MNet: A Framework to Reduce Fruit Image Misclassification

Vishal A. Meshram[1*], Kailas Patil[1], Sahadeo D. Ramteke[2]

[1] Department of Computer Engineering, Vishwakarma University, Pune 411048, India
[2] ICAR-National Research Center for Grapes, Manjri Farm P.O., Pune-Solapur Road, Pune 412307, India

Corresponding Author Email: vismeshram1980@gmail.com

ABSTRACT

Fast and accurate fruit classification is a major problem in the farming business. To achieve the same, the most popular technique used to build a classification model is "Transfer Learning", in which the weights of pretrained models are used in a new model to solve different but related problems. This technique assures the fast model building with a reduction in generalization error. After testing a popular image classification models namely, DenseNet161, InceptionV3, and MobileNetV2 on created dataset in which a "misclassification" is observed as a major problem which is overlooked by many researchers. This paper proposed a novel framework called "MNet: Merged Net" which not only improves the accuracy, but also addresses the misclassification problem. In this framework, the fruit classification problem is divided into small problems and build a separate model for each. In the final stage, the results of these models are combined. Two models called as FC_InceptionV3 (Fruit Classification InceptionV3) and MFC_InceptionV3 (Merged Fruit Classification InceptionV3) are created based on popular pretrained model InceptionV3. MFC_InceptionV3 is based on proposed framework. In this work, a dataset consisting of 12000 color images of top fruits in India with "Good" and "Bad" quality labels was created and published. The dataset consists of a total of 12 classes. The proposed framework MNet is tested on the most popular deep learning model called InceptionV3. The results of InceptionV3, FC_InceptionV3, and MFC_InceptionV3 are compared. The experimental results shows that the MFC_InceptionV3 model achieved 99.92% accuracy and moderates the image misclassification problem.

1. INTRODUCTION

India stood at the second position in the world in fruit producing after China. It is also ranked top in the list of countries who export fruits. In agriculture market as a whole, fruit markets has the highest economical weightage / share in the view of international or local market. Table 1 shows the statistics of the top fruits exported from India [1], along with the revenue generated from its sale. Thus while building our dataset high revenue generating fruits were considered. Farmers produce fruits, and then they either export the production to the international markets, or sell it to industries or to local market vendors. The fruits are processed in industry and different products like juice, pickles, jam, etc. are made which are either exported or sold at the market. Local market vendors sell the products to either small shopkeepers, supermarkets or directly to the customers. Thus, the three major categories of stakeholders for agriculture products are farmers, industries, and end-consumers or customers. Good fruit quality is the basic requirement of each stakeholder. The profit of business directly depends on the quality of fruit. Choosing good fruits in less time with high accuracy is the most vital requirement of the fruit business. The solution to this requirement must be user friendly and low cost which is affordable by all stakeholders of the agriculture products.

In India, fruit sorting is still being done manually, as shown in Figure 1. The manual forms of fruit classification/sorting/ grading can be very grueling, tiresome, back aching, and error prone, which directly affect the profit margin of stakeholders. Thus, by introducing technology to assist in fruit sorting as per the quality of grading, the manual stress required for the same task can be reduced and the efficiency will be increased. Every stakeholder at this end wants the good quality fruit and ignores the bad quality fruits. A good and bad quality fruit can be defined as: Good quality Fruit: a fruit having the required qualities or beneficial to stakeholders can be categorized as good quality fruits. Bad quality fruit: fruit of poor quality or low standard which is not desirable by stakeholders can be categorized as bad quality fruits. Consumers always desire the value of money while purchasing fruits from market.

The quality attributes of fresh fruits can be classified into three classes based on the occurrence of product characteristics as external, internal, and hidden respectively [2]. In the external class appearance (sight), feel (touch), and defects are considered as a quality attribute. In the internal class odor, taste, and texture are considered as a quality attribute. Many research works have been published on fruit classification based on internal and external quality attributes of fruits. The literature survey shows that systems based on internal classification are mostly sensor-based, complex, and costly because of extra hardware like NIR sensors.

Table 1. Top fruits exported from India

Sr No	Name of the Fruit	Year	Export Details		
			Qty (in MT*)	Value (In Rs. Crore)	Value (in US $ million)
1	Apple	2017-18	12,529.22	39.15	6.07
		2018-19	16,744.64	77.78	10.85
		2019-20	21,182.09	75.30	10.58
2	Banana	2017-18	1,01,314.37	348.79	54.07
		2018-19	1,34,503.45	412.09	59.35
		2019-20	1,95,745.87	658.58	91.59
3	Grapes	2017-18	1,88,221.16	1,899.97	294.63
		2018-19	2,46,133.79	2,335.24	334.79
		2019-20	1,93,690.54	2,176.88	298.05
4	Guava	2017-18	1,229.75	8,58,897.00	0.85
		2018-19	956.68	6,67,784.00	0.64
		2019-20	1,697.14	7,31,600.00	0.70
5	Lemons & Limes	2017-18	17,470.20	54.79	8.48
		2018-19	21,121.35	44.57	6.43
		2019-20	14,485.90	52.79	7.37
6	Mango	2017-18	49,180.46	382.31	59.28
		2018-19	46,510.27	406.45	60.26
		2019-20	49,658.67	400.21	56.11
7	Orange	2017-18	15,840.48	34.84	5.42
		2018-19	43,098.31	247.94	34.73
		2019-20	93,749.42	253.06	35.12
8	Pomegranate	2017-18	47,335.74	537.73	83.47
		2018-19	67,891.83	688.47	98.98
		2019-20	80,547.74	687.50	96.18

*MT: Metric Ton

(a)

(b)

Figure 1. Manual sorting of pomegranate in the wholesale fruit market, Pune, Maharashtra, India

Most recently, Ünal [3] and Meshram et al. [4] show that the emphasis has been on the development of computer vision systems based on machine learning and deep learning algorithms to solve problems in the agriculture field. Mostly these systems consider the external quality attributes like size, shape, gloss, and color for fruit classification. These systems have promising results; they are less complex, user friendly, and fast. In this work, the external parameters of fruits are considered for classification.

Singh et al. [5] presented a multilayer convolution neural network (MCNN) to classify the mango tree leaves infected by the anthracnose fungal disease with creating an own dataset

consisting of 1070 images of mango tree leaves. The obtained model achieved 97.13% accuracy. Kour and Arora [6] proposed a novel model called Fuzzy Rule-Based Approach for Disease Detection (FRADD) for apple fruit disease detection (called as apple scab) and classification. The Apple image dataset was created using a digital camera with different angles and distant values. The model achieved 91.66% accuracy. A new simple, effective, and lightweight model called DenseNet-16 was proposed to identify images of citrus diseases and pests in ref. [7]. The author created his own dataset of total images 12,562 (citrus pests = 9051 and citrus diseases = 3510) including seven different types of diseases and 17 species of citrus pests. The classification accuracy of the proposed model was 93.33%. Doh et al. [8] also worked on the automatic detection of citrus fruit diseases. Three machine learning algorithms K-means, SVM, and ANN were used in the proposed solution. The proposed solution uses color and texture features for classification. Hua et al., [9], presented a detailed review on latest advances in robots used for fruit harvesting. Use of robots for harvesting helps to increase the production and reduces the extra costs for apples, kiwifruit, sweet pepper, and tomatoes. Bresilla et al. [10] used a Convolution Neural Network (CNN) model with state-of-the-art, real-time object detection technique called You Only Look Ones (YOLO) for on-tree fruit detection. The authors created his own dataset manually labeling 5000 images of pear and apple fruits. The model achieved more than 90% accuracy. Hossain et al. [11] proposed two models with different frameworks to classify the fruits into 15 and 10 classes, respectively. The first model was built with six layers while the second was fine-tuned visual geometry group-16 pretrained Deep Learning (DL) model. Two datasets were used to evaluate the model in this work. The first dataset used was public and consisted of 2633 images with 15 classes; whereas another was created by using internet images

