TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

HIMALAYA COLLEGE OF ENGINEERING

A

MAJOR PROJECT REPORT ON

VEGETABLE RECOGNITION AND PRICE VIEWER SYSTEM

[CT-755]

SUBMITTED TO:

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

CHYASAL, LALITPUR

SUBMITTED BY:

RABIN GHIMIRE (073/BCT/28)
ANIL SHRESTHA (073/BCT/50)
AJIT DHAMI (073/BCT/51)

March 5, 2021

VEGETABLE RECOGNITION AND PRICE VIEWER SYSTEM

A FOURTH YEAR MAJOR PROJECT REPORT
[CT-755]

"A FOURTH YEAR PROJECT REPORT SUBMITTED FOR PARTIAL FULFILLMENT OF DEGREE OF BACHELOR'S IN COMPUTER ENGINEERING"

PROJECT SUPERVISOR
ER. SUDARSHAN SUBEDI

SUBMITTED TO
Tribhuvan University
Institute of Engineering
Himalaya College of Engineering
Department of Electronics and Computer Engineering
Chyasal, Lalitpur

SUBMITTED BY
RABIN GHIMIRE (073/BCT/28)
ANIL SHRESTHA (073/BCT/50)
AJIT DHAMI (073/BCT/51)

March 5, 2021

ACKNOWLEDGEMENT

In all humility and gratitude, we are over-helmed to acknowledge our profundity to all those who helped us put our ideas and assigned work well above the level of simplicity and into something concrete.

To our HOD, Er. Ashok GM Sir, we would like to express our sincere gratitude for the continuous support for the completion of the proposal of our Major Project, for his patience, motivation, enthusiasm and immense knowledge.

We would also like to show our gratitude to our project coordinator **Er. Sudarshan Subedi Sir** for sharing his pearls of wisdom with us during the course of this project. We are very thankful to our mentors for their valuable suggestions, guidance and encouragement.

Finally, we would also like to express our gratitude to all our friends and family for their moral support, cooperation and encouragement without which the completion of the project would be incomplete.

Rabin Ghimire (073/BCT/28)
Anil Shrestha (073/BCT/50)
Ajit Dhami (073/BCT/51)

ABSTRACT

The popular technology used in this innovative era is Computer vision for vegetables recognition. Compared to other machine learning (ML) algorithms, deep neural networks (DNN) provide promising results to identify vegetables in images. Currently, to identify vegetables, different DNN-based classification algorithms are used. However, the issue in recognizing vegetables has yet to be addressed due to similarities in size, shape and other features. This project discusses the use of deep learning (DL) for recognizing vegetables and its other applications. The project will also provide a concise explanation of convolution neural networks (CNNs) to recognize vegetables.

The programming language that we use to develop our system is Python. We will collect the images of different vegetables.Also the price will be viewed alongside the recognized vegetable . There will be different applications of the project. Consumers can use the system to get the price of the vegetable and its current price that helps them to prevent from getting paid overpriced.

Keywords:
AI, Deep Learning, machine learning, Neural Networks, Python

LIST OF ABBREVIATIONS

AI	Artificial Inteligence
OpenCV	Open Computer Vision
CNN	Convolutional Neural Networks
DNN	Deep Netral Network
ANN	Artificial Neural Networks
ML	Machine Learning

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1. INTRODUCTION

1.1 Background

Recognizing different kinds of vegetables is perhaps the most difficult task in supermarkets and fruit shops. Retail sales systems based on bar code identification require the seller (cashier) to enter the unique code of the given fruit or vegetable because they are individually sold by weight. This procedure often leads to mistakes because the seller must correctly recognize every type of vegetable and fruit; a significant challenge even for highly-trained employees. A partial solution to this problem is the introduction of an inventory with photos and codes. Unfortunately, this requires the cashier to browse the catalog during check-out, extending the time of the transaction. In the case of self-service sales, the species (types) and varieties of vegetables must be specified by the buyer. Unsurprisingly, this can often result in the misidentification of vegetables by buyers (e.g., Conference pear instead of Bartlett pear). Independent indication of the product, in addition to both honest and deliberate mistakes (purposeful indication of a less expensive species/variety of fruit/vegetable) can lead to business losses. The likelihood of an incorrect assessment increases when different fresh products are mixed up.

One potential solution to this challenge is the automatic recognition of vegetables and vegetables. The notation of recognition (identification, classification) can also be understood in different ways: as the recognition of a vegetables (distinguishing a fruit from another object, e.g., a leaf, a background), recognizing the species of a fruit (e.g., apple from a pear), and recognizing a variety of a given species of fruit (e.g., Golden Delicious apples from Gloster apple). In the case of retail systems, the last two applications have special significance. The concept of fruit classification best reflects the essence of the issue discussed in the article as a way of automatically determining the right species and variety of vegetables. Classification of vegetables is a relatively complex problem owing to the huge number of varieties. Considerable differences in appearance exist within species and varieties, including irregular shapes, colors, and textures. Furthermore, images range widely in lightning conditions, distance, and angle of the camera; all of which result in distorted images. Another problem is the partial or full occlusion of the object. These constraints have led to the lack of multi-class automated fruit and vegetable classification systems in real-life applications.

