Recognition and Classification of Fruits using Deep Learning Techniques

MSc. Research Project
MSc. Data Analytics

Shubham Kathepuri
Student ID: x18127398

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Shubham Balasaheb Kathepuri		
Student ID:	x18127398		
Programme:	MSc. Data Analytics	**Year:**	2019 - 2020
Module:	MSc. Research Project		
Supervisor:	Dr. Catherine Mulwa		
Submission Due Date:	28th September 2020		
Project Title:	Recognition and Classification of Fruits using Deep Learning Techniques		
Word Count:	9151	**Page Count:**	25

I hereby certify that the information contained in this (my submission) is information pertaining to the research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Shubham Balasaheb Kathepuri

Date: 27th September 2020

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on the computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Recognition and Classification of Fruits using Deep Learning Techniques

Shubham Kathepuri
x18127398

Abstract

Recently, there have been great advancements in the field of deep learning making it a popular choice for image processing applications. Recognition and classification of fruits using deep learning is one of the exciting applications of computer vision for commercial as well as agricultural applications. Nevertheless, the researchers still face challenges while the classification of fruits due to similarity of color, shape, and size. This project attempts to address some of the challenges faced by the previous researchers by developing a methodology for the recognition and classification of fruits. The deep learning models used for the project were CNN, VGG16, and ResNet50. These models were trained and tested using two sets of data one was pre-processed and the other was augmented. CNN performed well on the first set of data with high accuracy of 0.9691 and less computational time of 11.8 minutes whereas ResNet50 was able to achieve high accuracy of 0.9522 on the second set of data with a computational time of 24.87 minutes. Hence, the deep learning models were able to recognize and classify 131 categories of fruits accurately. However, it was observed that image augmentation did not improve the performance of the models.

1 Introduction

Computer vision is a field of data science that enables machines to gain a profound understanding of the human world by analyzing videos and pictures. With the help of deep learning algorithms computers are now capable of identifying and classifying the objects observed in the videos or pictures. The field of computer vision was first introduced in the 1950s when a neural network in its early stage of development was able to classify the object based on the edge detection[1]. With the introduction of the internet, a plethora of data was available which accelerated the development of this field. Just over a decade, the accuracy of these systems has improved drastically due to the vast amount of data being generated each day. The automation is continually on the rise with the increasing demand for goods across all industries which in turn has driven the use of computer vision and its applications. The impact of this technology can be widely seen in each field which is dependent on machines for analyzing the videos, images, etc. The objective of computer vision is to overcome the limitation of the traditional system by addressing the complexities and reach better efficiency. There are several applications where computer vision is used, however, this research discusses various methods for fruits classification using deep learning techniques.

[1] https://techsee.me/blog/computer-vision-applications/

It is a known fact that fruits are essential to lead a healthy lifestyle. Fruits are packed with vital nutrients like folic acids, vitamin C, potassium, and dietary fibers which are crucial for our health. It is proven that consuming fruits can lower the risk of high blood pressure, type 2 diabetes, some chronic disease, etc[2]. It is recommended to include at least one-fourth plate of fruits in a healthy meal[3]. There are numerous applications where fruit recognition and classification using deep learning and artificial intelligence can be employed. One of the applications is a fruit recognition system which on detection gives a detailed description of the respective fruit, such systems can be utilized for educational and shopping purposes (Khan and Debnath, 2019). One of the major applications of fruit detection and recognition is agricultural robots which are used in precision agriculture. The agricultural tasks are generally repetitive in nature which consists of harvesting, seeding, picking, weeding, sorting, feeding, and spraying of pesticides (Bresilla et al., 2019). Using robots for these tasks help farmer focus more on forming strategies to enhance the yields. Similarly, fruit recognition and classification are used in smart refrigerators which detects how many fruits are left, are the available fruits fresh enough to eat, and which fruits need to be replenished (Desai, 2019; Buzzelli, Belotti, and Schettini, 2018). One of the use cases could be a mobile application that can help a user identify whether the fruit is fresh or not. For these applications to work smoothly and minimising the damage caused to fruits in the process, a fast and accurate fruit recognition, and classification system is necessary.

1.1 Research Question

The topic of the fruits classification system is gaining high research attention in consideration of its increasing importance. Numerous works focus completely on fruit recognition and classification. Most of the researchers used techniques like SVM, statistical methods, YOLO, MobilNetv1, ResNet, VGG-16, Mask-RCNN, and Faster R-CNN, with GPU based computational power to reduce the computational time and obtain accurate results. However, the proposed works by the researchers are still facing some difficulties. Some of the common challenges faced by the researchers are differentiating between two similar-looking fruits like tomatoes and apples, differentiating between the same category of fruits but with different classes such as red apple from green apple, detecting the fruit among leaves, no to limited availability of real-life dataset consisting of fruits in nature, etc. (Khan and Debnath, 2019). To address these challenges, an ideal machine learning method should be competent enough to identify and differentiate the fruit from its background, accurate classification of the fruit given any shape and size, recognition, and classification with high accuracy and less computing power. Hence based on these challenges the following research question was formed.

RQ**:** *"How well can the deep learning techniques (CNN, VGG-16, ResNet50) recognize and classify fruits in order to reduce the human error, improve the process, and enhance profits for the end-user?"* Depending on the application of the project the process can be improved in terms of time and reliability.

---

2 https://www.betterhealth.vic.gov.au/health/healthyliving/fruit-and-vegetables
3 https://www.choosemyplate.gov/ten-tips-build-healthy-meal

Sub-RQ*: *"Can image augmentation improve the performance of the deep learning techniques (CNN, VGG-16, ResNet50) employed for fruits recognition and classification systems?"

To answer the above-mentioned questions following objectives were implemented and achieved.

1.2 Objectives and Contribution

To answer the research question and achieve the desired results following objectives in Table 1 were defined and implemented. These objectives include a comprehensive review of the literature from 2010 to date, contributing to existing literature, and implementing models using deep learning techniques for recognition and classification of fruits.

Table 1: Objectives

<table>
<tr><th>Obj. no.</th><th>Description</th><th>Evaluation Metrics</th></tr>
<tr><td>1</td><td>Develop a fruit recognition and classification methodology and contribute to the existing literature.</td><td></td></tr>
<tr><td>2</td><td>Acquire and pre-process the images of fruits.</td><td></td></tr>
<tr><td>3</td><td>Implement data augmentation techniques using ImageDataGenerator.</td><td></td></tr>
<tr><td>4</td><td>Implement and evaluate fruit classification models and data augmentation of the images</td><td></td></tr>
<tr><td>4 (i)</td><td>Implementation, evaluation, and results of CNN with and without data augmentation.</td><td rowspan="5">Accuracy & Time</td></tr>
<tr><td>4 (ii)</td><td>Implementation, evaluation, and results of VGG-16 with and without data augmentation.</td></tr>
<tr><td>4 (iii)</td><td>Implementation, evaluation, and results of ResNet50 with and without data augmentation.</td></tr>
<tr><td>5</td><td>Comparison of all implemented models.</td></tr>
<tr><td>6</td><td>Comparison of implemented models with state-of-the-art models.</td></tr>
</table>

Contributions: The major contribution of this project is that the developed models can recognize and classify 131 varieties of fruits with high accuracy and reasonable computational time which will reduce the associated labour cost, time, and eliminate the possibility of human error for stakeholders which include customers, retail shops, supermarkets, and farmers. This will ensure smooth and hassle-free execution of tasks across a wide array of applications. The minor contribution resulting from this project was reviewed literature and identified gaps. Besides, it was observed that the models that were implemented with image augmentation did not improve the performance of the models compared to the models implemented without image augmentation.

To perform this study, a critical review of the existing literature was carried out which is presented in chapter 2, it is followed by the proposed methodology for the recognition and classification of fruits in chapter 3. The development of different models and methods to recognize and classify fruits is carried out in chapter 4 which is followed by results and discussion in chapter 5. And finally, the conclusion of the study addressing the objectives, research question, and future work in chapter 6.

2 Related Work

Fruit recognition and classification are considered a complex task and are still faced with certain challenges as stated above. To develop a perfect fruit recognition and classification technique these challenges need to be overcome. This section presents the work of researchers who have attempted to solve some of these problems using various methods such as statistical, machine learning, and deep learning methods. The following studies present the work that has been carried out over the last decade starting from 2010 to date. Further, each study is critically evaluated based on what problem the researcher is addressing, what kind of data is used for the study, the techniques used, future work, and its practical application. The following section is divided into four sections based on which techniques are used: 1) Machine Learning Techniques, 2) Statistical Techniques, 3) Deep Learning Techniques, 4) Critique, Identified Gaps, and Conclusion.

2.1 Critical Review of Machine Learning Techniques Employed

According to Shukla and Desai (2016), fruit recognition is a difficult task due to their shape, color, texture, and size. To add to this, different imaging conditions in which pictures of fruits are captured act as another hurdle. This study is an attempt to automate and address these issues using machine learning. The dataset consists of 9 categories of fruits with 155 images in total. Otsu's method is used for converting the color image to greyscale image to binary image. Combination features like shape, color, and texture, are used and passed to Support Vector Machine (SVM) and KNN which act as multiclass classifiers. Results show that K-Nearest Neighbour (KNN) performs better than SVM and gives the best accuracy for k=2 value. Best results are obtained for a combination of color, shape, and texture rather than for any two features. It is said that in the future, images of fruits in the natural environment should be used, and rather than a single instance of fruit, it should contain multiple fruits of the same or different kind.