consisting of total 5946 images with 10 classes. Kirk et al. [12] proposed a Deep Neural Network (DNN) based strawberry fruit detection system which was tested on an existing dataset as well as on the real strawberry dataset. The model was evaluated with its F1 score, the harmonic mean of precision, and recall methods. Whereas Altaheri et al. [13] proposed a machine vision system to categorize the date fruit. Transfer learning from two famous CNN models, AlexNet and VGGNet were used to build three classification models to classify date fruit according to their maturity stage, type, and whether they are harvestable or not. Own dataset was created consisting of 8000 date images of five verities of dates in different maturity and prematurity stages. To conduct the Precision Agriculture (PA) practice with the objective of increasing yield and crop marketability before harvest, Bauer et al. [14] developed a platform that chains up-to-date ML techniques, modern computer vision, and integrated software engineering practices to measure yield-related phenotypes from ultra-large aerial imagery named as AirSurf. This system was modified by combining computer vision algorithms and DL based classification model trained with 100,000 labeled lettuce signals. Zhang et al. [15] developed a harvesting robot for autonomous harvesting which consisted a low-priced gripper and a technique for detection of cutting point. The purpose of the study was to develop an autonomous harvester system which can harvest any crop with a peduncle rather than damaging its flesh. Onishi et al. [16] proposed a new system consisting of Single Shot MultiBox Detector (SSD) and a stereo camera for autonomous detection and harvesting of fruits. The experiment was conducted on an apple tree called "Fuji". Zhang et al. [17] developed two improved model, namely, GoogLeNet and Cifar10, to identify 9 types of maize leaf diseases. A dataset of 500 images used in the experimentation was built from different sources like Plant, Village, and Google website which was divided into nine classes. Augmentation techniques like rotation, grayscale were used before training and testing the dataset. As a result, models GoogLeNet and Cifar10 achieved 98.9% and 98.8% identification accuracy, respectively. Liu et al. [18] proposed a novel pipeline consisting of segmentation, 3D localization, and frame-to--to-frame tracking for accurately counting the fruits from images. FCN (Fully Convolutional Network) was used to obtain the accurate fruit segmentation in the first stage of pipeline. Then the Hungarian technique applied to the Kalman Filter corrected Kanade-Lucas-Tomasi (KLT) tracker to track the fruit through images. In the last stage, they have used SfM for localizing the fruit to correct the counting. This model was evaluated on the orange and apple fruits dataset. For multiple fruit classification, a Pure Convolutional Neural Network (PCNN) was proposed by Kausar et al. [19]. This model was built using seven convolution layers and Global Average pooling (GAP) layer to overcome the problem of over fitting. The performance of PCNN was tested on Fruit 360 dataset and the result shows that PCNN achieved 98.88% fruit classification accuracy. Femling et al. [20] proposed a system to atomize the fruits and vegetables detection process in the retail market. This system was built by using hardware Raspberry Pi and camera Module v2 and ML techniques. Two CNN architectures, Inception and MobileNet, were used to classify 10 different verities of fruits or vegetables.

Zhang et al. [21] studied how the dataset augmentation methods are related to the predication results of the classification results. The training dataset was built by using 200 colour images of five maturity levels, collected during different daylight conditions. Two augmentation operations are proposed to increase the dataset, which will help to raise the final predication accuracy, namely, Geometric Transformation and Random Noise. Authors developed their own CNN based model to classify tomatoes based on ripeness. The model architecture consists of five layers for feature extraction and one fully connected layer for classification. Cheng et al. [22] developed a system for early yield predication with ANN. A new approach was proposed by considering the parameters of fruits and tree canopy while building the model, Two back propagation neural network models for yield predication were developed, one for the early period after June drop and another for the maturing period respectively. Puttemans, et al. [23] developed a robust pipeline by linking an object classification framework with application-specific pre-filtering, scene and proficient cluster segmentation for strawberry and apple fruits detection. The correct output of the trained object detector is always depending on the input image annotations. Annotation biasing is the major problem which can affect the final detection accuracy. Stein et al. [24], developed an efficient yield estimation model using multisensory framework and the DL algorithm. The proposed system was able to detect, track, localize, and map every fruit in a mango farm. The latest DL algorithm called faster R-CNN was used to detect the fruits from images. To solve the problem of occlusion, a multiple viewpoint approach was used which avoids the need for labor intensive field calibration.

Yamamoto et al. [25] developed an on-tree tomato fruit classifier using ML algorithms to classify it into mature, immature, and young tomatoes. Classification decision was based on the size, texture, shape, and colour parameters extracted from the images. Font et al. [26] developed automated fruit harvester which consists of a low-priced camera and a gripper tool consisting of a robotic arm. In (Suchet Bargoti and James Underwood, [27]), a new object detection framework, Faster R-CNN was presented in the paper. The model was tested for detection of mangoes, almonds, and apple fruits in orchards. A tiling approach on images was introduced in the study, which improved the detection accuracy of Faster R-CNN with F1-score greater than 0:9 for mango and apples.

The DL based classifier to classify Cavendish banana grade was explored in Ucat et al. [28]. The model was developed using Python, OpenCV, and Tensorflow. The model achieved more than 90% classification accuracy. Ireri et al. [29] proposed a machine vision system for postharvest tomato grading. The system works on RGB images given as input to the system. Own datasets were created by manually labeling the tomato images into four categories according to their defect, healthy, and ripeness parameters. Four different models were built to classify the image into one of the categories according to the matching features. Total 15 features were considered while taking the decision. The result shows that RBF-SVM performed well compared to others for category 1 i.e. healthy or defected with 0.9709 detection accuracy. Piedada et al. [30] developed a system for banana (Musa acuminata AA Group 'Lakatan') classification using ML techniques based on tier-based. A noninvasive tier-based technique was used in this study. ANN, SVM, and RF classifiers were used to classify bananas into extra class, class I, class II, and rejected classes. The result showed that the random forest algorithm outperformed compared to others with 94.2% accuracy. An automated real–time grading system

for quality inspection for apple fruit was developed by Sofu et al. [31]. The developed system comprises a roller, transporter and class conveyor joined with an enclosed cabin with camera, load cell, and control panel units. The system not only classifies the apples on the basis of color, size, and weight parameters, but also identifies defective apples. A grading and sorting system based on machine vision for date fruit was developed by Yousef Al Ohali [32]. The system was able to categorize the dates into three classes (grade 1, 2, or 3) using the given RGB image as an input. A back-propagation algorithm was tested in the study which showed 80% accuracy. Fruits and vegetable quality depends on their parameters like shape, size, texture, color, and defects. Different methods have to apply to classify the fruits and vegetables according to their quality parameters like data collection, preprocessing of data, image segmentation, feature extraction, and finally classification.

