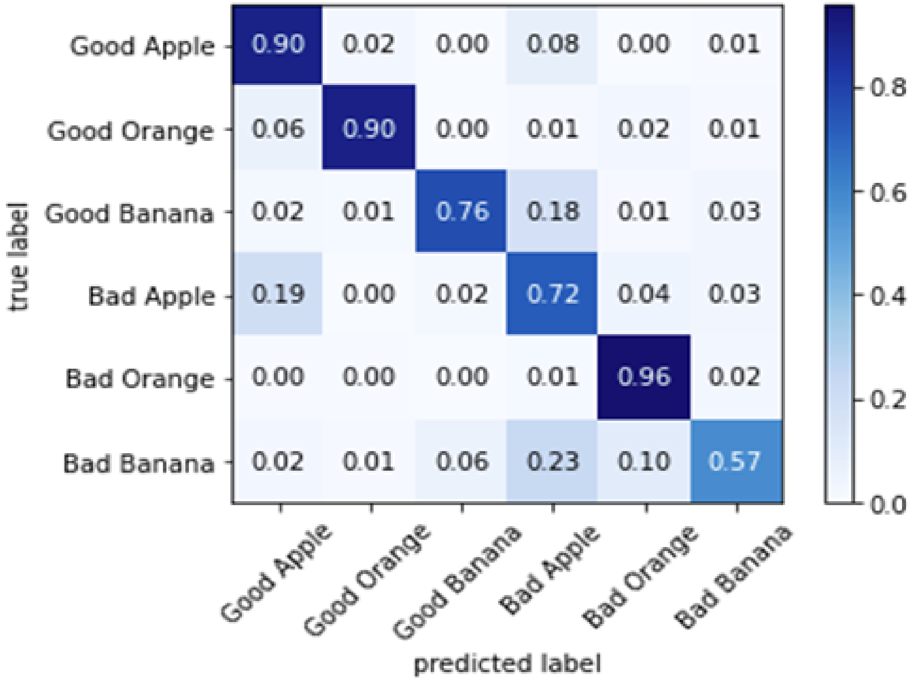
Table - 11

- Total no of correctly classified = 2165 / 2698
- Total no of wrongly classified = 533 / 2698
- Accuracy = 80.24462564862861



This graph shows the accuracy vs loss of NN model. We can see that as number of epochs increased accuracy of model has increased and loss has decreased simultaneously.

Figure - 6 - Loss vs Accuracy



This figure shows the confusion matrix of NN model. Diagonal elements represent the precision.

Figure - 7 - NN Confusion Matrix

CNN Results

Table - 12

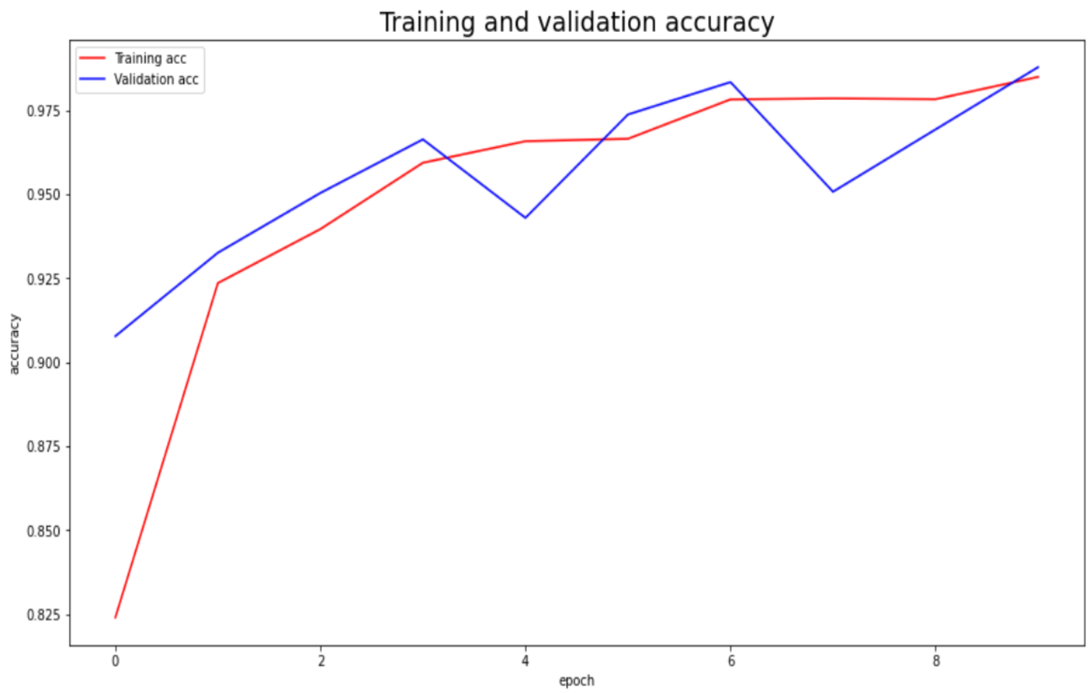
Dataset Name	Total Training Images	Total Validation Images
Kaggle (Apple, Banana, Orange)	10901	2698

