# Chapter 1

# INTRODUCTION

## ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction. Particular applications of AI include expert systems, speech recognition and machine vision. These technologies reduce human effort. Now in many industries, people are using this technology to develop machine slaves to perform the different activity. Using the machine for the work speed up your process of doing work and give you an accurate result.

The introduction of AI brings the idea of error free world. This technology will slowly introduce in all the sector to reduce human effort and give accurate and faster result.

## SPYDER

Spyder is a powerful scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It offers a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.

## ALGORITHM

**DECISION TREE CLASSIFIER**:

A decision tree is a flowchart-like tree structure where an internal node represents feature (or attribute), the branch represents a decision rule, and each leaf node represents the outcome. The topmost node in a decision tree is known as the root node. It learns to partition on the basis of the attribute value.

**NAVIE BAYES ALGORITHM:**

In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter.

**SVM ALGORITHM:**

In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very.

## PROJECT SCOPE

Recognition of fruits in the natural environment becomes challenging due to the involvement of complex backgrounds. This problem has also been addressed in the literature. Their approach is based on the principle that there is a big difference in curvatures between edge points on the edge produced by occlusion and the edge points on the edge without occlusion. In the classification based methods, first the system train a classifier with the feature descriptors of the training images and finally classify the test images to a category using trained classifier and feature descriptor of the test image.

## Project objectives

- In this project we will classifies the fruit accordingly.
- The data is taken from the dataset and it is read in the program.
- The data in the dataset is manipulated and checked whether the value matches correctly.
- The output is registered and it is displayed to the user.

## Project goals

- **Planned Objective:** It will classify the fruit whether it is orange, lemon or apple according to data we have given.
- **Accuracy:** The way we made in this project give accurate results.
- **Reliability:** The unwavering quality of the application framework will be high because of the above reasons. The accuracy of the data will be in a best way because of the expanded dependability.
- Eliminating manual classification of fruit with this device saves lots of time and energy and side-by-side decreases the error.

# Chapter 2

# LITERATURE REVIEW

Different creator, over the years, have intend distinct methods for classification of fruits. Bhanu and Navneet usage shape (area, perimeter, major axis length and minor axis length), color (mean and variance of HIS, HSV, L*A*B and YbCbCr components) and texture (GLCM texture form of contrast, dissimilarity, angular second moment, energy and entropy) form and categorized them using artificial neural net Hashem Tamimi converse a fruit and vegetable acknowledgment system which necessity histograms of dye and purity components and performance and Chi-rule minimum distance technique as classifier. Chromaticity interior the RGB Time blended with shape capabilities extracted the use of Fourier descriptors, Hu moments and simple geometrical capabilities are utilized by Farid for the assortment second-hand synthetic satisfy before neural network.

Another advanced uses shape (area, perimeter and roundness value), color (signify of RGB value), and structure(entropy) with next adjoining classifier. The intend process describe fruit acknowledgment extending multiple form supported analysis that includes texture, color and shape. To recognize the texture of a fruit the Log Gabor filter out has been used, mean dye has been calculated for color and shape has been analyzed by counting perimeter and area pixel. Thus, SVM is found to be a better classifier due to its higher accuracy. Thus, color and texture features should be taken into consideration for accurate identification and classification of fruits using BPNN.

# Chapter 3

# SYSTEM ANALYSIS AND DESIGN

## 3.1 PROBLEM DESCRIPTION

- Consider a scenario where you are working peacefully in your fruit stock factory.
- The problem may occur in our presence and maybe in our absence also.
- We cannot always monitor the situations. So we introduce this to classify.
- But you are wasting your precious time in classifying the fruits.

## 3.2 EXISTING SYSTEM

Identification of fruits and vegetables are fulfilling indifferent areas. The most usual areas are identification in the retail business, and in areas where the intention is to ease the harvest in the perspective of agriculture. In the retail business, the identification is mostly done manually by a cashier, or via the particular-service systems in a supply Extends existing orchard management systems and control layer. Adapt platforms from adjacent sectors (especially arable farming) to fruit -specific requirements. Implement a generic artificial intelligence and control platform in the fruit domain. This can also be implemented using image processing technique.

### 3.3 PROPOSED SYSTEM

In the former works, intra-class acknowledgment of fruits has been execute mainly using outline shape features. However, the shape can change depending on the view point of the camera with respect to the object of interest. Moreover, some fruits like grapes, are always available in gather, and extracting single shapes from a gather captured at different sight angles, is a non-ordinary task. To address these problems, the immediate work uses a combination of color and texture for intra-class recognition.