The classification methods studied as above have one or more limitations such as:

1. The classification doesn't take into account the needs of the real stakeholders and thus the proposed models are not helpful to farmers, industries, and customers.
2. Few works are either trained on public dataset or images from internet instead of creating own datasets, as in Doh et al. [8], Hossain et al. [11], Onishi et al. [16], Zhang et al. [17], Kausar et al. [19], and Bargoti et al. [27].
3. The available public or private datasets consists of general classes of fruits and do not mention the quality of fruits.
4. Few classification systems are costly as they require extra hardware, as in Femling et al. [20].

To overcome these limitations, this paper presents a framework which makes the model robust, efficient, and accurate for fruit classification. The major contributions of this paper are as follows:

1. Published own dataset of top Indian fruits consists of more than 12000 colour images taken in different light conditions with different angles with quality labels [33].
2. Presented/Studied misclassification problems in prebuilt models.

3. Proposed novel framework to build a model called as MNet is proposed to solve the misclassification problems.
4. A popular pretrained model, InceptionV3 is used to test the framework on random images from Google and the results show that we are able to increase the accuracy and mitigate the problem of misclassification.

The rest of the paper is organized as follows: Section II presents the misclassification problem in popular published models. Section III describes in detail the dataset used for the experimentation. Section IV describes the proposed framework in detail. Section V reports the experimental results. Then discussion and conclusion with a summary and of future work given in section VI and Section VII.

2. PROBLEM IDENTIFICATION

In this work, four pretrained models for image classification are tested which are state-of-art popular and widely used. A Python script is written by using torchvision. Models package in the PyTorch libraries to test these four models, namely, densenet161, inception_v3, mobilenet_v2, vgg19, on our fruit dataset. These models were trained on a very popular dataset called "ImageNet", which consists of 1000 image classes. Out of 1000 classes, only four are matched with our dataset, and those are banana, lime, orange, and pomegranate. Thus, models were tested only on these four classes. Table 2 shows the results of the pretrained models. The Top-1 accuracy considered as in the Table 2.

The result shows that in some cases like lime, orange, and pomegranate, the misclassification problem exists. Misclassification problems can be defined as an act or an instance of wrongly assigning someone or something to a group or category. This problem is severe and may hamper the performance of the system/model in real time. The misclassification may cause financial losses to stakeholders.

Therefore, there is an urgent need to address the problem.

Table 2. Results of Pretrained models tested on own dataset

Sr. No.	Fruit Classes	Name of the Pre-trained Model							
		InceptionV3		VGG19		densenet161		MobileNet V2	
		Correctly classified (%)	Misclassification (%)	Correctly classified (%)	Misclassification (%)	Correctly classified (%)	Misclassification (%)	Correctly classified (%)	Misclassification (%)
1	Banana Bad	31.7	68.3	14.8	85.2	41	59	12.4	87.6
2	Banana Good	99.6	0.4	99.7	0.3	100	0	99.2	0.8
3	Lime Bad	3.5	96.5	20	80	12.6	87.4	22.4	77.6
4	Lime Good	19.9	80.1	8.2	91.8	5.9	94.1	6.3	93.7
5	Orange Bad	42.6	57.4	13.4	86.6	18.3	81.7	24.6	75.4
6	Orange Good	8.1	91.9	2.4	97.6	0.9	99.1	0.3	99.7
7	Pomegranate Bad	79	21	82.7	17.3	93.3	6.7	85.1	14.9
8	Pomegranate Good	44.8	55.2	6.1	93.9	5.9	94.1	9.1	90.9

3. DATASET

The success of a research experiments depends on how large and clean dataset you have. It plays an essential role in success of the research project. It is very difficult to get ready-made dataset which you can use as it is in your project. Therefore, our major goal was to build a dataset that contained many fruit images, where each fruit is represented with many images as possible. Thus to build our own dataset we considered the top fruits exported from India [1] based on

revenue generated, as shown in Table 1. Datasets can be used to solve two major problems, either used for object classification or for object category recognition and detection. For this work, we have created a dataset useful for object classification. The dataset was created in Vishwakarma University campus, Pune, India, in July and August 2019. The images were captured inside and outside the lab, at the very high resolution of either 2448×3264×3 or 3024×3024×3 pixels by using 13+ megapixel mobile rear cameras.

For experimental purpose, the images were down sampled to 256×256×3 resolution. The dataset comprises of six different categories with two subcategories of each and number of images for each class. Figure 2 shows the sample images from the dataset of each class. Images were collected at various times on different days for the same category with natural background. These features improve the dataset variability and represent a more realistic scenario. The images had large variations in quality, lighting, and background. Illumination is one of those variations in imagery, illumination can make two images of the same fruit less similar than two images of different kinds of fruits. The fruit dataset was collected under relatively unconstrained conditions. Images were also captured with different light conditions in the laboratory or outside in the sunlight with different background scenarios. Basic conditions which are considered during dataset collection are:

1. Fruits in different positions.
2. Popular categories of fruits in India for export and consumption.
3. Used high resolution mobile cameras to capture to images.
4. Different lighting conditions with different background.
5. Photos were taken inside and outside the lab.

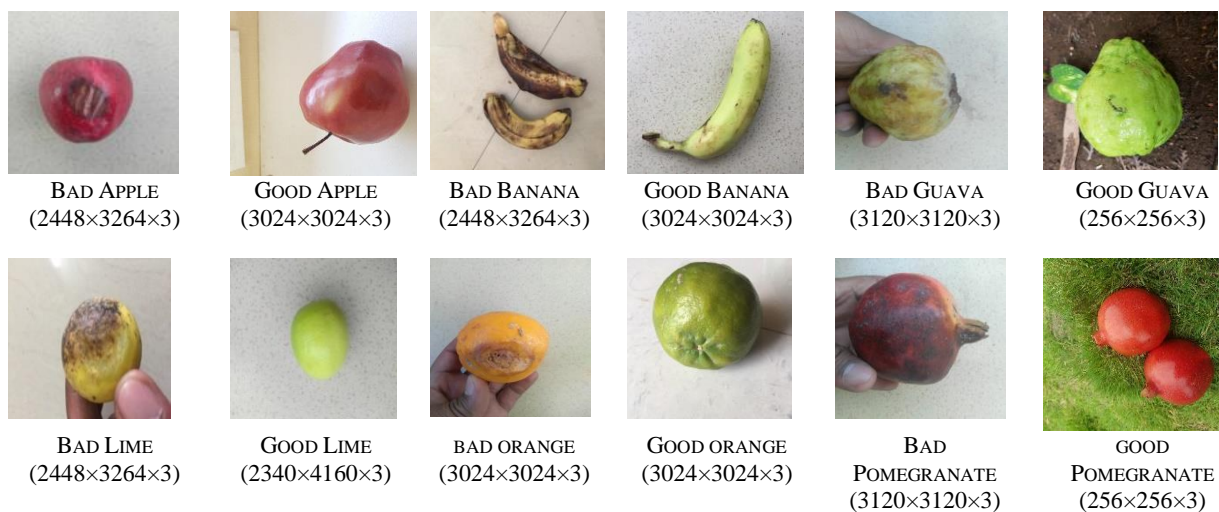Dataset of 12000 images [33] was created and used in the experimentation. ImageDataGenerator API in Keras is used for generating augmented images. This package overcomes the problem of small dataset.

The ImageDataGenerator package also provides functions for image pre-processing. For experimental purpose, the images were down sampled to 256×256×3 resolution. The dataset comprises of six different categories with two subcategories of each and number of images for each class. Figure 2 shows the sample images from the dataset of each class. Images were collected at various times on different days for the same category with natural background. These features improve the dataset variability and represent a more realistic scenario. The images had large variations in quality, lighting, and background. Illumination is one of those variations in imagery, illumination can make two images of the same fruit less similar than two images of different kinds of fruits. The fruit dataset was collected under relatively unconstrained conditions. Images were also captured with different light conditions in the laboratory or outside in the sunlight with different background scenarios.