Vegetable have great relevance for humans because of their nutritional value. Consequently, research on vegetable processing is very important for several economic sectors, both for the wholesale and retail markets, as well as for the processing industries. Hence, different methods have been developed to automatically process vegetable, either to classify them or to efficiently estimate their quality and price.

1.2 Objectives and Scope

The main objective of our system are as follows:

- To recognize the vegetable from the real time live camera.
- To get the current market price of the recognized vegetable.

2. LITERATURE REVIEW

Computer vision is one of the most used technological tools in the agro-industrial field, both in automatic fruit harvesting, fruit sorting machines, and fruit scanning in supermarkets. All vision systems typically include different types of data generated by sensors or cameras. This data can be RGB images, RGB depth images (RGB-D), hyperspectral images, among many other types. So that, due to different computational methods and algorithms, required features must be extracted and processed to perform the corresponding task to the fruit industry sector. For example, in supermarkets, a fruit recognition process is required or in an orchard for harvest, the accurate detection of fruit [1].

In the area of image recognition and classification, the most successful results were obtained using artificial neural networks. This served as one of the reasons we chose to use a deep neural network in order to identify vegetables from images. Deep neural networks have managed to outperform other machine learning algorithms. They also achieved the first superhuman pattern recognition in certain domains. This is further reinforced by the fact that deep learning is considered as an important step towards obtaining Strong AI. Secondly, deep neural networks ,specically convolutional neural networks have been proved to obtain great results in the field of image recognition. We will present a few results on popular datasets and the used methods.

Nowadays, artificial intelligence (AI) is a field with several practical applications in a wide range of industries and active research topics. The main challenge for AI is to solve the tasks that people intuitively solve, but hard to implement computationally. Therefore, AI systems must have the ability to acquire their knowledge, extracting raw data patterns, which is known as machine learning. Thus, AI-based techniques are very useful to solve complex problems where traditional methods would not be efficient [2].

Machine learning (ML) allows researchers and developers to computationally address problems related to the knowledge of the real world. ML endows computers with the ability to act without being explicitly programmed, building algorithms to recognize patterns on the data and make predictions based on it. ML-based systems are applied in several areas, such as information analysis, agriculture, ecology, mining, urban planning, defense, space exploration, among others [3].

Deep Learning is the sub-field of Machine Learning, which is the sub-field of Artificial Intelligence. It is a collection of techniques that model high-level abstractions in data. In deep learning, a computer-based statistical model understands and learns from pictures, sound, or text to conduct analysis. These models can attain state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers in term of accuracy [4].

While the concept of deep learning was first put forward back in the 1980s, the idea subsequently became popular because of two reasons: it needs a huge amount of labeled data and substantial computing power. The number of deep learning applications has been experiencing research growth in the last decade, including natural language processing, image classification, and information retrieval, etc. The deep learning term could be divided into two parts and understand them individually: deep and learning. Learning is about taking previous understanding and information and creating an inner depiction of the matter that the agent can use to act. Typically, the internal depiction is a compact representation for summarizing the data. The field of Machine Learning offers different functions and techniques for learning automatically from the available information, and this learning from the information is used for forecasting and projections in the future [5].

Currently, deep learning (DL) is one of the most used ML-based methods. An important characteristic of DL is that it has high levels of abstraction and the ability to automatically learn patterns present in images. Particularly, Convolutional Neural Network (CNN) is the main DL architecture used for image processing. CNNs is a kind of artificial neural networks (ANNs) that use convolution operations in at least one of their layers . Since 2012, when Krizhevsky et al. won the ImageNet competition (ILSVRC), CNNs have gained great popularity as an efficient method for image classification in many fields. Specifically in agriculture, CNN-based approaches have been used for fruit classification and fruit detection[6].

Convolutional neural networks (CNN) are part of the deep learning models.Such a network can be composed of convolutional layers, pooling layers,ReLU layers, fully connected layers and loss layers. In a typical CNN architecture, each convolutional layer is followed by a Rectified Linear Unit (ReLU) layer, then a Pooling layer then one or more convolutional layer and finally one or more fully connected layer. A characteristic that sets apart the CNN from a regular neural network is taking into account the structure of the images while processing them. Note that a regular neural network converts the input in a one dimensional array which makes the trained classifier less sensitive to positional changes. Among the best results obtained on the MNIST dataset is done by using multi-column deep neural networks. As described in paper, they use multiple maps per layer with many layers of non-linear neurons. Even if the complexity of such networks makes them harder to train, by using graphical processors and special code written for them. The structure of the network uses winner-take-all neurons with max pooling that determine the winner neurons. Another paper further reinforces the idea that convolutional networks have obtained better accuracy in the domain of computer vision. In paper an all convolutional network that gains very good performance on CIFAR-10 is described in detail. The paper proposes the replacement of pooling and fully connected layers with equivalent convolutional ones. This may increase the number of parameters and adds inter-feature dependencies however it can be mitigated by using smaller convolutional layers within the network and acts as a form of regularization. In what follows we will describe each of the layers of a CNN network[7].