In another study, automatic fruit classification using machine learning techniques is proposed. The proposed methodology consists of stages like pre-processing, feature extraction, and classification. In the pre-processing stage, the images are resized to 90 x 90 px to reduce the color index of the images. For feature extraction, two different methods are used namely shape and color algorithm and Scale Invariant Feature Transform (SIFT) algorithm. In the classification stage, the classification of fruits is carried out using SVM, Random Forest (RF), and K-NN. The dataset used to train and evaluate the results contained 178 images of apples, strawberries, and oranges. The results of these models are compared and it concluded that RF acquired higher accuracy compared to the other two algorithms (Zawbaa et al., 2014).

Automation in the field of agriculture is continually on the rise to improve productivity and quality of products. Classification of fruits plays a crucial role in sorting various kinds of fruits. SVM and Genetic Algorithm (GA) are proposed by Mahajan (2016) to carry out this task. The dataset used for this task consist of 178 images of three types of fruits namely apples, grapes, and bananas. These images are first pre-processed then features like color, shape, and texture are extracted. The extracted features are optimized using GA to enhance the training speed. GA also helps in reducing the number of features by selecting the best features only. The classification of the fruits is performed using SVM which was able to achieve an accuracy of 0.9677.

Furthermore, for assisting automated harvesting, Jana, Basak, and Parekh (2017) propose an efficient and accurate classification method for a variety of fruits. The images in the dataset are pre-processed to extract the fruit from its background. For feature selection, texture and color features are extracted with the help of Gray-level Co-occurrence Matrix (GLCM) and segmented image. The SVM classification model is created by combining these two features using a single feature descriptor. The proposed model was able to achieve an overall accuracy of 0.8333.

To reduce the cost of the agricultural task of picking and improve the automation adaptability, Peng et al. (2018) propose a machine learning method. The dataset consisted of images of 6 categories of fruits which were pre-processed using Gaussian filter and image brightness histogram equalization. To identify edges of the fruits Canny Edge Detection was used while for image segmentation OTSU algorithm was used. For feature extraction shape invariant moment and HIS color model were used. For classification SVM was used which achieved the highest accuracy of 0.9750 for citrus and the lowest accuracy of 0.8000 for bananas.

For assessing the quality and reliable differentiation of fruits this study presents a method that ranks four kinds of fruits based on their quality. In this method first, the color was extracted and then the fruit was split from its background using the split and merge method. Features such as geometrical color, textural, and statistical are extracted from the images. A comparison of machine learning and deep learning techniques was carried out. It was observed that SVM most accurately detects fruits with an accuracy of 0.9848. SVM also differentiated between Rank1, Rank2, and defected fruits with an accuracy of 0.9527 which was found to be best amongst all (Bhargava and Bansal, 2019).

2.2 Critical Review of Statistical Techniques Employed

In this study, a fusion of two features namely color and texture is used. To recognize a given fruit this approach uses a minimum distance classifier which is based upon statistical and co-occurrence features. A dataset procured from a supermarket containing 2633 images of 15 categories fruits was used of which 0.5000 was used for training and rest was used for testing. The recognition rate using only color features was 0.4549 while using only texture features was 0.7085. The proposed approach worked better when both color and texture were used in combination which gave recognition accuracy of 0.8600. To enhance the flexibility and functionality of this system in future features like size and shape of fruits can also be used alongside color and texture. Also, to improve the recognition rate more images can be included in the dataset (Arivazhagan et al., 2010).

Further, an image processing approach is discussed by Desai (2019) for the classification of fruits of four kinds. The approach was proposed for an application of refrigerator inventory management which keeps a count of fruits of each kind. First, the images are converted to HSV for color detection. Second, for boundary detection Fourier descriptors are used and for determining the texture of image statistical moments are used which extends the algorithm used for simple object detection to the application of refrigerator inventory management.

2.3 Critical Review of Deep Learning Techniques Employed

An accurate, reliable, and low-cost image-based system for strawberry detection which uses convolution neural networks was proposed by Lamb and Chuah (2018). Single Shot Multibox Detector (SSD) neural network framework was used for implementation as the fruit detector. A sparse, three-layer convolutional layer was used as a classifier which was modified in numerous ways to enhance precision and speed. To boost the performance of the network, first, the input image is compressed to 360 x 640 px and a color mask is applied to separate the regions of interest. 160 lowly weighted filters are removed by compressing the entire network and then it is retrained. The developed model could augment the individual frames in a live video with depth information to locate each recognizable strawberry in 3D. This system can be employed for mass harvesting in turn reducing human intervention, enhancing efficiency, and decreasing damage to the fruit.

In a similar study, Hossain, Al-Hammadi, and Muhammad (2019) proposed an efficient fruit classification framework using deep learning for applications like identifying fruits in a supermarket and help people to identify whether the fruit meets their dietary requirements. The proposed framework was based on two different deep learning models. First, the model used was a light framework consisting of a 6 layered convolution neural network. Whereas the second model used was a pretrained Visual Geometry Group 16 (VGG-16) model. These models are trained using two datasets of color fruit images spread across 15 categories of fruits. One of the datasets contains clear fruits images (2633 images) while images in which it is hard to classify fruits were included in the second dataset (5946 images). The accuracy of classification achieved by the first model on the first and second dataset was 0.9949 and 0.8543. The classification accuracy achieved by the VGG-16 model on the first and second datasets was 0.9975 and 0.9675. For future work more variety of fruits and veggies can be added in the dataset, different model parameters can also be tuned.

Furthermore, in Indonesia fruits are considered a commodity and a high potential crop. Fruits are produced on a large scale, but the harvesting process slows the production in turn decreasing the quality of fruits resulting in lower selling price. To address this issue Basri, Syarif, and Sukaridhoto (2018) proposed a Faster Regions with Convolutional Neural Networks (Faster R-CNN) deep learning method which uses MobileNet as the base model to classify different fruits. The dataset used to train consists of 700 images of mangoes and pitaya each. The accuracy achieved was 0.9900 making it appropriate to be employed in sorting machine which can sort the fruits in real-time and maintain their quality.

In addition to this, according to Mai et al. (2020), Faster R-CNN shows a lack of detection advantage in the case of small fruits. The reason behind this is for localization of the proposed candidates only single level features are used. To overcome this challenge, the proposed method uses multiple classifier fusion strategies into the Faster R-CNN network for the detection of small fruit. To achieve this, a feature from three different levels are used to learn three classifiers for “objectness” classification in the stage of proposal localization. Further to generate the final “objectness” for proposal candidates the probabilities of the classifiers are combined. For training, a novel loss function is also presented. The training set consists of 384 images of almonds along with 1574 images with other bigger fruits in them. It is observed that the proposed model can detect small fruits with an F1 score of 0.8221.

Moreover, to speed up and reduce human-computer interactions, system design to identify fruits and vegetables in retail markets is proposed by Femling, Olsson, and Alonso-Fernandez (2018). For this purpose, 10 categories of fruits and vegetables are considered namely apple, banana, avocado, bell pepper, orange, clementine, pear, kiwi, potato, and tomato. 400 images of each category are used for this study which is obtained from ImageNet. Two different Convolutional Neural Network (CNN) architectures (Inception and MobileNet) are used for this purpose. The MobileNet outperforms Inception with top accuracy of 0.9700 with accurate predictions and fast identification.

To replace visual inspection of fruits which is an inconsistent and expensive process, Choi et al. (2018) proposed a smart fruit quality classification machine. In this system, along with other components a picture of the fruit is captured which is then pre-processed using segmentation techniques for computation of fruits shape features. A dataset consisting of 1800 images of pears is used for training an Artificial Neural Network (ANN) feed-forward model. Activation function ReLu (Rectified Linear Unit) and Adams optimizer were selected for the model. The proposed model achieved a classification accuracy of 0.9740.

Classification of fruits is considered as one of the challenging operations because of the similarity in shape, color, texture, and its vast variety. To perform the classification of fruits accurately, an approach based on Pure Convolution Neural Network (P-CNN) is proposed along with a minimum number of parameters. The proposed model consists of 7 convolutional layers. It was observed that the Global Average Pooling (GAP) overcomes the overfitting problem and provides better performance than the Fully Connected (FC) layer. The model was trained using the Fruit-360 dataset which includes 81 classes of fruits and 55244 images. P-CNN was accurately able to classify most of the fruits with 0.9888 accuracy (Kausar et al., 2018).

As an application of smart kitchen/refrigerator, Buzzelli, Belotti, and Schettini (2018) proposed a vegetable and fruit recognition method based up on CNN. Two datasets were used for this study namely Fruits 360 and VegFru. Different neural network architectures such as ResNet-34, ResNet-50, and NASNet are fine-tuned with the help of Stochastic Gradient Descent (SGD) to carry out the task of recognition. Two sets of experiments were carried out. In the first experiment, the above-mentioned models are trained and evaluated against state-of-the-art neural architectures. Results from the first of the two experiments showed that NASNet outperformed other models. In the second set of the experiment, a solution was designed and tested to exploit the hierarchical nature of such classes to enhance the performance of the system. The second experiment results showed that the performance of ResNet-50 improved whereas no significant change was observed in the case of NASNet.

