

**Tribhuvan University**

**Institute of Science and Technology**

**Himalaya College of Engineering**

**[Code: CSC-404]**

A

Final Year Project Defense Report

On

“**FRUIT RECOGNITION SYSTEM USING CNN ALGORITHM**”

By

**Kushal Poudel[20]**

**Mahesh Panday[22]**

**Samrat Ghimire[34]**

**Upendra Acharya[45]**

Submitted to Department of Computer Science and Information Technology in partial fulfillment of the requirement for the Bachelor’s degree in Computer Science and Information Technology

**Department of Computer Science and Information Technology**

[February 9, 2021]

# ACKNOWLEDGEMENT

We would like to express our sincere thanks to the department of BSc.CSIT for providing us the opportunity to explore our interest and ideas in the field of computer science and technology through this project.

Without the support, patience and enthusiasm of the following people this paper would not have been completed. It is to them, we owe our sincere gratitude. We would like to acknowledge and express our big thanks to everyone for his/her support and encouragement which has inspired us for this project. We would like to acknowledge and extend our heartfelt gratitude to the following persons who have helped us for this project:

Er. Himal Chand Thapa Head of Department, BSc.CSIT for his vital encouragement and support; Mr. Drish Mali, Project Supervisor, for his guidance and effort to provide us inspiration for this project. Last but not the least we would like to thank all our colleagues for providing continuous feedback and support in the course of this project.

# ABSTRACT

The popular technology used in this innovative era is Computer vision for fruits recognition. Compared to other machine learning (ML) algorithms, deep neural networks (DNN) provide promising results to identify fruits in images. Currently, to identify fruits, different DNN-based classification algorithms are used. However, the issue in recognizing fruits 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 fruits and its other applications. The project will also provide a concise explanation of convolution neural networks (CNNs) to recognize fruits.

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

***Keywords:***

***AI, Deep Learning, machine learning, Neural Networks, Python***

TABLE OF CONTENTS

[ACKNOWLEDGEMENT 2](#_Toc64031909)

[ABSTRACT 3](#_Toc64031910)

[LIST OF ABBREVIATION 5](#_Toc64031911)

[LIST OF FIGURES 1](#_Toc64031912)

[1. INTRODUCTION 2](#_Toc64031913)

[1.1 Introduction 2](#_Toc64031914)

[1.2 Objectives and Scope 3](#_Toc64031915)

[2. LITERATURE REVIEW 4](#_Toc64031916)

[3. FEASIBILITY STUDY 6](#_Toc64031917)

[3.1. Economic Feasibility 6](#_Toc64031918)

[3.2. Technical Feasibility 6](#_Toc64031919)

[3.3. Operational Feasibility 6](#_Toc64031920)

[4. METHODOLOGY 7](#_Toc64031921)

[4.1 Collection of Fruits images 7](#_Toc64031922)

[4.2 Creating and training the images 7](#_Toc64031923)

[4.3 Set up for the live images from camera module 7](#_Toc64031924)

[4.4 Web Scrapping 7](#_Toc64031925)

[4.5 Algorithm 8](#_Toc64031926)

[4.5.1 CNN Algorithm 8](#_Toc64031927)

[4.6 Tools 10](#_Toc64031928)

[4.7 Diagrams 11](#_Toc64031929)

[4.7.1 Flowchart Diagram 11](#_Toc64031930)

[4.7.2 Sequence Diagram 12](#_Toc64031931)

[4.7.3 Class Diagram 13](#_Toc64031932)

[5. Testing 14](#_Toc64031933)

[5.1 Unit Testing 14](#_Toc64031934)

[5.1.1 Testing for pickle file 14](#_Toc64031935)

[5.1.2 Testing for CNN model 14](#_Toc64031936)

[5.1.3 Testing for camera module 14](#_Toc64031937)

[5.1.4 Testing for recognition of fruit 14](#_Toc64031938)

[5.2 Integration Testing 15](#_Toc64031939)

[6. Conclusion 16](#_Toc64031940)

# LIST OF ABBREVIATION

AI : Artificial Intelligence

OpenCV : Open Computer Vision

CNN : Convolutional Neural Networks

DNN : Deep Netral Network

ANN : Artificial Neural Networks

ML : Machine Learning

# LIST OF FIGURES

[Figure 4‑1 Convolutional Neural Networks Algorithm [8] 8](file:///D:\kushal%20project\final%20year%20project\Tribhuvan%20University%20.docx#_Toc64031879)