CNN Accuracy Table -13

Training accuracy - 99.43% Validation accuracy - 98.78%

CNN Loss Table -14

Training loss - 1.66% Validation loss - 3.33%



This graph shows training and validation accuracy of CNN model. We can see that as epochs increased model training and validation accuracy has increased.

Figure - 8 - Training and Validation Accuracy

This graph shows training and validation loss of CNN model. We can see that as epochs increased model is training and minimizing the loss.

Figure - 9 - Training and Validation Loss

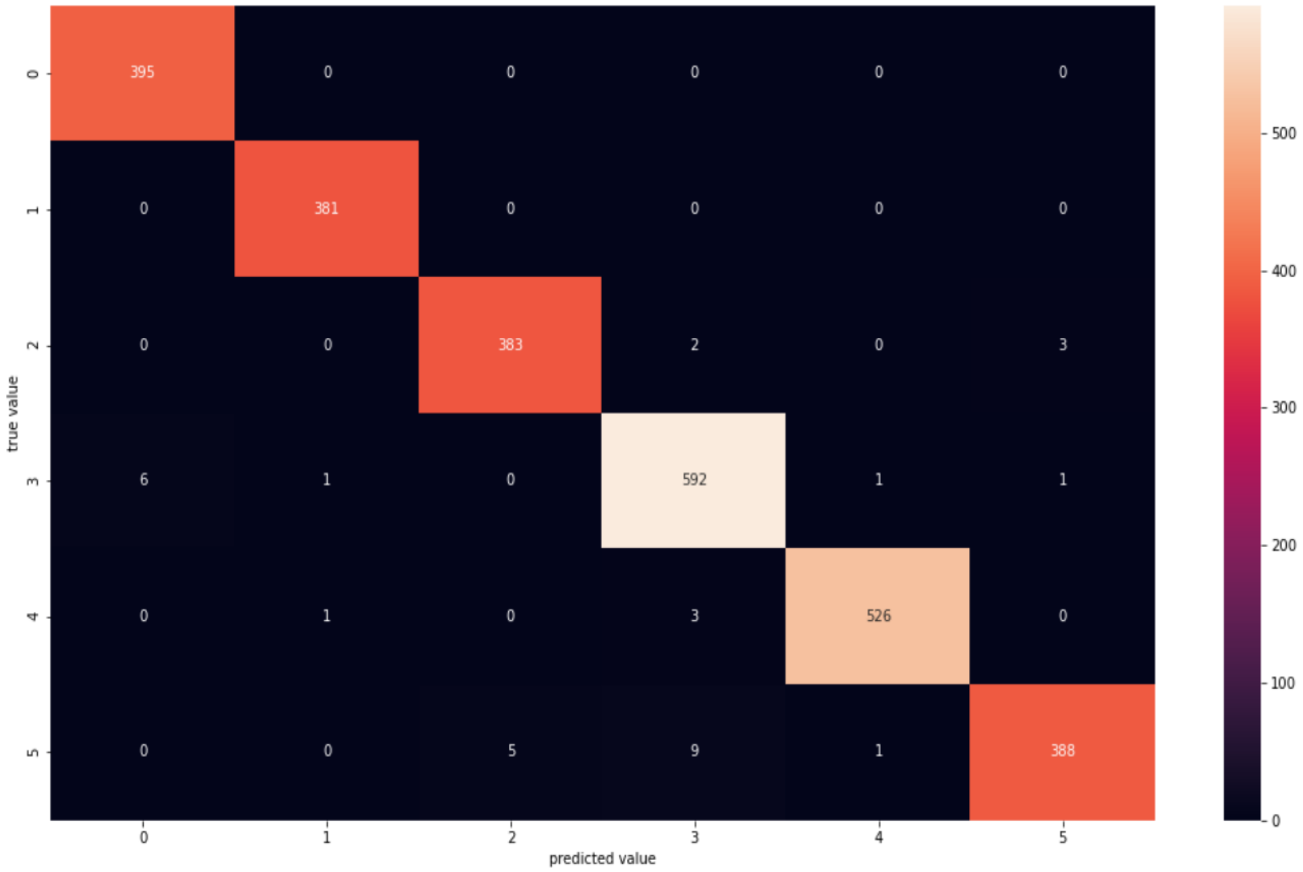


Figure - 10 - CNN Confusion Matrix

This figure shows confusion matrix of CNN model. Diagonal elements represent correctly classified images and rest of the elements in this matrix represent misclassified images.

CONCLUSION

CNN VS ANN

- CNN is faster than ANN.
- CNN has higher accuracy than ANN.

CNN VS ANN

Table - 15

CNN	ANN
Training Accuracy - 99.43	Training Accuracy - 81.6
Validation Accuracy - 98.78	Validation Accuracy - 80.2

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

CF - Confusion Matrix

TP - True positives

FP - False positives

TN - True negatives

FN - False negatives

SVM - Support Vector Machine

ROC - Receiver Operating Characteristic Curve

AUC - Area Under the ROC Curve

PCA - Principal Component Analysis

HSV - Hue, Saturation and Value

RGB - Red Green Blue

CNN - Convolutional Neural Network

ANN - Artificial Neural Network

DNN - Deep Neural Network