Similarly, to address the challenges for accurate fruit recognition in a smart refrigerator Zhang et al. (2018) proposed a multi-model fusion coupled with a multi-source data fusion method. For multi-model fusion three single shot multibox detection models (ResNet, VGG-

16, and VGG-19) are used, as different models extract different features. The output from these models is passed over to a back-propagating neural network. It is coupled with a multi-source data fusion method that captures weight information of fruits. Two datasets were considered for this study, one was acquired from the internet (25000 images) and the other was self-developed (5000 images). It was observed that the proposed model performed better in comparison to ResNet, VGG-16, VGG-19, and multi-model fusion with a recognition accuracy of 0.9700.

Zhang et al. (2014) proposed a classification method that uses Feed Forward Network (FNN) and Fitness-Scaled Chaotic Artificial Bee Colony (FSCABC). The dataset used included 1653 color images of 18 categories of fruits. The images were pre-processed using a split and merge algorithm which is an image segmentation technique used for removing background noise. For feature extraction, a composite feature space was used with which 7 texture features, 64 color features, and 8 shape features were extracted. To further reduce the number of features, PCA (Principal Component Analysis) was implemented which covered 0.9500 variance of the original image. Cross-validation is used to improve the generation capacity of the classifier and through trial and error value of k for k-fold is determined to 5. Classification accuracy of 0.8910 was achieved using FSCABC-FNN which was best among other models.

CNN based on the YOLO model, a deep learning technique is proposed by Bresilla et al. (2019) for accurate and fast detection of fruits. The need to extract hard-coded features like shape, color, size, etc. is eliminated when using a deep learning technique. The overall size of the model was reduced to 11 layers with a grid size of 26 x 26 from the original size of 13 layers with a grid size of 13 x 13. The model was trained using 5,000 images of apples. The F_1 score obtained for the proposed model was 0.7900 which improved significantly (0.90) when the dataset size was increased to 20,000.

Moreover, to address the manual sorting of fruits which is quite a time-consuming task, an approach to automate the process has been discussed by Saranya et al. (2020). Five types of fruits are considered in this study. The images are pre-processed to standardize and normalize the images using different filters and eliminate the noise. To convert images to grayscale OTSU threshold is used and to smoothen the edges morphological operations are performed. Features such as height, size, color, and width are considered. A comparison of machine learning and deep learning models is carried out. It is observed that a simple CNN outperforms the rest with an accuracy of 0.9649.

2.4 Identified Gaps and Conclusion

After critically reviewing the literature, it was observed that all the methodologies, techniques, and methods can obtain good results but are heavy on the computational side as the fruit images are colored high dimensional images which increases the computational time and storage memory (Zhang et al., 2014). Further, it was also noted that for one of the studies, data augmentation i.e. random zoom, flipping, horizontal and vertical shift, and rotation improved the accuracy of the model significantly (Hossain, Al-Hammadi and Muhammad, 2019). Table 2 summarises some of the best results achieved by the researchers in the reviewed techniques.

Table 2 Comparison of Reviewed Techniques

Author	Fruit Type / Dataset Size	Technique	Accuracy
Hossain, Al-Hammadi and Muhammad (2019)	15 categories Dataset 1 = 2633 images Dataset 2 = 5946 images	VGG-16	Dataset 1 = 0.9975 Dataset 2 = 0.9675.
Basri, Syarif and Sukaridhoto (2018)	Mango and Pitaya Dataset = 1400 images	Faster R-CNN	0.9900
Femling, Olsson and Alonso-Fernandez (2018)	10 categories Dataset = 4000 images	MobileNet	0.9700
Choi et al. (2018)	Pear Dataset = 1800 images	ANN	0.9740
Kausar et al. (2018)	81 categories Dataset = 55244 images	P-CNN	0.9888

One of the drawbacks which were common in most of the studies was the size of the dataset used for the study and the variance of the fruits under consideration. Also, it was seen that the reviewed techniques and methodologies for recognition and classification of fruits using deep learning showed good results. Hence, to investigate how well the deep learning techniques can perform on a larger dataset with a very wide variety of fruits and can image augmentation can enhance the obtained results, based on the reviewed literature this project will employ deep learning models (CNN, VGG-16, ResNet) along with image augmentation on a dataset with 131 categories of fruits and 90,483 images. The recognition and classification of fruits using deep learning techniques being computationally heavy were evaluated using execution time and accuracy of predictions. This chapter achieves objective 1 mentioned in chapter 1 Table 1.

3 Research Methodology

This section discusses the modified methodology based on CRISP-DM and architecture design used for the development of this project.

3.1 Recognition and Classification of Fruits Methodology

The methodology used for this project is based upon CRoss-Industry Standard Process for Data Mining (CRISP-DM) which offers guidelines and framework for data mining projects **(**Azevedo and Santos, 2008**)**. CRISP-DM is deemed fit for this project as it includes the business aspect which is necessary for the complete understanding of the project. The following Figure 1 shows the graphical representation of the modified CRISP-DM approach applied to the recognition and classification of fruits which is followed by a detailed explanation of the process.

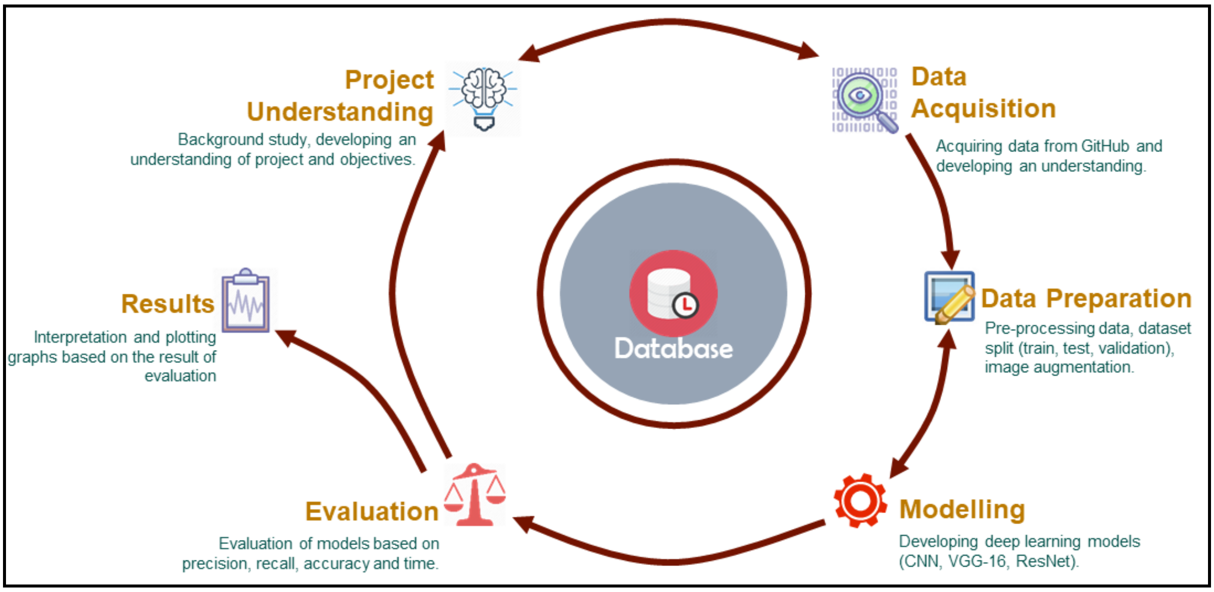

Figure 1 Recognition and Classification of Fruits Methodology

Project Understanding: In this phase, an understanding of the requirements and objectives of the project was developed from a business standpoint. The knowledge gained from this was used to develop a problem statement and to achieve the project objectives an initial plan was designed.

Data Acquisition: In this step of the methodology, data is needed to be acquired from the project resources. If required, the acquisition is followed by the loading of data for understanding purposes. The data is examined for any patterns, quality problems, and verification.

Data Preparation: Following the acquisition of the data, pre-processing was performed based on the knowledge obtained from the previous step. Then the dataset was split into train, test, and validation set. Further, image augmentation was performed on all the images.

Modeling: This step involved the selection and development of different models i.e. CNN, VGG-16, and ResNet. These models were trained and validated using train and validation sets. After training these models to assess their performance they were employed on test sets

Evaluation: This step dealt with evaluating the performance of developed models based on parameters such as accuracy and execution time. The models were carefully reviewed and evaluated to make sure if they were able to achieve the set business objectives.

Results: This is the final step in the methodology which focused on the interpretation of the gained knowledge and results. Further, the obtained knowledge and results were organized and presented with help of graphs.

3.2 Design Specifications

Figure 2 is an architectural diagram that was employed and deemed fit for carrying out this project. It illustrates a two-tier architecture diagram consisting of two layers namely the client layer and business logic layer. It also contains the technologies which are used in each step of the project.