3. REQUIREMENT ANALYSIS

3.1. FEASIBILITY STUDY

The following points describes the feasibility of the project.

3.1.1. Economic Feasibility

The total expenditure of the project is just computational power. No other expenses is required to perform this project. The tools required for the project are easily accessible in internet and are free of cost. Thus, the project is economically feasible.

3.1.2. Technical Feasibility

The proposed project will be developed using python programming language. There are various libraries available which could be used to perform various operations involving natural language recognition. We will be making use of CNN (Convolutional Neural Network) Algorithm. The algorithms which is proposed is efficient on cost as well as complexity basis. The project is technically feasible

3.1.3. Operational Feasibility

The proposed system is feasible in terms of real world application as well. The proposed system is feasible in terms of real-world application as well. It performs its task very well. As talking about task, recognition of fruit and labelling its price is the main task that this project does. So, it is feasible operationally.

3.2 REQUIREMENT ANALYSIS

4.1 Functional Requirements

- The system must provide the accurate recognition of vegatable.
- The system must have an easy to use interface for using the system for all the users.
- System should provide a graphical interface for the prediction generated and for the User Interface.
- The system must be able to view the accurate price of the vegatable at today's market.
- The dataset of the images must be available for the system.
- Internet is required to fetch the price from the web.

4.2 Non-Functional Requirements

- <u>Scalability</u>

  The proposed system should be scalable and accommodate a number of vegetables based on the implementation.
- <u>Error Free</u>

  The proposed system should be error free and should function in proper manner.
- <u>Portability</u>

  For the portability of the system, it should run properly in different hardware.
- <u>Maintainability</u>

  The Proposed system should be easy to maintain.