Basic conditions which are considered during dataset collection are:

1. Fruits in different positions
2. Popular categories of fruits in India for export and consumption.
3. Used high resolution mobile cameras to capture to images.
4. Different lighting conditions with different background.
5. Photos were taken inside and outside the lab.

Dataset of 12000 images [33] was created and used in the experimentation. ImageDataGenerator API in Keras is used for generating augmented images. This package overcomes the problem of small dataset. The ImageDataGenerator package also provides functions for image pre-processing.

| BAD APPLE (2448×3264×3) | GOOD APPLE (3024×3024×3) | BAD BANANA (2448×3264×3) | GOOD BANANA (3024×3024×3) | BAD GUAVA (3120×3120×3) | GOOD GUAVA (256×256×3) |
| BAD LIME (2448×3264×3) | GOOD LIME (2340×4160×3) | BAD ORANGE (3024×3024×3) | GOOD ORANGE (3024×3024×3) | BAD POMEGRANATE (3120×3120×3) | GOOD POMEGRANATE (256×256×3) |

Figure 2. Examples of images from each class with resolution

4. METHODOLOGY

Machine learning (ML) and Deep Learning (DL) are popular and high-performance computing technologies used in computer vision systems for object classification and detection. For the fast and accurate image classification model building, the technique called "Transfer Learning" has been used. Transfer learning assures the fast model building with a reduction in generalization error. There are two ways in which the transfer learning technique can be used, either only weights of the pretrained model or freeze certain layers (add or prune layers) as per requirement and retrain the model on your dataset. This reduces the cost of training and as the model is retrained on, your dataset, quality can be assured. Transfer learning technique is used on a popular pretrained model called InceptionV3 to build our own model. The model is built using both techniques as mentioned to mitigate the misclassification problem. In the proposed research work, our

own dataset consisting of total 12000 images consisting of 12 classes is used. The proposed research work is mainly focusing on how to minimize the image misclassification problem with raising accuracy of the model.

4.1 Transfer learning using CNN pretrained models

A detail survey on transfer learning is presented by Tan et al. [34]. The "Network-based" technique is used in which a partial network (weights, number of layers) of the built model is used to develop a new model. The newly built model is retrained on the dataset again. This approach works well for a small dataset. We built a model called FC_InceptionV3, based on a very popular pretrained model InceptionV3 published by Google [35]. Figure 3 shows the architecture of FC_InceptionV3 model.

The steps for building the model are as follows:

Step 1) Freeze all layers and used only weights and train the model.
Step 2) Used the data augmentation function ImageDataGenerator() for image preprocessing and to overcome the limitation of the small dataset.

train_datagen = ImageDataGenerator (rotation_range=15,
                                    rescale=1./255,
                                    shear_range=0.1,

                                    zoom_range=0.2,
                                    horizontal_flip=True,

width_shift_range=0.1,

        height_shift_range=0.1)
Step 3) add our own layers:
#we added our own layers on last output (only Convolution layers can be added)
        x=layers.GlobalAveragePooling2D()(last_output)
        x=layers.Dense(1024, activation='relu')(x)
        x=layers.Dropout(0.25)(x)
        x=layers.Dense(12, activation='softmax')(x)
Step 4) we freeze 80% layer and retrained the model again

In the first step, weights of "InceptionV3" pretrained model were used to build more accurate model in minimum time. In second step, "data augmentation" technique was used to apply different transformations to original images in the dataset by applying random functions like rotation, shifts, rescale, and flip etc. in each epoch. This will make the model more generalizable and robust. The original InceptionV3 model can classify images into 1000 object categories, whereas for the proposed work the output classes are 12. So, in the third step, new layers are added with "softmax" layer having 12 output classes.

4.2 Transfer learning with divide & conquer method

In the computer branch, the divide and conquer technique is popular and used to solve many problems. The basic idea of divide and conquer is, if a problem posed by the given input is sufficiently simple, it is solved directly, otherwise it is decomposed into independent subproblems, the subproblems are solved, then the results are composed. In simple words, we can say, divide and conquer works on the principle, 'decompose, solve, and compose' [36]. We applied this technique for solving the misclassification problem because of the following reasons: the dataset is small, the dataset consists of similar feature objects which makes it difficult for the model to take a decision, and it makes the solution modular and means you can add or remove the feature from the model anytime. For example: in this work we consider two features, fruit class (i.e., fruit category) and quality class (i.e., good or bad). As per the need of problem, you can add the features like size, shape, etc. or if you want, you can remove any parameters.

Mathematically it is represented like:

Input = Xo
Decomposes into subinputs {x1, x2,…..xn}
=f1(x1) + f2(x2) +….. fn(xn) = {z1,z2,…zn}
Feasible outputs from these subinputs {z1, z2,…..zn}
after composed feasible output with respect to the input
Xo = Zo

A new model called MFC_InceptionV3 based on our proposed framework as shown in Figure 4 is built. The fruit classification with quality problem is divided into two subproblems. First, identification of fruit class and second identify the quality of the fruit. Separate models are built for each problem. In the final stage, the output is merged into a single output. While building the individual model, the same philosophy is used for FC_InceptionV3 model.
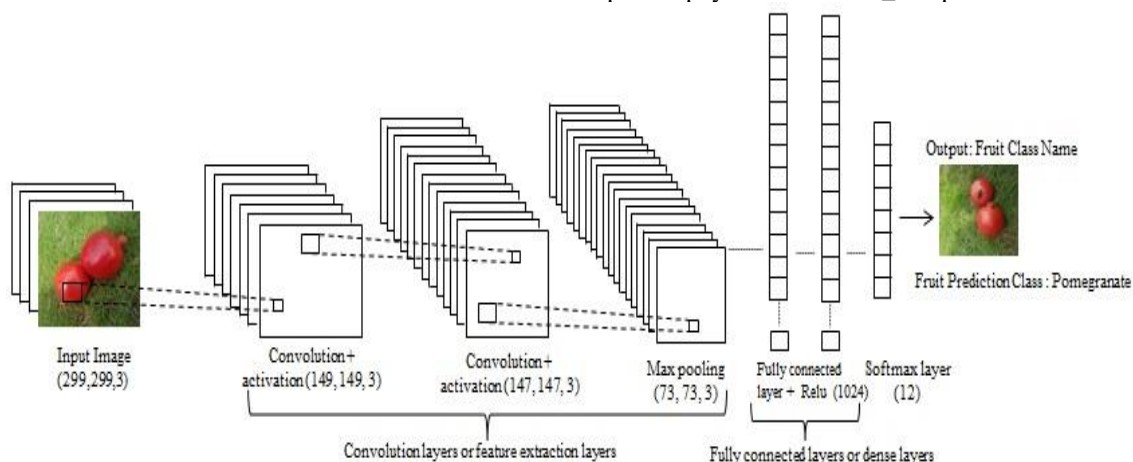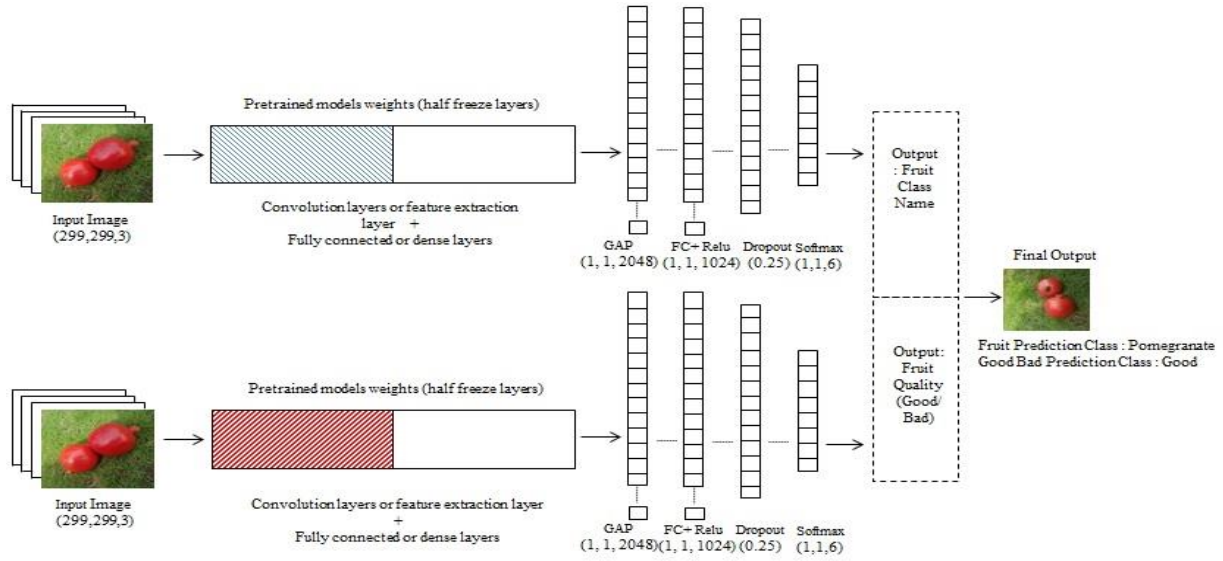