[Figure 4‑0‑2 CNN used for the project 9](file:///D:\kushal%20project\final%20year%20project\Tribhuvan%20University%20.docx#_Toc64031880)

[Figure 4‑3 Flowchart of the system 11](file:///D:\kushal%20project\final%20year%20project\Tribhuvan%20University%20.docx#_Toc64031881)

[Figure 4‑4 Sequence diagram of the system 12](#_Toc64031882)

# LIST OF TABLES

Table

# 1. INTRODUCTION

## Introduction

With the lively improvement of our human society, additional attention has been paid to the superiority of our life, particularly the food we eat. Over the last few years, computer visions have been widely used in fruit recognition methods. In the field of image recognition and classification, Deep Neural Network (DNN) is used to identify fruits from images. DNN performs better than other machine learning algorithms. Convolutional Neural Networks (CNNs) are classified as a deep learning algorithm. In deep learning, CNN are the most commonly used type of Artificial Neural Networks (ANNs). It is being used several visual recognitions analyzing which includes video and image recognition, face recognition, handwritten digit recognition, and fruit recognition etc. The accuracies in these fields including fruit recognition using CNN have reached human-level perfection

CNN has a very similar architecture as ANN. There are several neurons in each layer in ANN. Hence, the weighted sum of all the neurons of a layer becomes the input of a neuron of the next layer adding a biased value. In CNN the layer has three dimensions. Here all the neurons are not fully connected instead they are connected to the local receptive field. A cost function is generated in order to train the network. It compares the network’s output with the desired output. Accurate and efficient fruit recognition is of great importance in the field of robotic harvesting and yield mapping. An ideal fruit recognition system is accurate that can be trained on an easily available dataset, shows real-time predictions and acclimates various types of fruits. Therefore, in our research, we implement a fruit recognition classifier using CNN. The input image is taken as 100×100 pixels of RGB image. For the network’s best performance, we used various combinations of hidden layers for five cases and observe the accuracies. The final experiment result shows the much-improved fruit recognition rate. The mathematical model of the network is executed in python with Tensorflow. OpenCV has been used to open the live camera module in order to detect the fruit. After detecting, price is shown by web scraping using BeautifulSoup from the Kalimati vegetable and fruit website. For the dataset, Fruit360 from Kaggle is used where more than 90000 photos are present of different Vegetables and Fruits.

Fruits have great relevance for humans because of their nutritional value. Consequently, research on fruits 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 fruits, 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 fruits from the real time live camera.
* To get the current market price of the recognized fruits.

# 2. LITERATURE REVIEW

Computer vision is one of the foremost used innovative devices within the agro-industrial field, both in programmed fruit gathering, fruit sorting machines, and natural product checking in general stores. All vision frameworks ordinarily incorporate diverse sorts of information created by sensors or cameras. This data can be RGB pictures, RGB depth pictures (RGB-D), hyperspectral pictures, among numerous other sorts. So that, due to distinctive computational strategies and algorithms, required features must be extracted and handled to perform the comparing task to the fruit industry segment. For example, in grocery stores, a natural product recognition process is required or in a plantation to collect, the precise detection of fruit [1].

In the range of image recognition and classification, the most effective results were gotten using artificial neural networks. This served as one of the reasons we chose to utilize a deep neural network in order to identify fruits 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, especially 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) may be a field with a few practical applications in a wide range of industries and active research topics. The main challenge for AI is to solve the tasks that individuals naturally solve, but difficult to execute computationally. In this manner, AI frameworks must have the capacity to obtain their information, extricating raw information patterns, which is known as machine learning. In this way, AI-based strategies are exceptionally valuable to solve complex issues where traditional strategies would not be effective [2].

Machine learning (ML) permits analysts and developers to computationally address issues related to the knowledge of the real world. ML blesses computers with the capacity to act without being explicitly programmed, building algorithms to recognize patterns on the data and make predictions based on it. ML-based systems are connected in a few zones, such as information analysis, farming, biology, mining, urban arranging, defense, space investigation, 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 methods that show high-level abstractions in data. In deep learning, a computer-based statistical model understands and learns from pictures, sound, or text to conduct examination. These models can achieve state-of-the-art precision, in some cases exceeding human-level execution. Models are trained by employing a huge set of labeled data and neural network architectures that contain numerous 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 considering 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. FEASIBILITY STUDY

The following points describes the feasibility of the project.