### 3.4 BENEFITS OF PROPOSED SYSTEM

- Proven practical usability in the fruit domain
- No change of software vendor necessary
- Long term commitment of the vendor to the fruit sector
- Domain knowledge available at the vendors

### 3.5 SOFTWARE REQUIRED

- Spyder.
- SVM Library.

# Chapter 4

# DIAGRAMS

## 4.1 BLOCK DIAGRAM

The developed methodology is consisting of three steps. In the first step we will analyze the given data. In the second step feature is extracted from the dataset. In this we proposed a texture feature to enhance the classification accuracy. In the last step classification of fruit take place on the basis of knowledge gain during training phase.

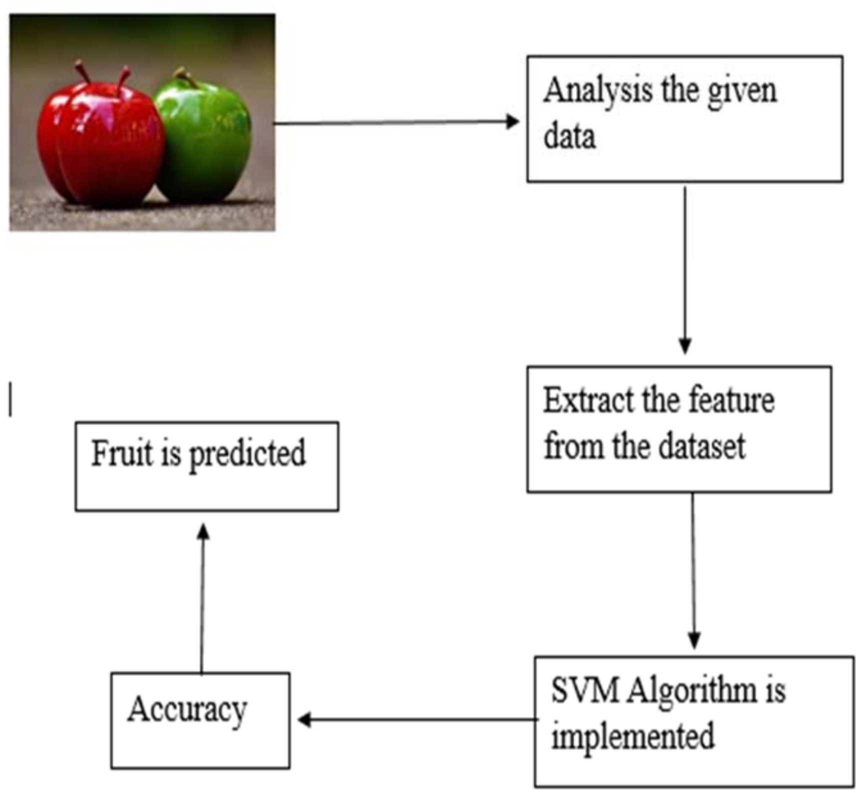


**Figure 4.1.1 (Block diagram)**

## 4.2 FLOW DIAGRAM

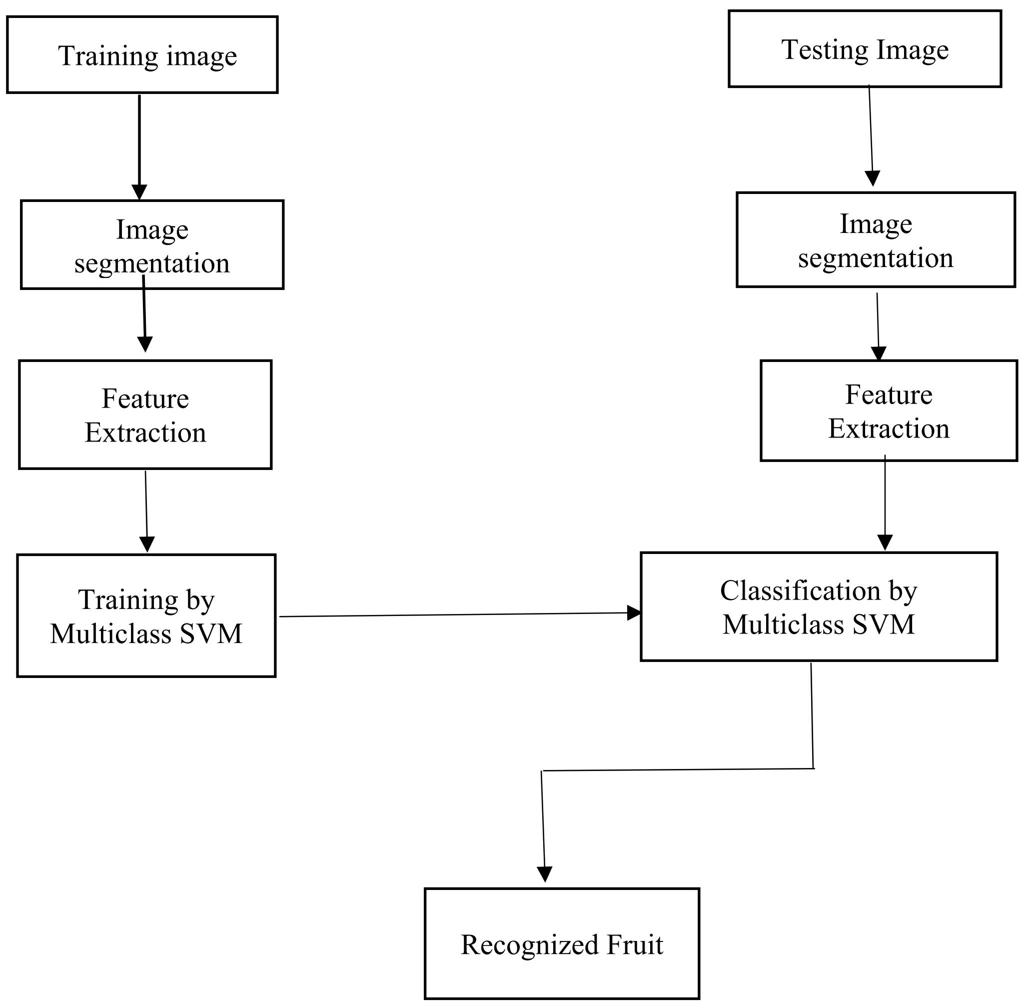


**Figure 4.2.1 Flow diagram**

## Chapter 5

## IMPLEMENTATION

### 5.1 DATASET

| fruit_label | fruit_name | fruit_subtype | mass | width | height | color_score |
|---|---|---|---|---|---|---|
| 1 | apple | granny_smith | 192 | 8.4 | 7.3 | 0.55 |
| 1 | apple | granny_smith | 180 | 8.0 | 6.8 | 0.59 |
| 1 | apple | granny_smith | 176 | 7.4 | 7.2 | 0.60 |