4. SYSTEM DESIGN

4.1 Diagrams

4.1.1 Flowchart Diagram

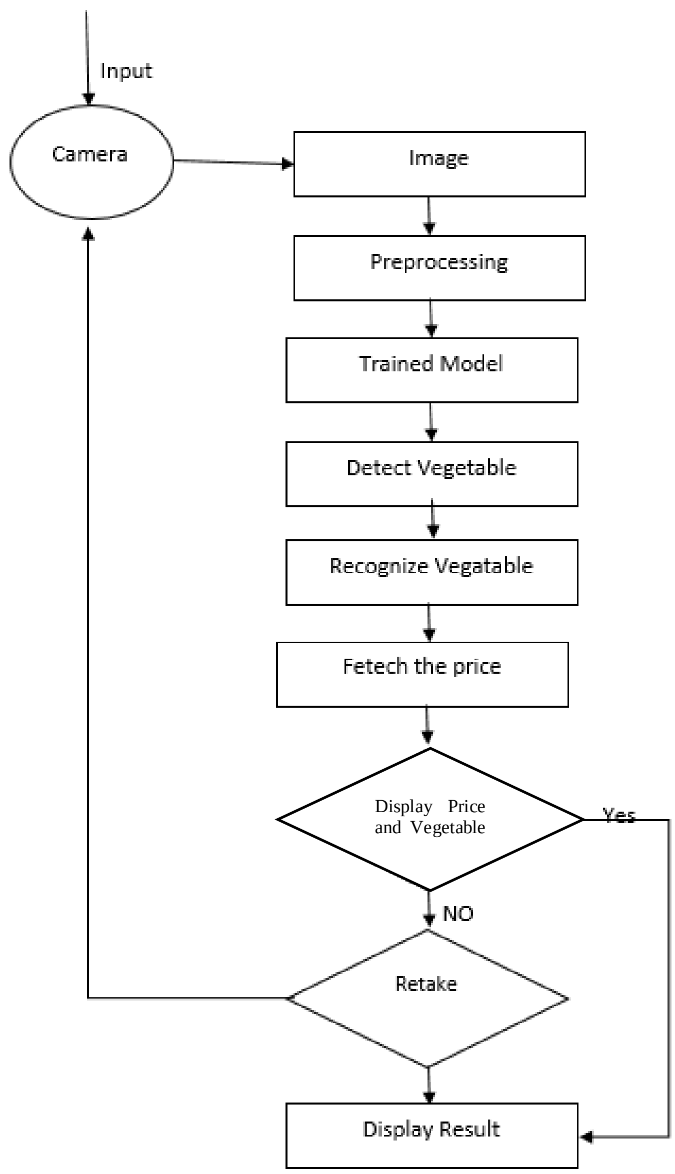

Figure 1 Flowchart Diagram

4.1.2 Block Diagram of System

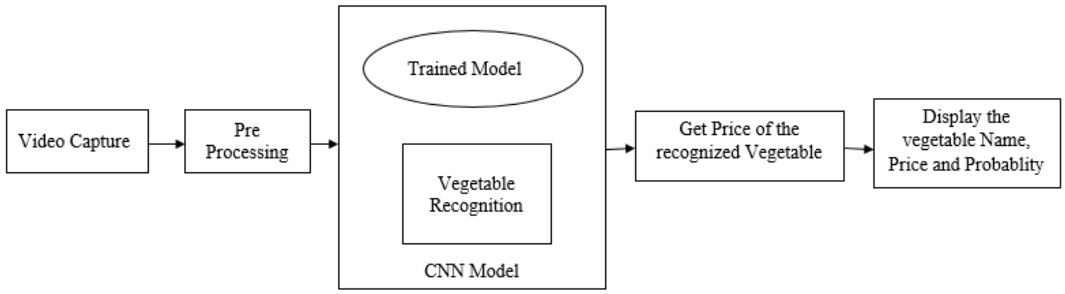

Figure 2 Block Diagram of System

4.1.2 Sequence Diagram

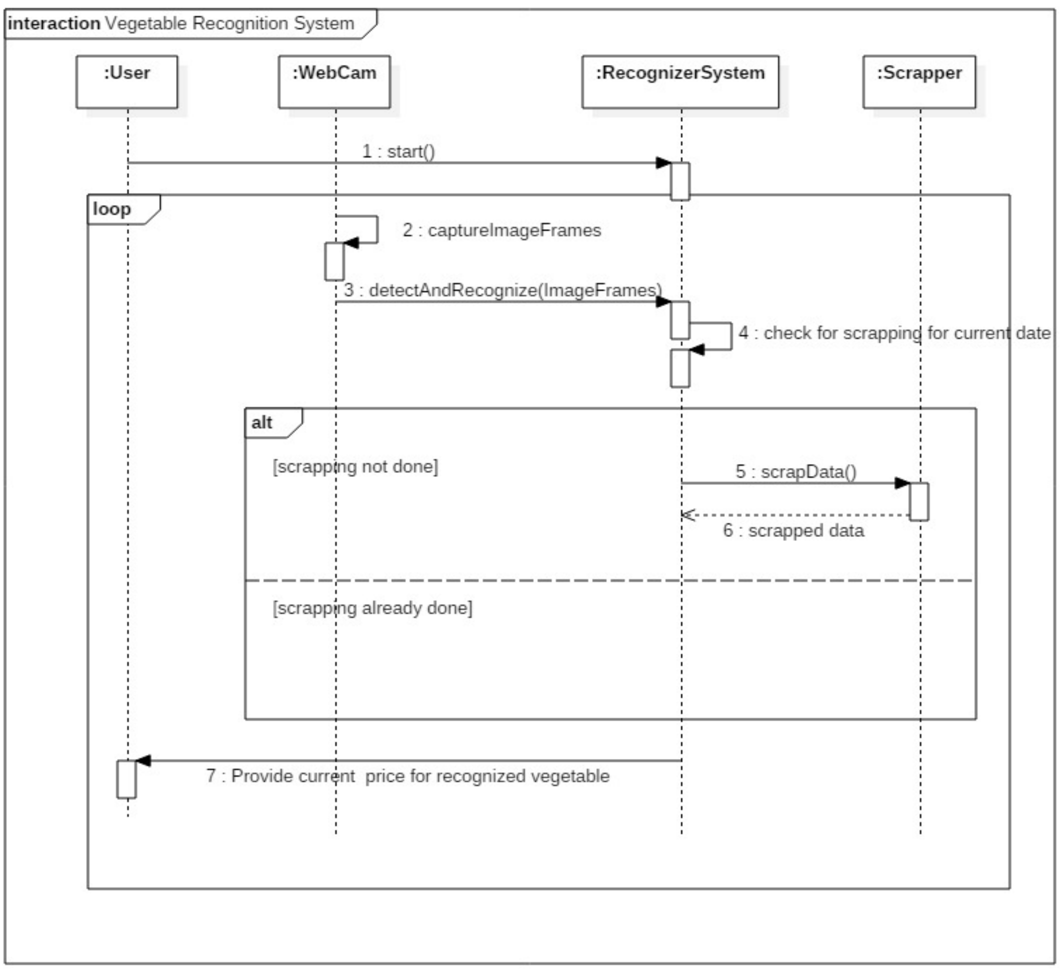

Figure 3 Sequence Diagram

4.1.3 Activity Diagram

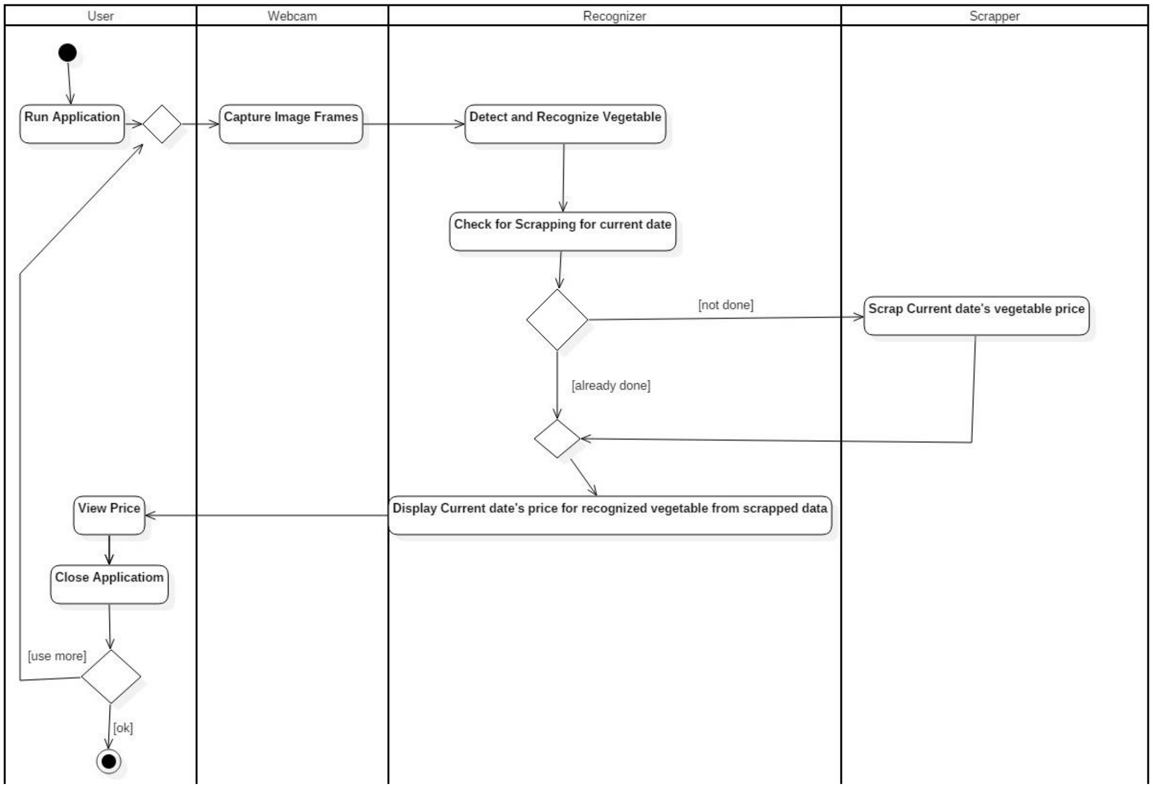

Figure 4 Activity Diagram

4.1.4 Use Case Diagram

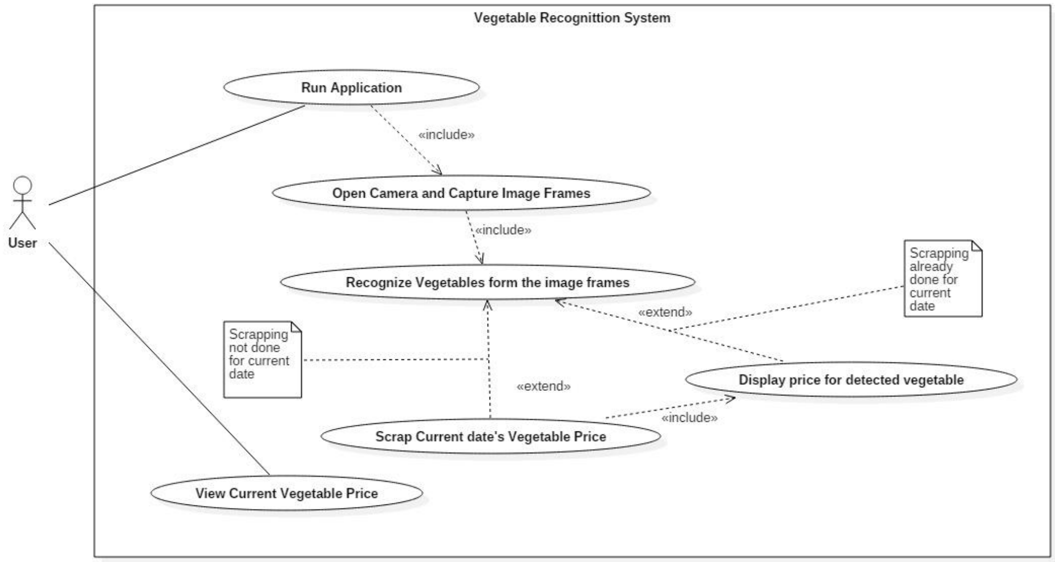

Figure 5 Use Case Diagram

4.2 Tools

- Python
- Tensorflow
- OpenCV
- Pandas
- Keras
- Matplotlib
- NumPy
- Anaconda
- PyCharm
- BeautifulSoup
- PyQt5

5. METHODOLOGY

The different steps required for the development of the system are as explained below:

5.1 Collection of Vegetables images

When recognizing real-world objects, they always vary a lot in their background, image quality, lighting etc. Taking this in account and trying to create as realistic dataset as possible. Different images are captured in different conditions such as different background and lighting conditions. In order to maintain the accuracy, the collections of images plays the vital role. Fruit 360 was used where more than 90000 picture of fruits and vegetables are present. In this project, all the required images are used from the fruit 360 dataset. More than 1300 picture of onion, potato and tomato were taken from the dataset that has white background.

5.2 Preparing the Images to Train

The images of three different vegatables; Onion, Tomato and Potato has been taken from fruit360. The image size used for the project is 50 * 50. Those imaged were saved in data folder with individual folders for each set of vegetables. Thus, the images were classified into four classes where class 0 contain tomato, class 1 contain potato and class 2 contain potato. These images are kept into different paths/directory and are labeled.