## 3.1. 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.

## 3.2. 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. CNN (Convolutional Neural Network) Algorithm has been used in this project. The algorithms which is proposed is efficient on cost as well as complexity basis. The project is technically feasible

## 3.3. Operational Feasibility

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.

# 4. METHODOLOGY

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

## 4.1 Collection of Fruits 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. For testing, fruit 360 was used where more than 90000 picture of fruits and fruits are present. In this project, all the pictures are captured through mobile camera. More than 1000 picture of onion, potato and lemon were captured and has white background. In order to maintain the accuracy, the collections of images plays the vital role.

## 4.2 Creating and training the images

The collected images of fruits should be classified and dataset should be made. This can be done with the help of Tensorflow and Convolutional neural network algorithm. Thus, the images were classified into four classes where class 0 contain tomato, class 1 contain lemon and class 2 contain potato. These images are kept into different paths/directory and are labeled. Now, Tensorflow records is created. Training pipeline is made to train the images. The training process may take time depending on the system. The dataset .csv is generated as the dataset for fruits detection. CNN has input, hidden and output layer. Convolution of filter size 3\*3 and 32 filter layers are made and the input image is given. Again the convolution output is given for the convolution of filter size 3\*3 and 64 filter layers two times. Then the output is flatten and dense. ReLU (Rectified Linear Unit) layer is used for the activation function.

## 4.3 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 fruits is detected as soon as the system detects any images of fruits in the live video.

## 4.4 Web Scrapping

With the technique of web scrapping, as soon as the system detects any fruits in the live video, the system gives the current market price of detected fruits. Python library Beautiful Soup is being used to carry the web scrapping operation. BeautifulSoup is a python package for parsing HTML and XML documents. It creates a parse tree for parsed pages that can be used to extract data form HTML, which is used for web scraping. 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 fruits from kalimatimarket.gov.np websites. For eg: if apple is shown in the camera, the system should detect the apple and show Apple: Rs. 50 in the video, where Rs 50 is the current price of apple.

## 4.5 Algorithm

### 4.5.1 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 overlaps to cover the entire visual area.

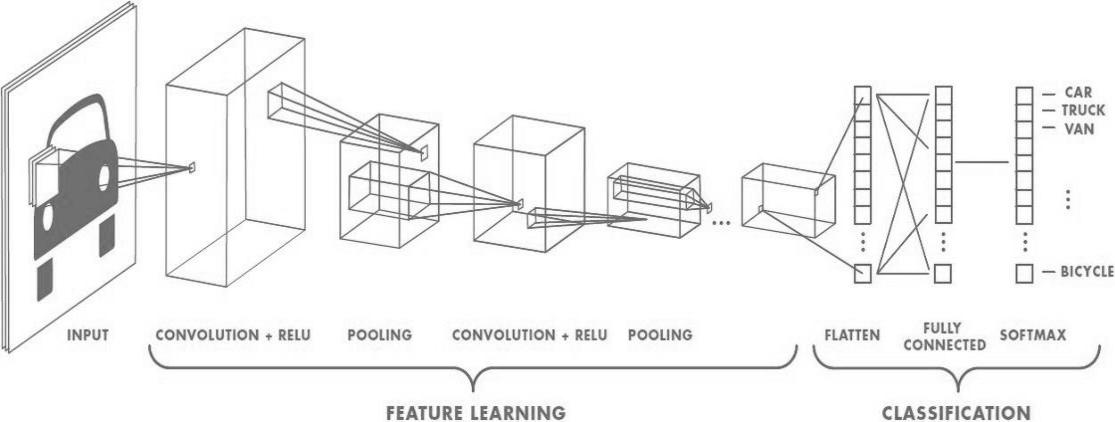


Figure 4‑1 Convolutional Neural Networks Algorithm [8]