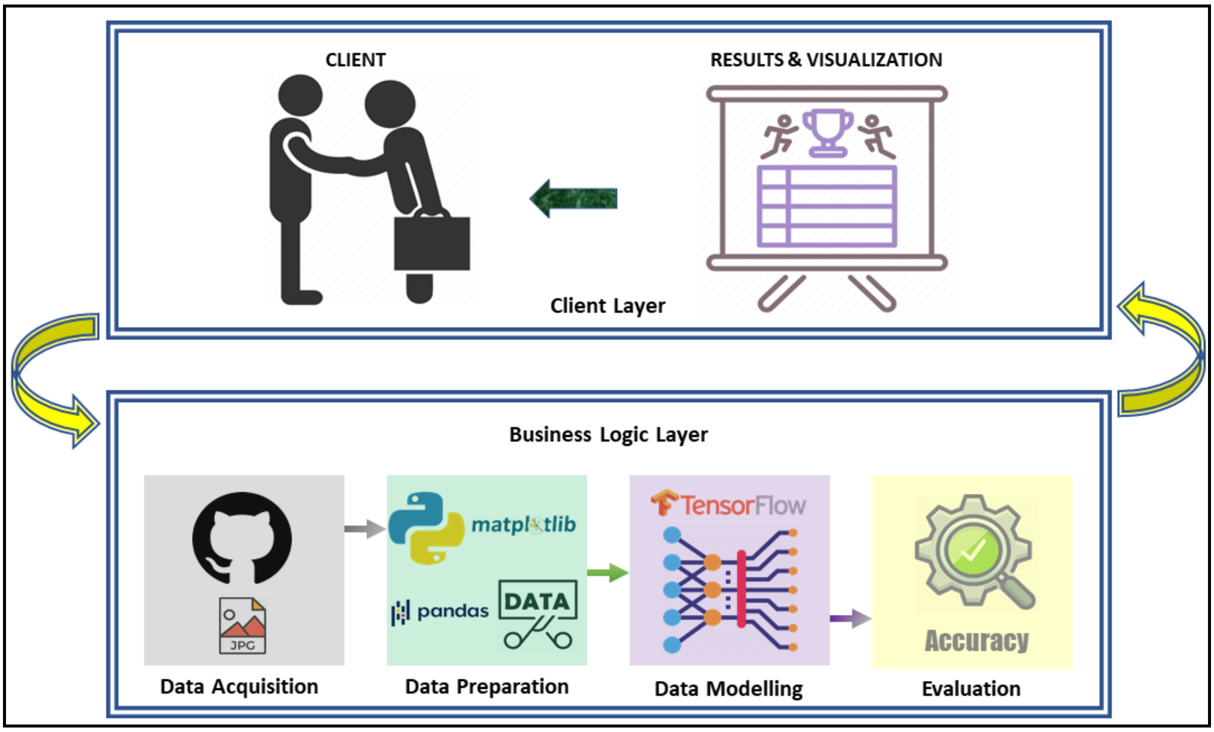

Figure 2 Design Specification

The above-mentioned methodology and design specifications were used for the implementation of this project.

4 Implementation, Evaluation, and Results of Recognition and Classification of Fruits

4.1 Introduction

This chapter is an extension of the methodology chapter. The following sections present in detail about specifications and technologies used for data pre-processing, image augmentation, and modeling which helped in achieving objectives 2, 3, and 4 from chapter 1 Table 1.

Implementation: In this section three models CNN, VGG16, and ResNet50 were developed from scratch. The implementation section follows a pattern of model theory, implemented model, base model results, and model with data augmentation results. To perform a comparative study and considering the computational limitations of the system, a batch size of 16 was kept constant for all models. Also, optimum results were obtained over 5 number of epochs. The validation split was 0.3 throughout the implementation.

Evaluations: After critically reviewing the literature, it was observed that in most of the studies, the evaluation metrics used for measuring the performance of the model was accuracy. Hence the evaluation metrics used for measuring the performance of the implemented models was accuracy. Accuracy can be defined as a ratio of correctly classified data points (True Positive and True Negative) to the total number of data points. Following is the equation of accuracy.

$$\textbf{Accuracy} = \textbf{(TP + TN) / (TP + TN + FP + FN)} \qquad (1)$$

In addition, to understand whether a model is computationally heavy or not, the execution time of each model was recorded using the time library.

4.2 Data Acquisition

The dataset was acquired from GitHub[4]. The dataset was artificially created by mounting the fruits on a rotating shaft and a short video of 20 seconds was recorded using a Logitech C290 webcam. The fruit images were then created from the video using a dedicated algorithm due to inconsistency in the lighting conditions. Further, the dataset contains 131 different types of fruit and 90,483 images with one fruit per image. The pixel size of the images is 100 x 100 with fruit in the forefront and white background. To understand the quality and get some insights regarding the pre-processing of the images, random images were chosen from each category and plotted with help of the matplotlib library.

4.3 Data Pre-processing and Data Augmentation

In this step, after acquiring the dataset, it is extracted with the help of a user-defined function "extract_dataset" using the zipfile library. To understand the distribution of the number of images in each category, 131 categories along with its number of images was printed. To verify images in each of the 131 categories are correct and to check their quality, random images from each of the categories are plotted using the matplotlib library (Figure 3). There was no need to divide the dataset into train and test sets as the acquired dataset already had two separate folders for train and test sets. Further, the dataset details such as the total size of the dataset, train and test dataset size, and the total number of categories were printed and the dataset was loaded with the help "load_dataset" function. The training dataset consisted of 67,692 images while the test set consisted of 22,688 images. After loading the dataset, the images were converted into a pixel array using the PIL library. The images were also rescaled while the conversion by dividing each pixel by 255. The final result was two arrays (train and test) with a 4D shape of (1, 100, 100, 3) which represents axis, image size, and channel. Further, as the images in the dataset were already clean, no additional pre-processing was performed in addition to the aforementioned processing.

Further, another set of train dataset was loaded with help of ImageDataGenerator, which augments the images in real-time without creating a physical dataset. The image augmentation technique is employed to increase the diversity of the data without having to collect additional data. By generalizing better, it also enhances the performance of the models and thus reduces the problem of overfitting. Rescaling, rotation, width shift, height shift, zoom, and horizontal flip operations were performed on the images for augmentation. The training dataset was split into train and validation set in the ratio of 70:30. Another reason why it is important to perform image augmentation is the way the dataset was formed, the dataset lacks the variance between the fruits as not all fruits are identical in all ways. The

4 https://github.com/Horea94/Fruit-Images-Dataset

result of image augmentation is shown in Figure 4. Hence performing augmentation will also improve the variance of the data. Objective 2 and 3 from chapter 1 Table 1 has been achieved.

Figure 3 Random Images from Few Class

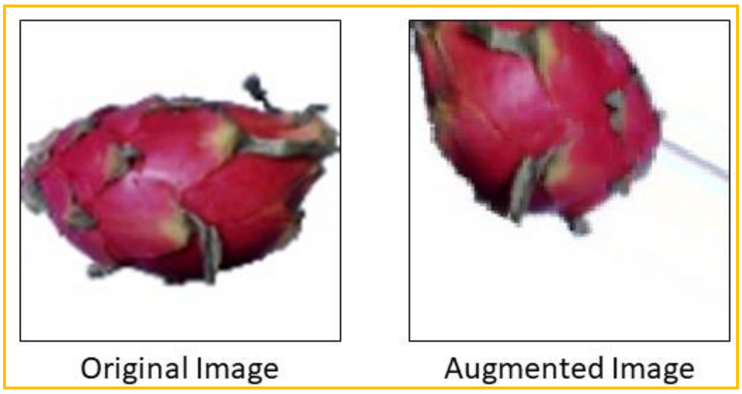

Figure 4 Before and After Augmentation Image

4.4 Implementation of Convolution Neural Network Model

CNN is a deep learning algorithm, which has a feed-forward architecture, can capture an input image, being able to distinguish among objects, and assign importance i.e. biases and learnable weights. CNN is developed based on the biological concept of neurons in the human brain through which they can learn highly abstract features and efficiently classify in between objects and have a remarkable capability to generalize. One of the main reasons to choose CNN is its ability to share weights which in turn reduces the parameters essential for training. This stimulates smooth training and overcomes the issue of overfitting. The general CNN architecture consists of multiple blocks of convolution layers, activation function, pooling layers, and a fully connected layer which assists feature extraction and classification (Albawi, Mohammed, and Al-Zawi, 2017)

The CNN model used in this study was developed from scratch. After several trial and error attempts, the best results were obtained for a model consisting of four convolution layers with a kernel size equal to 2, padding as 'same' and activation function as 'ReLu'. The input size of the image is 100 x 100 x 3. The number of filters used in each convolution layer is doubled in the next layer i.e. the first consist of 16 filters while the last consist of 128 filters. Each convolution layer is followed by a maxpooling layer with a pool size of 2 it reduces the dimensionality while retaining maximum features. After stacking these layers, a dropout layer is added with a probability value of 0.3. The output from the dropout layer is flattened to convert the 3D matrix of features to a vector and passed on to a fully connected dense layer with 150 units and ‘relu’ as an activation function. This is followed by a final dropout layer with a probability of 0.4 and a dense layer with units equal to the number of fruit categories in the dataset i.e. 131 and ‘softmax’ as activation function which classified the images.

The developed CNN model was first directly employed (trained and tested) on the dataset and was evaluated using Accuracy and Execution Time. Secondly, the model was trained and tested on the dataset after applying image augmentation. The effect of image augmentation was observed using the evaluation parameters. Hence, objective 4 (i) from chapter 1 Table 1 is achieved.