Figure 3. The architecture of FC_InceptionV3 model

Figure 4. The architecture of the MFC_InceptionV3 Model based on MNet: Merged Net framework

5. EXPERIMENTAL RESULTS

The experiments are carried out to evaluate the performance of the model based on our proposed framework. The results of MFC_InceptionV3, which are based on our framework are compared with the FC_InceptionV3 which is built on the standard transfer learning technique. We first prepared a dataset by equally dividing 1000 images in each class. Total 12000 images divided into 12 classes are considered as shown in Figure 5. This assured balanced dataset for the experiment. The Google Colaboratory or "Colab" is used to write and run the experiments. We used GPU support provided by Google Cloud server. We used Keras API written in Python provided for image classification, which runs on top of the TensorFlow machine learning platform. Figures 3 and 4 explains the detail framework designed for fruit classification. Each architecture consists of three parts. The first part is to input images having shape (299, 299, 3). In the second part, we used transfer learning by using pretrained model weights by freezing half layers and retrained the model on our dataset. Here Why do we freeze the layers and then unfreeze again to train the model? Transfer learning is used when a model wants to get the knowledge of the pretrained model and use this knowledge to identify the required objects.
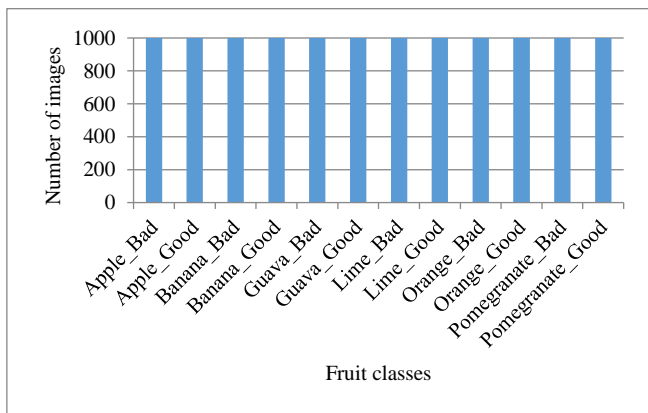
Figure 5. Total 12000 images were equally divided into 12 classes

This head-start of imparting knowledge into the model can be very helpful for the model. It is usually done for models where the dataset is small and the computational power is less, but it is a perfect way to train a model in an efficient manner. In the third part Global Average Pooling (GAP) is used and some dense layers are applied on the top layers along with a dropout of 25% to avoid overfitting. Dense layers are applied with relu and softmax as the activation function to classify the objects.

Global average pooling (GAP) layers are used instead of fully connected layers to avoid the overfitting. GAP layer reduced the parameters in the model which makes it more robust [37].

Out of several activation functions like linear function, Sigmoid Function (σ), Tanh Function; ReLU [38] is one of the most popular functions in many classification tasks. The function is denoted as Eq. (1).

$$F(x) = \max(0, x) \quad (1)$$

where, x represents the input to a neuron. The equation can also be represented as Eq. (2).

$$f(x) = \sum \begin{matrix} 0, \, for \, x<0 \\ x, \, for \, x=>0 \end{matrix} \quad (2)$$

We exploited ReLU in our experiment; to activate it needs only a threshold value. As it doesn't need extra calculations, it is faster and makes the model lighter. Another activation function is called softmax [39], which is very popular and used in all most all multiclass neural network models in the last layer. The function is denoted as Eq. (3).

$$p_j = \frac{\exp(x_j)}{\sum_{i=1}^{k} \exp(x_i)} \quad (3)$$

where, Pj represents the probability of each class, in our case, it is six. This returns the probabilities of each class and the class having the highest probability is our target or output class.

6. RESULTS AND DISCUSSION

6.1 Results of CNN model based on transfer learning (FC_InceptionV3)

The classification results and misclassification of FC_InceptionV3 model are presented in this section. The model presented in Figure 2, is built using the standard way of transfer learning technique based on a popular pretrained model InceptionV3. The model achieved 99.75% accuracy. Confusion matrix is used to represent the performance of the model as shown in Figure 6.

6.2 Misclassification CNN model based on transfer learning (FC_InceptionV3)

To test the FC_InceptionV3 model, we created another dataset which consists of 360 images downloaded from Google. To the best of our knowledge, no dataset is available with quality labels, so we created an own dataset to test the misclassification problem. We equally divided 30 images per class and test the model. A Python script is written to test our model on this dataset. Table 3 shows the results of

FC_InceptionV3 model. But after testing the model in real scenarios it misclassified the fruits even after achieving the 99.53% accuracy. This happened because of the similar features of the fruits. Many fruits like lime, oranges, guava and green apples have similar features like shape, color, or texture etc. because of which the model was misclassifying the fruits as shown in the Figure 7. The FC_InceptionV3 model classified bad apple as Apple_Good, bad guava as Guava Good, and good pomegranate as Pomogranate Bad.