ReLU - Rectified Linear Unit

ELU - Exponential Linear Unit

REFERENCES

1. Anuja Bhargava and Atul Bansal (2019), Journal of King Saud University, Fruits and Vegetables Quality Evaluation Using Computer Vision

2. Chandra Sekhar Nandi, Bipan Tudu, and Chiranjib Koley, Member, IEEE (2016), IEEE Sensors Journal, A Machine Vision Technique for Grading of Harvested Mangoes based on Maturity and Quality.

3. Nashwa El-Bendary,Esraa El Hariri,Abdul Ella Hassanien, Amir Badra (2015), ElseVier, Using machine learning techniques for evaluating tomato ripeness.

4. Santi Kumari Behera, Amiya Kumar Rath, Abhijeet Mahapatra, Prabira Kumar Sethy (2020), Journal of Ambient Intelligence and Humanized Computing Identification, classification & grading of fruits using machine learning & computer intelligence.

5. Aniket Harsh, Kishan Kumar Jha, Shashwat Srivastava,Abhinav Raj, Raghav, (2020), International Research Journal of Modernization in Engineering Technology and Science, FRUIT FRESHNESS DETECTION USING CNN APPROACH.

6. Shuxiang Fan, Jiangbo Li, Yunhe Zhang, Xi Tian,Qingyan Wang, Xin He, Chi Zhang, Wenqian Huang, (2020),Journal of food engineering, On line detection of defective apples using computer vision system combined with deep learning methods.

OUR CODE IMPLEMENTATIONS

NN

https://github.com/bt21rs02/BTP/blob/main/Deep_Neural_Net_N_layered.ipynb

CNN

https://github.com/bt21rs02/BTP/blob/main/CNN_Apples_Banana_Orange_Tensorflow.ipynb