4.4.1 Base CNN

The developed CNN was trained and tested using the train and test tensors created during the pre-processing stage. The ‘rsmprop’ optimizer was used along with ‘categorical_crossentropy’ as the loss function and ‘accuracy’ were used as metrics. The training accuracy achieved was 0.9991 while the test accuracy of the model was 0.9646. The time taken by the model to complete training was 11.8 minutes. To verify the test accuracy random 16 images were selected and passed to the model for predictions (Figure 5). The model predicted all 16 correctly. Figure 6 shows the variance of model accuracy and model loss for training and validation sets per epoch. As the gap between the train and validation accuracies goes on decreasing whereas the train and validation loss have good learning rates this implies the model is good i.e. its neither overfitting nor underfitting.

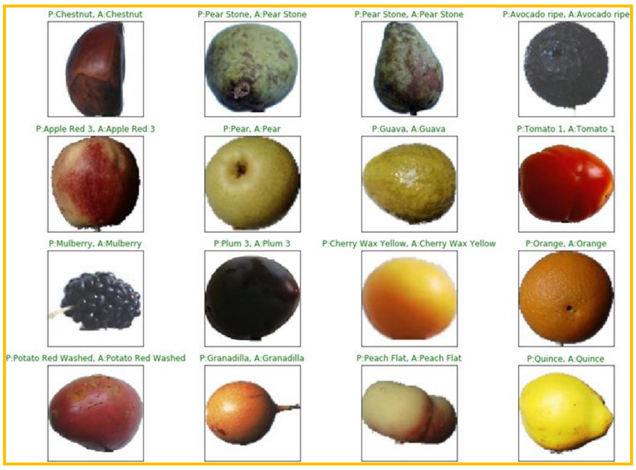

Figure 5 Random Predictions CNN

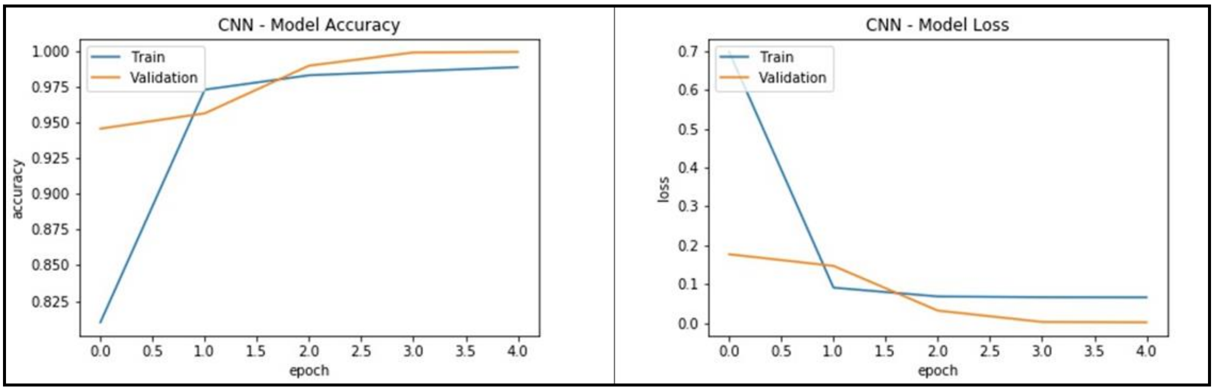

Figure 6 Accuracy and Loss Graphs of CNN

4.4.2 CNN with Image Augmentation

Similarly, the model with the same specifications as the base CNN model was trained using the augmented images. The training and test accuracies obtained were 0.9548 and 0.9383, respectively. The time taken during the training was 20.05 minutes. The model predicted 14/16 correct prediction when verified (Figure 7). It can be observed that in Figure 8, the gap between train and validation accuracies is negligible i.e. no overfitting. Also, the validation loss is seen to have a good learning rate till the 3rd epoch and then there is a sudden spike for the 4th epoch and drop for 5th epoch whereas train loss is almost constant implying model underfitting.

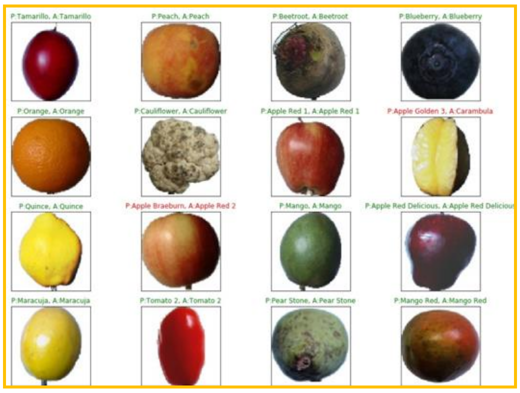

Figure 7 Random Prediction Augmented CNN

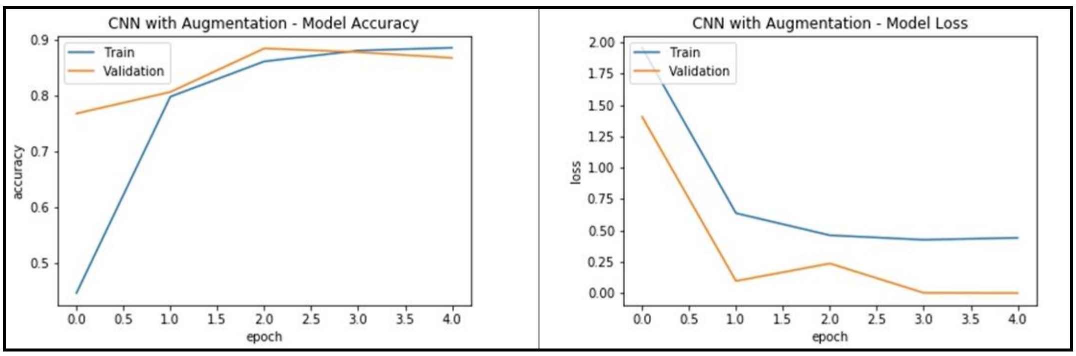

Figure 8 Accuracy and Loss Graphs of Augmented CNN

4.5 Implementation of Visual Geometry Group 16 Model

VGG16 is one of the well-known transfer learning models in deep learning. As the name suggests it consists of 16 layers – 13 convolution layers, 2 fully connected layers, and one softmax classifier. These 16 layers are divided into 5 blocks. The first 2 blocks contain 2 convolution layers followed by 1 maxpooling layer. The next 3 blocks contain 3 convolution layers followed by 1 maxpooling layer. These 5 blocks are followed by 2 fully connected layers and softmax classifiers. The image input size for the model is fixed to 224 x 224. The volume size hence obtained is handled using maxpooling layers. This model is trained on the ImageNet dataset – a large database of millions of images of a wide variety. VGG16 generalizes well and can achieve the state of the art results. The pre-trained networks being freely available for VGG16, it is widely used for image classification problems (Liu and Deng, 2015)

In order to use the VGG16 for recognition and classification of 131 categories of fruits, the vanilla VGG16 model was followed by three additional layers. In the first, the layer of the output from the vanilla VGG16 model is flattened. This is then passed onto the second layer, a dropout layer with a probability of 0.3. Finally, a softmax classifier is used as the last layer. The 'layer.trainable' function for the VGG16 layers is kept as 'False'. Further, 'Adam' is used as an optimizer with a learning rate of 0.0001 along with 'categorical_crossentropy' as the loss function and 'accuracy' as the metrics. The validation split for both the base VGG16 model and VGG16 with image augmentation was kept as 0.3. The pretrained weights of ImageNet are used for training both the models. Even though the recommended input image size is 224 x 224, the original input image size (100 x 100) was retained to avoid possible blurring of the image. The results of both the base model and augmented model are presented below. Hence, objective 4 (ii) from chapter 1 Table 1 is achieved.

4.5.1 Base VGG16

The model was trained on the first set of train sets which is processed and mentioned in the first part of the pre-processing. The time taken by the 26.28 minutes. The training accuracy was 0.9998 whereas the test accuracy obtained was 0.9581. The model predicted all the images correctly (16/16) while verification (Figure 9). The plot of train vs. validation accuracy and loss (Figure 10) obtained was good indicating no overfitting or underfitting.