6.3 Results of CNN model based on Mnet: Merged net

Two sub frameworks as shown in Figure 3 for fruit classification and quality are developed, i. e. "Good" or "Bad" classification. The objective of designing two fruit classification algorithms is not only to increase the accuracy but to overcome the misclassification of fruits due to similar features of fruits. The fruit classification model achieved 100% accuracy and the quality model achieved 99.84%. Thus, when the output of these two models merged, the average accuracy obtained is 99.92%. This is higher than FC_InceptionV3 model. Confusion matrix is used to represent the performance of the model as shown in Figure 8 (a, b).
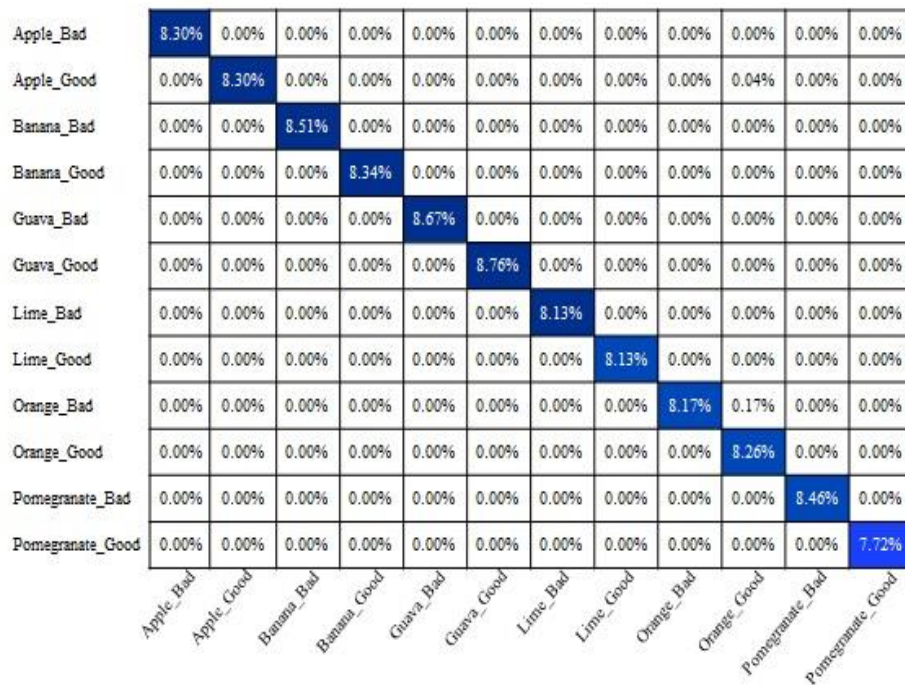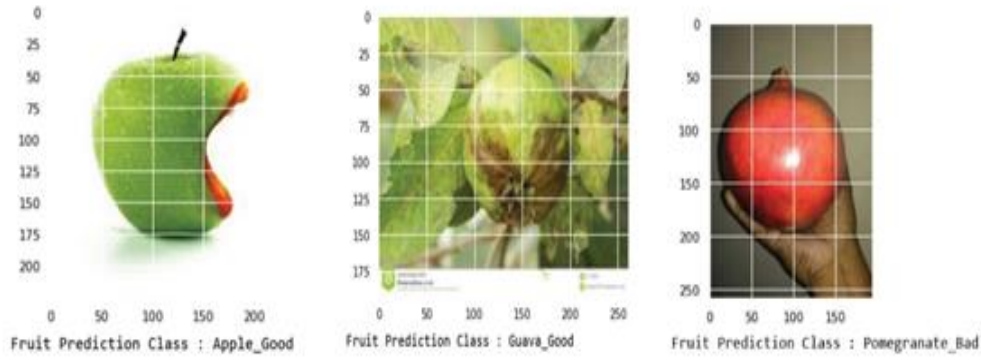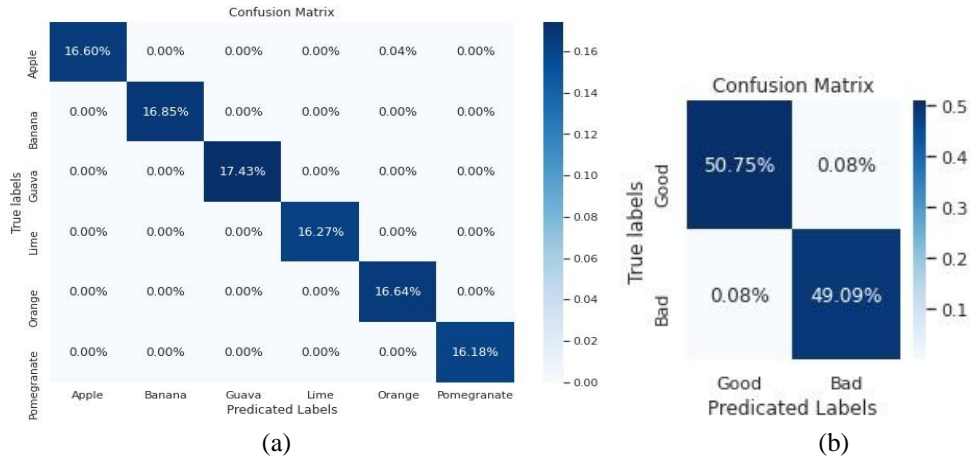

Figure 6. Confusion matrix of FC_InceptionV3 model

Table 3. FC_InceptionV3 misclassification result (Total Images considered 30)

Sr No	Name of the Fruit	FC_InceptionV3 Misclassification result (Total Images considered 30)			
		Total No of Images	Correctly classified (%)	Misclassified images	Misclass-ification (%)
1	Apple Bad	11	36.67	19	63.33
2	Apple Good	28	93.33	2	6.67
3	Banana Bad	29	96.67	1	3.33
4	Banana Good	25	83.33	5	16.67
5	Guava Bad	6	20.00	24	80.00
6	Guava Good	19	63.33	11	36.67
7	Lime Bad	9	30.00	21	70.00
8	Lime Good	11	36.67	19	63.33
9	Orange Bad	17	56.67	13	43.33
10	Orange Good	13	43.33	17	56.67
11	Pomegranate Bad	9	30.00	21	70.00
12	Pomegranate Good	13	43.33	17	56.67
	Average		52.78%		47.22%

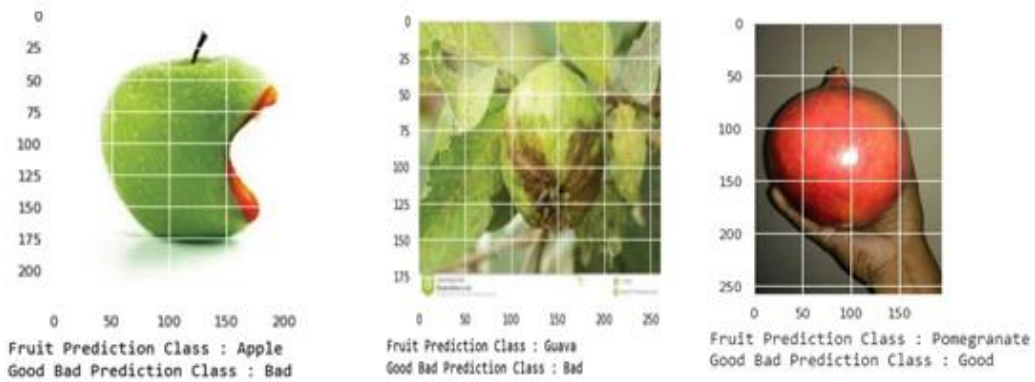**Figure 7.** Sample images misclassification output by whole inception model

(a)   (b)

Figure 8. a) Confusion matrix of fruit class model b) confusion matrix of quality classification model

Figure 9. Sample images correctly classified output by MFC_InceptionV3 model

**Table 4.** MFC_InceptionV3 misclassification result (Total Images considered 30)

Sr No	Name of the Fruit	MFC_InceptionV3 Misclassification result (Total Images considered 30)			
		Total No of Images	Correctly classified (%)	Misclassified images	Misclassification (%)
1	Apple Bad	10	33.33	20	66.67
2	Apple Good	29	96.67	1	3.33
3	Banana Bad	28	93.33	2	6.67
4	Banana Good	26	86.67	4	13.33
5	Guava Bad	10	33.33	20	66.67
6	Guava Good	20	66.67	10	33.33
7	Lime Bad	9	30.00	21	70.00
8	Lime Good	11	36.67	19	63.33
9	Orange Bad	15	50.00	15	50.00
10	Orange Good	14	46.67	16	53.33
11	Pomegranate Bad	13	43.33	17	56.67
12	Pomegr-anate Good	16	53.33	14	46.67
	Average	55.83%		44.17%	

6.4 Misclassification CNN Model based on MNET Framework MFC_InceptionV3

Same testing dataset downloaded from Google as mentioned in section 5.2 is used. Table 4 shows the results of MFC_InceptionV3 model. Because of two separate models the accuracy of the model is increased and the misclassification is reduced by 3.05%. Images which were misclassified by FC_InceptionV3 are correctly classified by MFC_InceptionV3 as shown in Figure 9. It classified the images correctly as apple bad, guava bad, and pomegranate good and so on.