5.3 Training the images

The images are ready to train. Convolutional Neural Netwok is used to train the images.The training process may take time depending on the system. The labels.csv contains all the information about the class id and name of vegetables. CNN has input, hidden and output layer. The input layer contains the 32 filter of filger size 3*3. Again the hidden convolution layer contains 64 filter of size 3*3 and 128 filter of filter size 2*2. Then the output is flatten and dense. ReLU (Rectified Linear Unit) layer is used for the activation function. These detail are mentioned in detail in chapter 5.6 of this report.

5.4 Set up for the live images from camera module

OpenCV is used for the live video capturing from the camera. It is the vast library that helps in providing various functions for image and video operations. It lets you create a video capture object which is helpful to capture videos through webcam and then you may perform desired operations on that video. To capture a video, VideoCapture object is created. Here, the vegetables is detected as soon as the system detects any images of vegetables in the live video. The frame of the video is taken and the frame is used to predict the vegetable that is shown in the camera.

5.5 Web Scrapping

With the technique of web scrapping, as soon as the system detects any vegetable in the live video, the system gives the current market price of detected vegetable. Python library Beautiful Soup is being used to carry the web scrapping operation. Beautiful Soup is a python package for parsing HTML and XML documents. It pull particular content from a webpage, remove the HTML markup, and save the information. So, it is used to pull the current price of detected vegetable from kalimatimarket.gov.np websites. For eg: if tomato is shown in the camera, the system should detect the tomato and show Tomato: Rs. 50 in the video, where Rs 50 is the current price of tomato.

5.6 GUI

PyQt5 is used for making simple GUI to make the system easy to use. A window is displayed that has a simple UI, consist of Button labeled “Open camera”. It opens the camera and starts to capture video. On upper left side of window, it displays the vegetable name, matching probablity and the today’s price of recognized vegetable. It also consist of label that display the name of our project.

5.7 Algorithm

5.7.1 CNN Algorithm

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

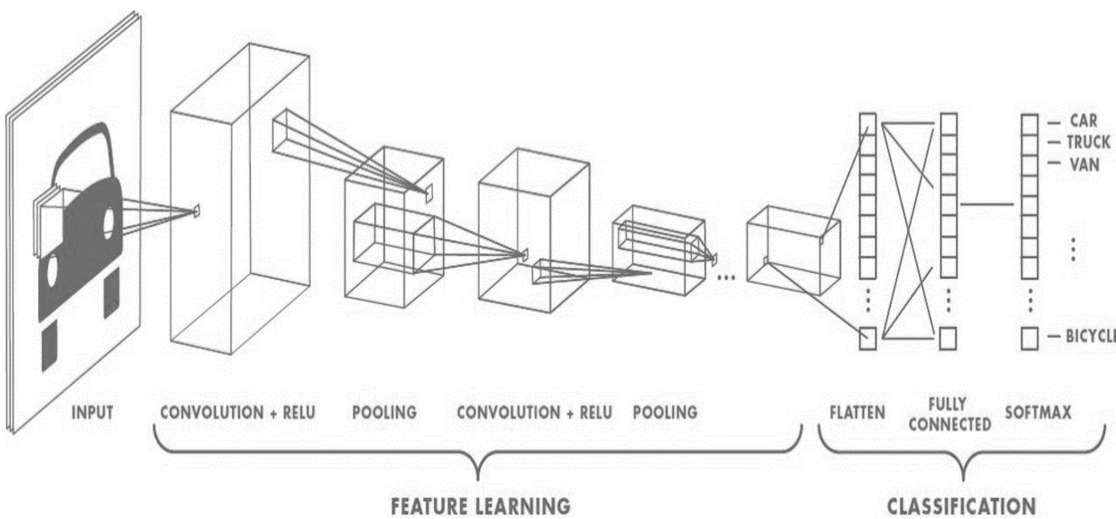

Figure 6 Convolutional Neural Networks Algorithm

CNN Architecture of our Project:

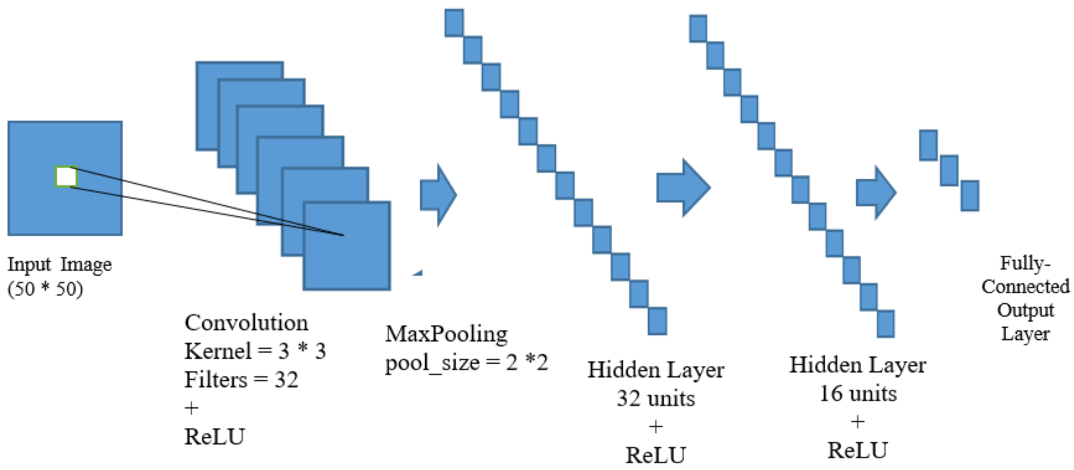

Figure 7 CNN Architecture of our Project

The Convolution Neural Network is sequentially build to get high-level features through their successive layers. The different layers in our CNN model are explained below:

Input Layer : All the input layer does is read the image. It has 32 units and the kernel size is (3*3). The activation function used in the layer is ReLU activation function.

Convolution Layer: The convolution layer uses filters that perform convolution operations while scanning the input image with respect to its dimensions. Its hyperparameters include the filter size 3x3 with the 32 filter layers. The resulting output is called the feature map or activation map and has all the features computed using the input layers and filters.

Pooling Layer: The pooling layer is used for downsampling of the features and is typically applied after a convolution layer. The two types of pooling operations are called max and average pooling, where the maximum and average value of features is taken, respectively.Maxpooling is performed with the pool_size of (2*2).

Hidden Layers: The Neural Network consist of 2 hidden layer. The first hidden layer has 32 filters and second hidden layer has 16 filters. Both hidden layer uses ReLU activation function.

Output Layer/Fully Connected Layer: The output layer in a CNN is a fully connected layer, where the input from the other layers is flattened and sent so as the transform the output into the number of classes as desired by the network. Here, The output layer is consist of 3 classes and activated using the softmax activation function to get the Fully Connected Layer.

6. SYSTEM TESTING

5.1 Unit Testing

5.1.1 Testing for data_aug.py

Test CaseID	Test Scenario	Test Step	Test Data	Actual Result	Observed Result	Pass/Fail
T1	Check the data argumentation operation.	Execute the data_aug.py	File from data directory	Two Pickle file should be generated containing information about dataset images.	Pickle file is generated	Pass

Table 1 Testing of data_aug

Testing is done on creating the pickle file. When data_aug.py file is executed. After the execution is completed, two pickle file X.pickle and y.pickle is created. This is the file which contain the information of the datasets and its images. X.pickle file contain the features of the images and y.pickle file contain labels of the images.

5.1.2 Testing for camera module

Test CaseID	Test Scenario	Test Step	Test Data	Actual Result	Observed Result	Pass/Fail
T2	Run cv2v2.py file	Camera module should open and view the video.	Images from camera.	Camera module should open and view the video.	Camera is opened and video is shown.	Pass

Table 2 Testing of camera module

In order to test if the camera module is opened or not, cv2v2.py file is run. After the completion, camera module is opened as shown in the table.

5.1.3 Testing for recognition of vegetable

Test CaseID	Test Scenario	Test Step	Test Data	Actual Result	Observed Result	Pass/Fail
T3	Recognition of Onion	Run cv2v2 and show onion near camera.	Onion image from camera.	Should show class [0]: Onion , probability value and price of the displayed vegetable	Result is shown as expected.	Pass
T4	Recognition of Potato	Run cv2v2 and show potato near camera.	Potato image from camera.	Should show class [2]: Potato , probability value and price of the displayed vegetable	Result is shown as expected.	Pass

Table 3 Testing for recognition of vegetable

Here, recognition of vegetable is tested where Onion and Potato are shown on the camera. The camera takes the images and is preprocessed then the processed image is compared with the trained module. When compared, it detects the vegatable and name, probability is displayed.

5.1.4 Testing for UI open camera

Test Case ID	Test Scenario	Test Step	Test Data	Actual Result	Observed Result	Pass/Fail
T5	Test the UI file.	Run testframe.py file	Execute the testframe.py	The window should appeared that has open camera button	The window is appeared that has open camera button	Pass
T6	Test the open camera button	Click the open camera button after executing testframe.py	Click the open camera button	Should open the camera module	Camera module is opened	Pass

Table 4 Testing for UI open camera

Here, the UI windows is opened by running testframe.py file where open camera button is located somewhere in the middle. On clicking the open camera button, camera module is opened.

5.2 Integration Testing

Test CaseID	Test Scenario	Test Step	Test Data	Actual Result	Observed Result	Pass/Fail
T7	Test for data argumentation	Run data_aug.py	Image Data from data filder	Output should be in pickle file	Result obtained as expected	Pass
T8	Train the images using CNN	Execute CNN.py	Images from data folder	CNN should execute in 10 epoch	Result obtained as expected	Pass
T9	UI and Open Camera Button	Execute testframe.py and Click the open camera button.	Execute and Click the open camera button	Should view the window and open camera button.	Result obtained as expected	Pass
T10	Detect the fruit	Show the onion after executing testframe.py and open camera button	Show onion near camera	Should show class[0]: Onion, probability and Price	Result obtained as expected	Pass

Table 5 Integration Testing

7. RESULT AND ANALYSIS

The overall system was tested for any faults and errors and the final result or output was analyzed. Firstly, the UI of the system is simple and user friendly.The CNN has been tuned to get the better accuracy in recognition. The plot of accuracy and plot of loss that consist of two curve training and validation can be seen below. The secondary aim of the project i.e to view the accurate price of the vegetable is also accomplished. After the integration of different module, the system perferms perfectly as expected.