A CNN has

* Convolutional layers
* ReLU layers
* Pooling layers
* A fully connected layer

A CNN

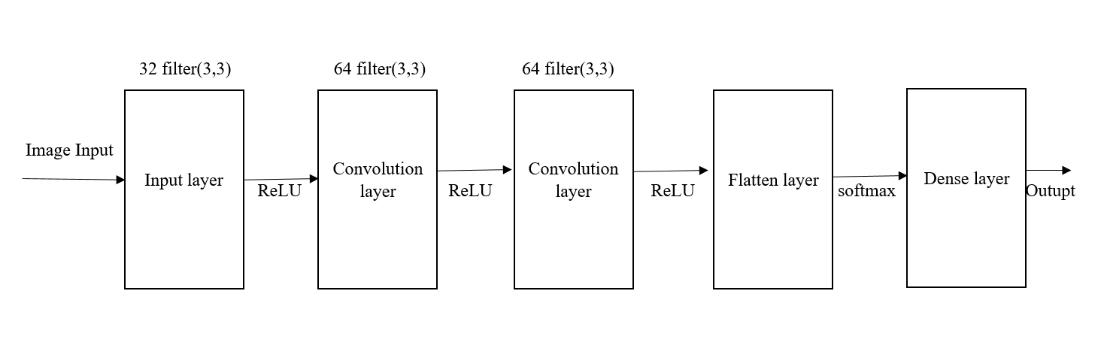
* Starts with an input image
* Applies many different filters to it to create a feature map
* Applies a ReLU function to increase non-linearity
* Applies a pooling layer to each feature map
* Flattens the pooled images into one long vector.
* Inputs the vector into a fully connected artificial neural network. Processes the features through the network. The final fully connected layer provides the “voting” of the classes that we’re after.
* Trains through forward propagation and backpropagation for many, many epochs. This repeat until we have a well-defined neural network with trained weights and feature

Figure 4‑0‑2 CNN used for the project

Convolutional Neural Network Algorithm is used as the backbone for the project. Convolution of filter size 3\*3 and 32 filter layers are made and the input image is given. Again, the convolution output is given for the convolution of filter size 3\*3 and 64 filter layers two times. Then the output is flatten and densed. ReLU (Rectified Linear Unit) layered for the activation function.

## 4.6 Tools

Python:

In this project, python programming language has been used. It is the best language for Machine learning and Deep Learning projects. Python's simple syntax and readability promote rapid testing of complex algorithms and make the language accessible to non-programmers.

Tensorflow:

Tensorflow is used for training the model of this project. Tensorflow set side by side with python as it is the programming language used in this project. It helps to train the image dataset and create the trained model.

OpenCV:

In order to get the real-time image, OpenCV library has been used. It helps to get the real time image of fruit.

PyCharm:

The best IDE for Python is PyCharm. As it provides a wide range of essential tools for python developers. So, PyCharm is used as the IDE for this project.

BeautifulSoup:

As the feature of displaying price of fruits detected, web scrapping is to be done. So, to get the price from the Kalimatimarket.gov.np, BeautifulSoup has been used.

## 4.7 Diagrams

### 4.7.1 Flowchart Diagram

Diagram

Description automatically generated

Figure 4‑3 Flowchart of the system

This is the flowchart of the system where input is given from the camera. The given input, live image of fruit, is firstly preprocessed where grayscale conversion, Histogram equalization and normalization is done. After preprocessing, the image is compared with the trained model where it predicts the given input. If the accuracy of predicted fruit is greater than 65% then its name and today’s price is displayed. If accuracy is less than 65 then nothing is displayed. This process is then repeated afterward.

### 4.7.2 Sequence Diagram

Diagram

Description automatically generated

Figure 4‑4 Sequence diagram of the system

User opens the application where open camera button can be seen. After clicking the button, camera open where desired fruit can be shown. Then the displayed fruit is compared with the trained CNN model. Then as explained above in flowchart diagram, if accuracy is greater than 65, it displays the name of the fruit else fruit not found is displayed. This is the sequential flow of the system which is represented by above sequence diagram.

### 4.7.3 Class Diagram

# 5. Testing

## 5.1 Unit Testing

### 5.1.1 Testing for pickle file

Table 5.1 Testing for pickle file

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.N | Input | Expected Output | Actual Output | Remarks |
| 1 | Run data file | Pickle file should be created containing information about dataset images. | Pickle file is created. | Is working properly |

### 5.1.2 Testing for CNN model

Table 5.2 Testing for CNN model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.N | Input | Expected Output | Actual Output | Remarks |
| 1 | Run CNN file | CNN model should be created | CNN model is created | Is working properly |

### 5.1.3 Testing for camera module

Table 5.3 Testing for Camera module

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.N | Input | Expected Output | Actual Output | Remarks |
| 1 | Run opencv file | Camera module should open to detect the fruit | Camera module is opened. | Is working properly |

### 5.1.4 Testing for recognition of fruit

Table 5.4 Testing for recognition of fruit

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.N | Input | Expected Output | Actual Output | Remarks |
| 1 | Show Apple(fruit) on camera | Should show Name and accuracy of the displayed fruit | Name: Apple  Accuracy:89.09%  Shown | Is working properly |

## 5.2 Integration Testing

# 6. Conclusion

**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.