6.5 Comparison of InceptionV3, FC_InceptionV3, and MFC_InceptionV3

Figure 10 shows the comparison of InceptionV3, FC_InceptionV3, and MFC_InceptionV3 in terms of accuracy, correctly classified images, and misclassification. The average of eight classes for result compression is considered to maintain the consistency between all models. Classes namely Banana Bad, Banana Good, Lime Bad, Lime Good, Orange Bad, Orange Good, Pomegranate Bad, and Pomegranate Good are considered. Result tested on our testing dataset consist of the original model i.e. InceptionV3 published by Google claimed 78.1% accuracy on ImageNet dataset is one of the widely used model for image classification. The FC_InceptionV3 model is customised and based on the original model claimed 99.75% accuracy trained on the own dataset. It succeeded to address the misclassification problem reducing it by 2%. MFC_InceptionV3 model is built using MNet architecture based on the original model, claims 99.92% accuracy, and achieved success to reduce the misclassification problem, reducing it by 6% as compared to the original model and 4% compared to FC_InceptionV3 Model. The comparison is shown in Figure 10.
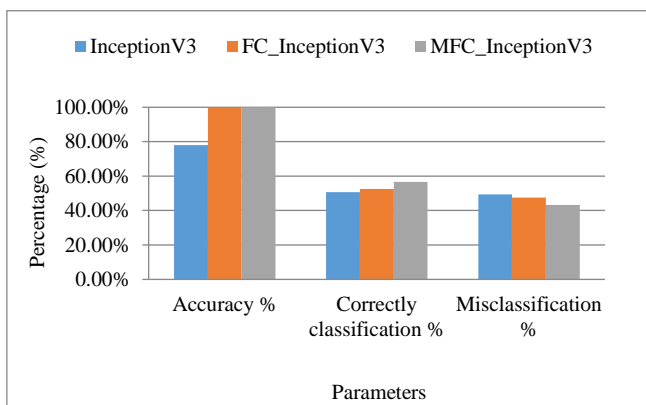
Figure 10. Comparasion between varients of InceptionV3 models

7. DISCUSSION

This section discusses the limitations of the presented work. In this work, an attempt is made to meet the current needs of the users, but still, we acknowledge a few limitations. We created our own dataset of 12000 images considering a top Indian fruit depending on consumption or exported. Each fruit can have subcategories which are not considered in the study. For example, "Pomegranate" is a general category of fruit, but in India four verities of pomegranate are popular, namely,

Arakta, Bhagwa, Ganesh, and Ruby [40]. A standard method called data augmentation is used to overcome the small dataset problem.

Four popular models Inception, Resnet152V2, Xception, and MobileNet are evaluated, but the list of pretrained models is large. The output of the model showed the class of fruit with quality, for example "Apple Bad" or "Apple Good", but it doesn't do the fruit detection of the bunch of fruits. It also doesn't show labels on the output. For this we can use the latest object detection algorithm called YOLO (you only look ones). The major limitation of this work is the solution which is computationally intensive, i.e., we need to create a separate model for each parameter. However, the advantages of this approach promise more accuracy, this reduces the misclassification problem, the models become modular. This approach is good when your dataset is small and has similar feature images.

8. CONCLUSION

Instead of classifying the fruits in general categories, the fruits classified in good or bad fruit, which is the need of stakeholders. A framework to increase the accuracy of the deep neural network model for fruit classification and to address the problem of misclassification is proposed. It is observed that splitting the model will help to achieve more accuracy with better classification. Popular CNN models densenet161, InceptionV3, MobileNetV2, and VGG19 were investigated to highlight the misclassification problem. In this work, we focused on only one model, i.e., InceptionV3 for further investigation. By using transfer learning, two different models are built, namely, FC_InceptionV3 and MFC_InceptionV3. Then the results are compared and found that the MFC_InceptionV3 based on MNet has achieved 99.92% accuracy and alleviated the misclassification problem by 5.98% compared to the original InceptionV3 and 4.17% compared to FC_InceptionV3.

The framework can be extended to develop the smartphone app for the fruit classification system for users.

REFERENCES

[1] Agricultural & Processed Food Products Export Development Authority (Ministry of Commerce & Industry, Govt. of India). http://agriexchange.apeda.gov.in/indexp/product_direct ory.aspx.

[2] Costa, Fiori G., Noferini M., Ziosi V. (2007). Safety and quality of fresh fruit and vegetables: a training manual for trainers, united nations New York and Geneva. ISHS Acta Horticulturae 712: IV International Conference on Managing Quality in Chains - The Integrated View on Fruits and Vegetables Quality. http://dx.doi.org/10.17660/ActaHortic.2006.712.38

[3] Ünal, Z. (2020). Smart farming becomes even smarter with deep learning—a bibliographical analysis. IEEE Access, 8: 105587-105609. http://dx.doi.org/10.1109/ACCESS.2020.3000175

[4] Meshram, V., Patil, K., Hanchate, D. (2020). Applications of machine learning in agriculture domain: A state-of-art survey. International Journal of Advanced Science and Technology, 29(8): 5319-5343.

[5] Singh, U., Chouhan, S., Jain S., Jain S. (2019).

Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. IEEE Access, 7: 43721-43729. http://dx.doi.org/10.1109/ACCESS.2019.2907383

[6] Kour, V., Arora, S. (2019). Fruit Disease Detection Using Rule-Based Classification. In: Tiwari S., Trivedi M., Mishra K., Misra A., Kumar K. (eds) Smart Innovations in Communication and Computational Sciences. Advances in Intelligent Systems and Computing, vol 851. Springer, Singapore. https://doi.org/10.1007/978-981-13-2414-7_28

[7] Xing, S., Lee M., Lee, K. (2019). Citrus pests and diseases recognition model using weakly dense connected convolution network. Sensors, 19(14): 3195. http://dx.doi.org/10.3390/s19143195

[8] Doh, B., Zhang, D., Shen, Y., Hussain, F., Doh R., Ayepah, K. (2019). Automatic citrus fruit disease detection by phenotyping using machine learning. 25th International Conference on Automation & Computing, Lancaster University, Lancaster UK, pp. 1-5. http://dx.doi.org/10.23919/IConAC.2019.8895102

[9] Hua, Y., Zhang, N., Yuan, X., Quan, L., Yang, J., Nagasaka, K., Zhou, X.G. (2019). Recent advances in intelligent automated fruit harvesting robots. The Open Agriculture Journal, 13: 101-106. http://dx.doi.org/10.2174/1874331501913010101