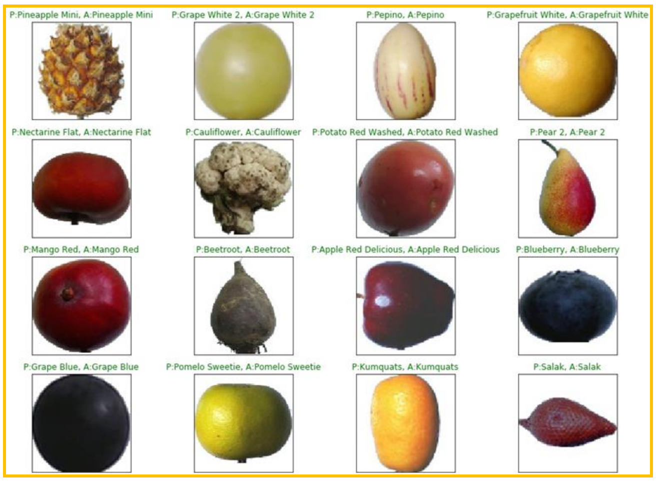

Figure 9 Random Prediction VGG16

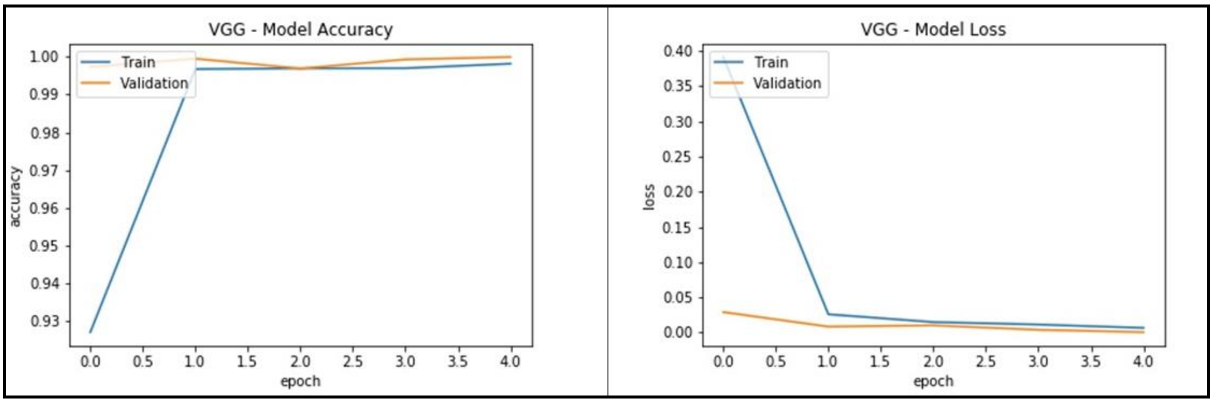

Figure 10 Accuracy and Loss Graphs of VGG16

4.5.2 VGG16 with Image Augmentation

The model was like the base model, the only difference being it was trained on the second set of train datasets i.e. the augmented one. The time taken by this model while training was 21.83 minutes. 0.9265 and 0.8376 train and test accuracy were achieved by the model. From Figure 12 it is evident that the model is overfitting and the validation set does not provide enough information to analyze the generalization ability of the model. Thus, while verifying the test accuracy the model predicted 11/16 images correctly (Figure 11).

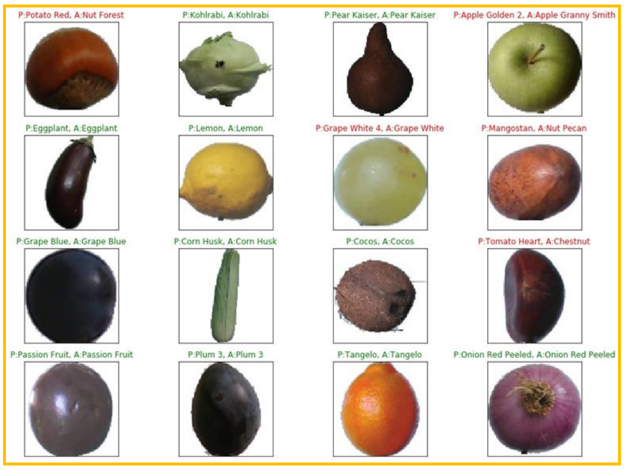

Figure 11 Random Prediction Augmented VGG16

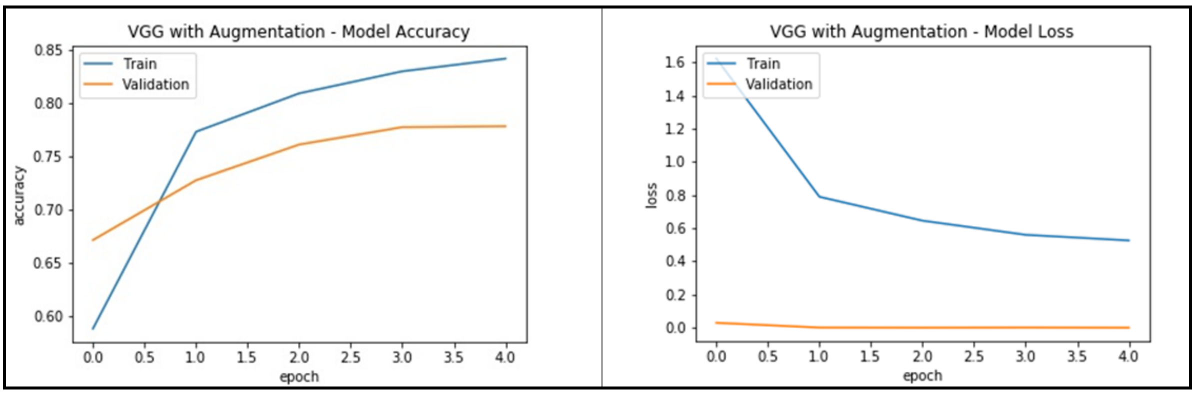

Figure 12 Accuracy and Loss Graphs of Augmented VGG16

4.6 Implementation of Residual Network 50 Model

ResNet as known as Residual Networks is one of the popular models in the pool of deep learning models and has also won the ImageNet challenge. As the name suggests this model consists of 50 layers and quite deeper than the prevision two models used in this project. A typical block of residual network contains layers consists of convolutional, ReLu, and batch normalization layers. The advantages of using ResNet50 is that it encourages the utilization of features, reinforce feature propagation, and cut down the number of parameters substantially. The specialty of ResNet is the identity connection which skips connection in between layers and adds the output from the previous layer to the layer is connected to in turn making the model deeper (He et al., 2016).

The vanilla ResNet50 model with 'imagenet' weights is modified by add 3 additional layers similar to the VGG16 model. The output from ResNet50 is flattened then passed to a dropout layer with a probability of 0.3 and finally, it passed to a softmax classifier with 131 units. The layers within the ResNet50 are kept untrainable. The model is complied with 'Adam' optimizer along with 'accuracy' metrics and 'categorical_crossentropy' as the loss function. However, to improve the performance of the model, the optimizer was changed to 'SGD'. The input image size was kept the same as the original size for the above-mentioned reasons. The validation split in the case of both the following models was kept as 0.3. The results of both the base model and augmented model are presented below. Hence, objective 4 (iii) from chapter 1 Table 1 is achieved.

4.6.1 Base ResNet50

This model similar to the previous two base models was trained using the first train set. The time taken by this model to complete training was 29 minutes. The accuracy attained by the model for training was 0.9962 and the test set was 0.9692. The model predicted all images correctly when tested on 16 random images of fruits from the test dataset (Figure 13). Figure 14 shows that the accuracy plot is a good fit whereas the loss plot indicates that the validation set does not provide the model sufficient information for generalization.

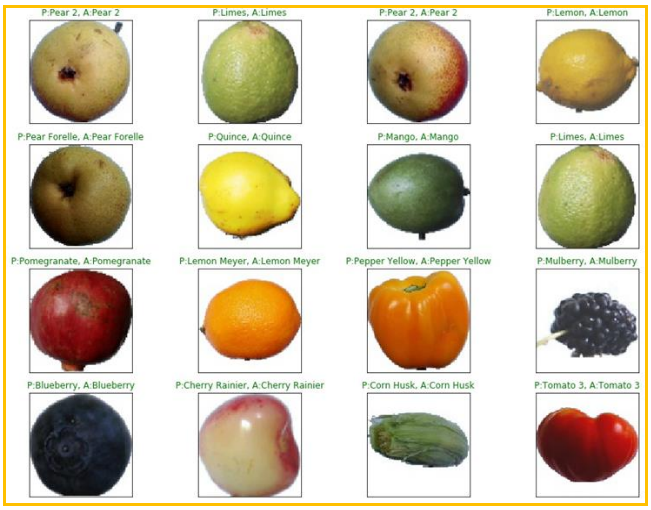

Figure 13 Random Prediction ResNet50

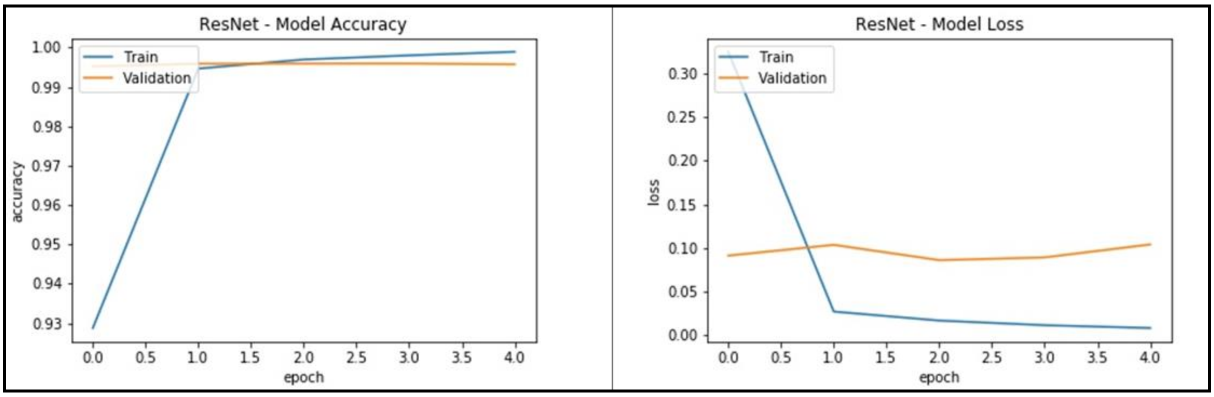

Figure 14 Accuracy and Loss Graphs of ResNet50

4.6.2 ResNet50 with Image Augmentation

The developed model was trained using the second set of training data. The time consumed by this model for training was 24.87 minutes. The accuracies achieved for training was 0.9728 and the test set was 0.9522. On testing against 16 random images from the test dataset, the model predicted 16/16 correctly (Figure 15). Further, the accuracy and loss plots shown in Figure 16 indicates that the model seems like a good fit with no underfitting/overfitting.