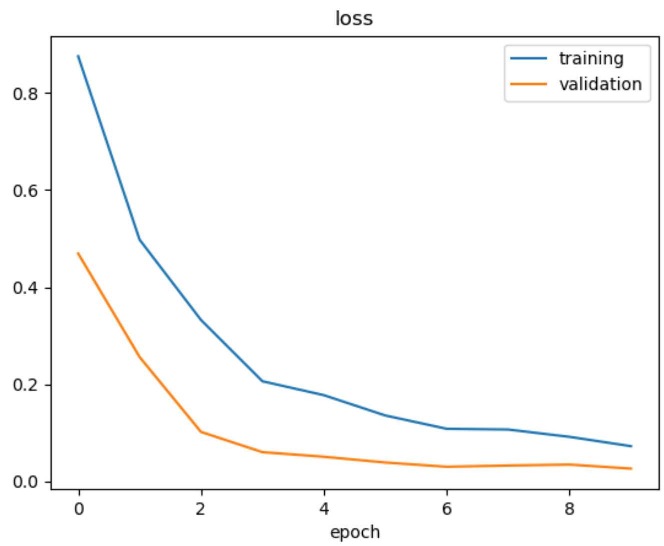

Figure 8 Loss Plot

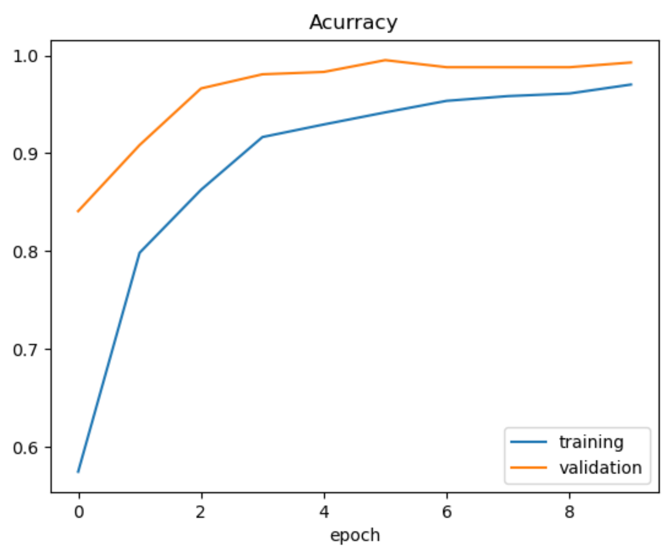

Figure 9 Accuracy Plot

8. CONCLUSION AND FUTURE ENHANCEMENT

Overall the project has been a success. A fully functional system has been developed in order to support the users. Apart from the minimum requirements that has been implemented some extra enhanced functionality is provided by the application. At the end the evaluation showed that the analysis, extraction of requirements design and implementation were handled correctly. All the deadlines have been met meaning that the schedule chosen was suitable for the project. After completing the project the following result are obtained.

- the vegetable recongnition from the real time live camera.
- the current market price of the recognized vegetable is displayed.

As far as we could, we tried our best to make the project perfect and real. Even though there are some limitations that we were unable to implement in the application despite of our effort. Some of them are discussed below:

- Limited no of vegetales
- Requires high quality camera to detect.

Enormous knowledge has been gained throughout the project work. The importance of the background research, requirement analysis and specifications, well designing concept, and superior methodology were learnt. Also implementation techniques, testing, error handling, optimization issues and the predictability of problems such as when to perform a certain task, have been exercised. Thus we hope that the system developed will certainly contribute for the development of vegetable recognition and view the current market price.

References

[1] E. Rachmawati, I. Supriana and Khodra, "In Proceedings of the 2017 IEEE 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI),," in *Toward a new approach in fruit recognition using hybrid*, Yogyakarta, Indonesia, September 2017, pp. 19-21.

[2] B. Coppin, " Jones & Bartlett Learning," in *Artificial Intelligence Illuminated*, Burlington, MA, USA, Jones & Bartlett Learning, 2004.

[3] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu and F. Alsaadi, "A survey of deep neural network architectures and their Applications," in *Neurocomputing*, 2017.

[4] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu and M. Lew, "Deep learning for visual understanding," pp. 27-48, 2016.

[5] B. U. T. P. a. C. M. F Dayoub, "A fruit detection system using deep neural networks," Sensors, 2012, p. 1222.

[6] G. H. a. B. K. L Deng, "Rethinking Model Scaling for Convolutional Neural Networks," IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, p. 1905.

[7] L. D. Y. S. M. B. H. Cheng, "Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks," in *Journal of Imaging* , 2017.

APPENDIX