[10] Bresilla, K., Perulli, G.D., Boini, A., Morandi, B., Grappadelli, L.C., Manfrini, L. (2019). Single- shot convolution neural networks for real-time fruit detection within the tree. Frontiers in Plant Science, 10(611). http://dx.doi.org/10.3389/fpls.2019.00611

[11] Hossain, M.S., Al-Hammadi, M., Muhammad, G. (2019). Automatic fruit classification using deep learning for industrial applications. IEEE Transactions on Industrial Informatics, 15(2): 1027-1034. http://dx.doi.org/10.1109/TII.2018.2875149

[12] Kirk, R., Cielniak, G., Mangan, M. (2020). L*a*b*Fruits: A rapid and robust outdoor fruit detection system combining bio-inspired features with one-stage deep learning networks. Sensors, 20(1): 275. http://dx.doi.org/10.3390/s20010275

[13] Altaheri, H., Alsulaiman, M., Muhammad, G. (2019). Date fruit classification for robotic harvesting in a natural environment using deep learning. IEEE Access, 7: 117115-117133. http://dx.doi.org/10.1109/ACCESS.2019.2936536

[14] Bauer, A., Bostrom, A.G., Ball, J., Applegate, C., Cheng, T., Laycock, S., Rojas, S. M., Kirwan, J., Zhou, J. (2019). Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. Horticulture Research, 6(70). http://dx.doi.org/10.1038/s41438-019-0151-5

[15] Zhang, T., Huang, Z., You, W., Lin, J., Tang, X., Huang, H. (2020). An autonomous fruit and vegetable harvester with a low-cost gripper using a 3D sensor. Sensors, 1: 93. http://dx.doi.org/10.3390/s20010093

[16] Onishi, Y., Yoshida, T., Kurita, H., Fukao, H., Takanori, A., Iwai, A. (2019). An automated fruit harvesting robot by using deep learning. Robomech Journal, 6: 13. http://dx.doi.org/10.1186/s40648-019-0141-2

[17] Zhang, X., Qiao, Y., Meng, F., Fan, C., Zhang, M. (2018). Identification of maize leaf diseases using improved deep convolutional neural networks. IEEE Access, 6: 30370-30377.

[18] Liu, X, Chen, S.W., Aditya, S., Sivakumar, N., Dcunha, S., Qu, C., Taylor, C.J., Das, J., Kumar, V. (2018). Robust fruit counting: combining deep learning, tracking, and structure from motion. arXiv:1804.00307v2 [cs.CV].

[19] Kausar, A., Sharif, M., Park, J., Shin, D.R. (2018). Pure-CNN: A framework for fruit images classification. International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, pp. 404-408. http://dx.doi.org/10.1109/CSCI46756.2018.00082

[20] Femling, F., Olsson, A., Alonso-Fernandez, F. (2018). Fruit and vegetable identification using machine learning for retail applications. arXiv:1810.09811v1 [cs.CV], 2018.

[21] Zhang, L., Jia, J., Gui, G., Hao, X., Gao, W., Wang, M. (2018). Deep learning based improved classification system for designing tomato harvesting robot. IEEE Access, 6: 67940-67950. http://dx.doi.org/10.1109/ACCESS.2018.2879324

[22] Cheng, H., Damerow, L., Sun, Y., Blanke, M. (2017). Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. Journal of Imaging, 3(1): 6. http://dx.doi.org/10.3390/jimaging3010006

[23] Puttemans, S., Vanbrabant, Y., Tits, L., Goedem´e, T., (2016). Automated visual fruit detection for harvest estimation and robotic harvesting. International Conference on Image Processing Theory, Tools and Applications (IPTA'16) at: Oulu, Finland. http://dx.doi.org/10.1109/IPTA.2016.7820996

[24] Stein, M., Bargoti, S., Underwood, J. (2016). Image based mango fruit detection, localisation and yield estimation using multiple view geometry. Sensors, 16(11): 1915. http://dx.doi.org/10.3390/s16111915

[25] Yamamoto, K., Guo, W., Yoshioka, Y., Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. Sensors, 14(7): 12191-12206. http://dx.doi.org/10.3390/s140712191

[26] Font, D., Pallejà, T., Tresanchez, M., Runcan, D., Moreno, J., Martínez, D., Teixidó, M., Palacín, J. (2014). A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm. Sensors, 14(7): 11557-11579. http://dx.doi.org/10.3390/s140711557

[27] Bargoti, S., Underwood, J. (2017). Deep fruit detection in orchards. arXiv:1610.03677v2 [cs.RO].

[28] Ucat, R.C., Dela Cruz, J.C. (2019). Postharvest grading classification of cavendish banana using deep learning and tensorflow. 2019 International Symposium on Multimedia and Communication Technology (ISMAC), Quezon City, Philippines, pp. 1-6. http://dx.doi.org/10.1109/ISMAC.2019.8836129

[29] Ireri, D., Belal, E., Okinda, C., Makange, N., Ji, C. (2019). A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. Artificial Intelligence in Agriculture, 2: 28-37. https://doi.org/10.1016/j.aiia.2019.06.001

[30] Jr Piedada, E., Laradab, J.I, Pojasc, G.J., Ferrerc, L.V.V. (2018). Postharvest classification of banana (Musa acuminata) using tier-based machine learning. Postharvest Biology and Technology, 145: 93-100. https://doi.org/10.1016/j.postharvbio.2018.06.004

[31] Sofu, M.M., Er, O., Kayacan, M.C., Cetisli, B. (2016). Design of an automatic apple sorting system using machine vision. Computers and Electronics in Agriculture, 127: 395-405. http://dx.doi.org/10.1016/j.compag.2016.06.030

[32] Ohali, Y.A. (2010). Computer vision based date fruit grading system: Design and implementation. Journal of King Saud University Computer and Information Sciences, 23(1): 29-36. http://dx.doi.org/10.1016/j.jksuci.2010.03.003

[33] Meshram, V., Thanomliang, K., Ruangkan, S., Chumchu, P., Patil, K. (2020). FruitsGB: Top Indian fruits with quality. IEEE Dataport. http://dx.doi.org/10.21227/gzkn-f379

[34] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C. (2018). A Survey on Deep Transfer Learning. In: Kůrková V., Manolopoulos Y., Hammer B., Iliadis L., Maglogiannis I. (eds) Artificial Neural Networks and Machine Learning – ICANN 2018. ICANN 2018. Lecture Notes in Computer Science, vol 11141. Springer, Cham. https://doi.org/10.1007/978-3-030-01424-7_27

[35] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2015). Rethinking the inception architecture for computer vision. arXiv:1512.00567v3 [cs.CV].

[36] Smith, D.R. (1985). The design of divide and conquer algorithms. Science of Computer Programming, 5: 37-58. http://dx.doi.org/10.1016/0167-6423(85)90003-6

[37] Lin, M., Chen, Q., Yan, S. (2014). Network in network. arXiv:1312.4400v3 [cs.NE].

[38] Nair, V., Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, pp. 807-814.

[39] Salakhutdinov, R., Hinton, G. (2009). Replicated Softmax: An undirected topic model. International Conference on Neural Information Processing, NIPS'09. 1607-1614.

[40] Agricultural & Processed Food Products Export Development Authority (Ministry of Commerce & Industry, Govt. of India). http://apeda.in/agriexchange/Market%20Profile/MOA/Product/Pomegranate.pdf.