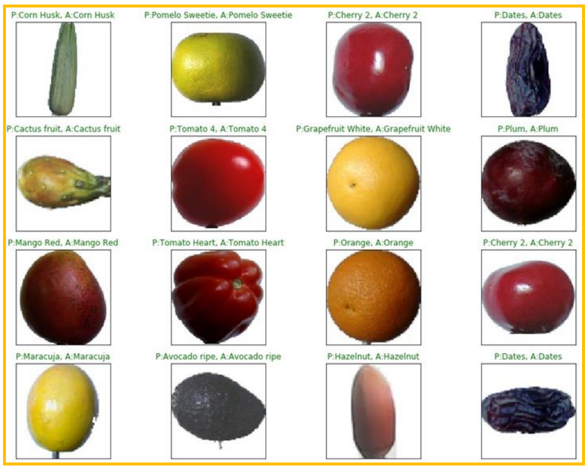

Figure 15 Random Prediction Augmented ResNet50

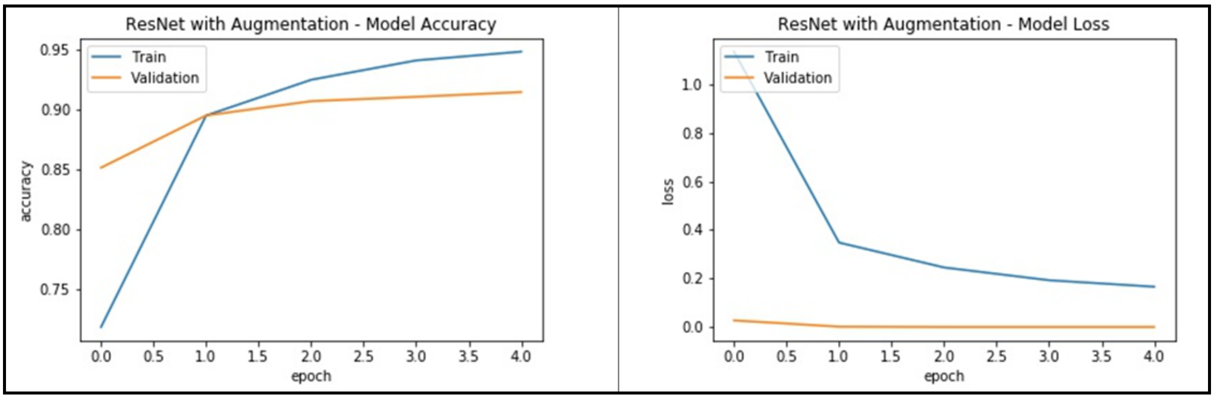

Figure 16 Accuracy and Loss Graphs of Augmented ResNet50

As per the objectives defined in Table 1 of chapter 1, this project successfully implements and achieves the objectives 1,2, 3, and 4. This in turn answers the research question (RQ) and sub-RQ.

5 Comparison of Developed Models and Discussion

The results of the models with and without image augmentation are summarised in Table 3. In the first part of the project, the deep learning models were employed on the first set of the training dataset. It is evident from Table 3 that all the base models accurately classified the fruits with all models achieving very high accuracies. The training and test accuracies of all models were more or less the same with a minor decimal difference. To verify the performance of the models, each model was tested by selecting random 16 images from the test set. All the base models correctly recognized and classified all the 16 images. Further, as evident from Figure 17, the execution time taken by the base CNN model was the least amongst all making it the best performing model in terms of computational power. In the second part of the project, the base models were trained using an augmented dataset to investigate whether it improves the performance of these models. However, it can be seen in Figure 17 and Table 3, that after image augmentation the computational time decreased for VGG16 and ResNet50 significantly, but the train and test accuracies also decreased for both the models. In the case of CNN, the computational time increased whereas both the train and test accuracies decreased. The decrease in accuracy of all three models can be attributed to the operations performed during augmentation which makes it harder for the model to make correct predictions. On the testing the augmented models against random 16 images, CNN classified 14/16 correctly, VGG16 classified 11/16 correctly and ResNet50 classified all images correctly. Amongst the augmented models, ResNet50 outperformed CNN and VGG16 in terms of accuracy but there was no significant difference in execution time as visible from Figure 17.

Table 3 Comparison of Developed Models

Models	**Base Model**	**Augmented Model**
CNN	Train Accuracy: 0.9994 Test Accuracy: 0.9691	Train Accuracy: 0.9319 Test Accuracy: 0.9256
VGG16	Train Accuracy: **0.9999** Test Accuracy: 0.9639	Train Accuracy: 0.8964 Test Accuracy: 0.8602
ResNet50	Train Accuracy: 0.9962 Test Accuracy: **0.9692**	Train Accuracy: **0.9728** Test Accuracy: **0.9522**

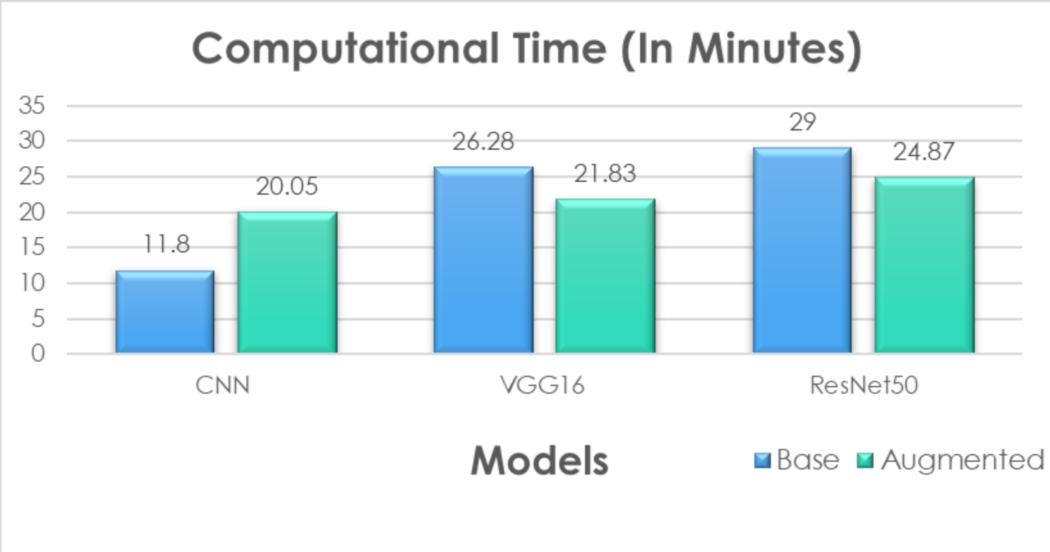

Figure 17 Comparison of Models based on Computational Time

Table 4 Comparison of Reviewed and Developed Techniques

Author	Fruit Type / Dataset Size	Technique	Accuracy
Hossain, Al-Hammadi and Muhammad (2019)	15 categories Dataset 1 = 2633 images Dataset 2 = 5946 images	VGG-16	Dataset 1 = 0.9975 Dataset 2 = 0.9675
Basri, Syarif and Sukaridhoto (2018)	Mango and Pitaya Dataset = 1400 images	Faster R-CNN	0.9900
Femling, Olsson and Alonso-Fernandez (2018)	10 categories Dataset = 4000 images	MobileNet	0.9700
Choi et al. (2018)	Pear Dataset = 1800 images	ANN	0.9740
Kausar et al. (2018)	81 categories Dataset = 55244 images	P-CNN	0.9888
Shubham Kathepuri (2020)	**131 categories Dataset = 90,483**	**Base CNN**	**Train: 0.9996 Test: 0.9691**

Further, on comparing the obtained results with the state-of-the-art models (Table 4), the developed model performs exceptionally well considering the large size of the dataset and the variety of fruits it can recognize and classify. In addition to that, the developed models also overcome the challenges faced by the researchers in differentiating between two similar-looking fruits such as tomatoes, and apple, and between fruits of the same category but with the different class such as red apple from green apple. Moreover, the project was also able to achieve the future work of using a dataset with a wide variety of fruits. If the developed models are commercialized, they will be able to significantly reduce human error, improve the process, and enhance profits for the end-user associated.

6 Conclusion and Future Work

In this project, a methodology for recognition and classification is developed after carrying out an extensive literature review. After a careful study of literature, deep learning models viz. CNN, VGG16, and ResNet50 were selected based on their performance in previous works. The project aimed to investigate how well can deep learning models (CNN, VGG16, ResNet50) perform when employed on a larger dataset with a wide variety of categories and does image augmentation enhances the performance of the models. To answer this, two sets of training dataset was formed, one was pre-processed while the other set was augmented using rescaling, rotation, width, and height shift, zoom, and horizontal flip operations. Further, the deep learning models were developed from scratch to achieve optimum results. The developed models were trained and tested against the two sets of data. ResNet50 performed better on both the dataset with training accuracy as 0.9962 and 0.9728 on the first and second sets of data, respectively. The test accuracy achieved by CNN on the first set was 0.9692 and on the second set was 0.9522. On passing 16 random images of fruits, ResNet50 was able to classify them correctly. However, the performance of VGG16 and CNN on the first set of data was similar to that of ResNet50 but it performed poorly when employed on augmented images. The developed models not only perform well on a large dataset with a wide variety of fruits but were also able to tend to the issues faced by the researchers in the literature work.

Although the developed models performed well and were able to achieve future works of previous researchers, the models had certain shortcomings of its own. First, the lack of images of fruits in the natural environment as opposed to what is used for this project. Second, the dataset used also lacked variance of fruit within a category i.e. a category say blueberry contains multiple images of single blueberry rather than different blueberries. Third, due to computational limits and available timeframe to complete the project, the models were not tested commercially. In the future, researchers can form a dataset with a wide variety of fruits images captured in the natural environment or with at least some background along with high intra category variance which will make the system more robust. Images of rotten fruits can also be included in the dataset which will expand the scope of the study and allow the models to differentiate between fresh and rotten fruits. The developed system can be employed for real-life applications and their performance can be monitor which only makes the system more reliable.

Acknowledgment

The completion of this project would not have been possible with continuous guidance from Dr. Catherine Mulwa even during these hard times. I would like to thank Dr. Catherine Mulwa, my family, and my friends for the immense support and encouragement they provided me throughout this project.

References

Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017) ‘Understanding of a convolutional neural network’, in *2017 International Conference on Engineering and Technology (ICET)*, Antalya, Turkey, 21-23 August 2017, pp. 1-6, IEEE Xplore. doi: 10.1109/ICEngTechnol.2017.8308186.

Arivazhagan, S., Newlin Shebiah, R., Selva Nidhyanandhan, S., Ganesan, L., (2010) ‘Fruit Recognition using Color and Texture Features’, Journal of Emerging Trends in Computing and Information Sciences, 1(2), pp. 90-94.

Azevedo, A., and Santos, M. F. (2008) ‘Kdd, Semma and Crisp-Dm: A Parallel Overview’, in *2008 IADIS European Conference Data Mining*, Amsterdam, Netherland, 24-26 July 2008 pp. 182-185.

Basri, H., Syarif, I. and Sukaridhoto, S. (2018) ‘Faster R-CNN Implementation Method for Multi-Fruit Detection using Tensorflow Platform’, in *2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, Bali, Indonesia, 29-30 October 2018, pp. 337–340, IEEE Xplore. doi: 10.1109/KCIC.2018.8628566.

Bhargava, A. and Bansal, A. (2019) ‘Automatic Detection and Grading of Multiple Fruits by Machine Learning’, Food Analytical Methods, 13(3), pp.751-761, Complementary Index. doi: 10.1007/s12161-019-01690-6.

Bresilla, K., Perulli, G., Boini, A., Morandi, B., Corelli Grappadelli, L. and Manfrini, L. (2019) ‘Single-Shot Convolution Neural Networks for Real-Time Fruit Detection Within the Tree’, Frontiers in Plant Science, 10, pp. 1-12.

Buzzelli, M., Belotti, F. and Schettini, R. (2018) ‘Recognition of Edible Vegetables and Fruits for Smart Home Appliances’, in *2018 IEEE 8th International Conference on Consumer Electronics,* Berlin, Germany, 2-5 September 2018, pp. 1–4, IEEE Xplore. doi: 10.1109/ICCE-Berlin.2018.8576236.

Buzzelli, M., Belotti, F. and Schettini, R. (2018) ‘Recognition of Edible Vegetables and Fruits for Smart Home Appliances’, *2018 IEEE 8th International Conference on Consumer Electronics (ICCE)*, Berlin, Germany, 2-5 September 2018, pp. 1–4, IEEE Xplore. doi: 10.1109/ICCE-Berlin.2018.8576236.

Choi, H. S., Cho, J. B., Kim, S. G. and Choi, H. S. (2018) ‘A Real-Time Smart Fruit Quality Grading System Classifying by External Appearance and Internal Flavor Factors’, in *2018 IEEE International Conference on Industrial Technology (ICIT),* Lyon, France, 20-22 February 2018, pp. 2081–2086, IEEE Xplore. doi: 10.1109/ICIT.2018.8352510.

Desai, M. (2019) ‘A Fruit Recognition Approach for Refrigerator Inventory Management’, in *2019 International Conference on Issues and Challenges in Intelligent Computing*

Techniques (ICICT), Ghaziabad, India, 1-6 September 2019, pp.1-6, IEEE Xplore Digital Library. doi: 10.1109/ICICT46931.2019.8977691.

Femling, F., Olsson, A. and Alonso-Fernandez, F. (2018) ‘Fruit and Vegetable Identification Using Machine Learning for Retail Applications’, *in 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, Las Palmas de Gran Canaria, Spain, 26-29 November 2018, pp. 9–15, IEEE Xplore. doi: 10.1109/SITIS.2018.00013.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) ‘Deep Residual Learning for Image Recognition’, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 27-30 June 2016, pp. 770–778, IEEE Xplore. doi: 10.1109/CVPR.2016.90.

Hossain, M. S., Al-Hammadi, M. and Muhammad, G. (2019) ‘Automatic Fruit Classification Using Deep Learning for Industrial Applications’, IEEE Transactions on Industrial Informatics, Industrial Informatics, 15(2), pp. 1027–1034.

Jana, S., Basak, S. and Parekh, R. (2017) ‘Automatic Fruit Recognition from Natural Images using Color and Texture Features’, in *2017 Devices for Integrated Circuit (DevIC)*, Kalyani, India, 23-24 March 2017, pp. 620-624, IEEE Xplore. doi: 10.1109/DEVIC.2017.8074025.

Kausar, A., Sharif, M., Park, J. and Shin, D. R. (2018) ‘Pure-CNN: A Framework for Fruit Images Classification’, in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, 12-14 December 2018, pp. 404–408, IEEE Xplore. doi: 10.1109/CSCI46756.2018.00082.

Khan, R. and Debnath, R. (2019) ‘Multi Class Fruit Classification using Efficient Object Detection and Recognition Techniques’, International Journal of Image, Graphics and Signal Processing, 11(8), pp.1-18.

Lamb, N. and Chuah, M. C. (2018) ‘A Strawberry Detection System Using Convolutional Neural Networks’, in *2018 IEEE International Conference on Big Data*, Seattle, WA, USA, 10-13 December 2018, pp. 2515–2520, IEEE Xplore. doi: 10.1109/BigData.2018.8622466.

Liu, S. and Deng, W. (2015) ‘Very Deep Convolutional Neural Network Based Image Classification using Small Training Sample Size’, in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, Kuala Lumpur, 3-6 November 2015, pp. 730-734, IEEE Xplore. doi: 10.1109/ACPR.2015.7486599.

Mahajan, M.S. (2016), ‘Optimization and Classification of Fruit using Machine Learning Algorithm’, International Journal for Innovative Research in Science & Technology, 3(1), pp. 377-381.

Mai, X., Zhang, H., Jia, X., and Meng. M. Q. (2020) ‘Faster R-CNN With Classifier Fusion for Automatic Detection of Small Fruits’, IEEE Transactions on Automation Science and Engineering, Automation Science and Engineering, 17(3), pp. 1555–1569.

Peng, H., Shao, Y., Chen, K., Deng, Y. and Xue, C. (2018) ‘Research on Multi-class Fruits Recognition Based on Machine Vision and SVM’, IFAC-PapersOnLine, 51(17), pp.817-821, Elsevier BV. doi: 10.1016/j.ifacol.2018.08.094.

Saranya, N., Srinivasan, K., Pravin Kumar, S. K., Rukkumani, V. and Ramya, R. (2020) ‘Fruit Classification Using Traditional Machine Learning and Deep Learning Approach’, Computational Vision and Bio-Inspired Computing, 1108, pp. 79-89.

Shukla, D. and Desai, A. (2016) ‘Recognition of Fruits using Hybrid Features and Machine Learning’, in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, Pune, India, 19-21 December 2016, pp. 572–577, IEEE Xplore. doi: 10.1109/CAST.2016.7915033.

Zawbaa, H. M., Hazman, M., Abbass, M. and Hassanien, A. E. (2014) ‘Automatic Fruit Classification using Random Forest Algorithm’, in *2014 14th International Conference on Hybrid Intelligent Systems,* Kuwait, 14-16 December 2014, pp. 164–168, IEEE Xplore. doi: 10.1109/HIS.2014.7086191.

Zhang, W., Zhang, Y., Zhai, J., Zhao, D., Xu, L., Zhou, J., Li, Z. and Yang, S. (2018) ‘Multi-Source Data Fusion using Deep Learning for Smart Refrigerators’, Computers in Industry, 95, pp. 15-21.

Zhang, Y., Wang, S., Ji, G. and Phillips, P. (2014) ‘Fruit Classification using Computer Vision and Feedforward Neural Network’, Journal of Food Engineering, 143, pp.167-177.