| 2 | mandarin | mandarin | 86 | 6.2 | 4.7 | 0.80 |
| 2 | mandarin | mandarin | 84 | 6.0 | 4.6 | 0.79 |
| 2 | mandarin | mandarin | 80 | 5.8 | 4.3 | 0.77 |
| 2 | mandarin | mandarin | 80 | 5.9 | 4.3 | 0.81 |
| 2 | mandarin | mandarin | 76 | 5.8 | 4.0 | 0.81 |
| 1 | apple | braeburn | 178 | 7.1 | 7.8 | 0.92 |
| 1 | apple | braeburn | 172 | 7.4 | 7.0 | 0.89 |
| 1 | apple | braeburn | 166 | 6.9 | 7.3 | 0.93 |
| 1 | apple | braeburn | 172 | 7.1 | 7.6 | 0.92 |
| 1 | apple | braeburn | 154 | 7.0 | 7.1 | 0.88 |
| 1 | apple | golden_delicious | 164 | 7.3 | 7.7 | 0.70 |
| 1 | apple | golden_delicious | 152 | 7.6 | 7.3 | 0.69 |
| 1 | apple | golden_delicious | 156 | 7.7 | 7.1 | 0.69 |
| 1 | apple | golden_delicious | 156 | 7.6 | 7.5 | 0.67 |
| 1 | apple | golden_delicious | 168 | 7.5 | 7.6 | 0.73 |
| 1 | apple | cripps_pink | 162 | 7.5 | 7.1 | 0.83 |
| 1 | apple | cripps_pink | 162 | 7.4 | 7.2 | 0.85 |
| 1 | apple | cripps_pink | 160 | 7.5 | 7.5 | 0.86 |
| 1 | apple | cripps_pink | 156 | 7.4 | 7.4 | 0.84 |
| 1 | apple | cripps_pink | 140 | 7.3 | 7.1 | 0.87 |
| 1 | apple | cripps_pink | 170 | 7.6 | 7.9 | 0.88 |
| 3 | orange | spanish_jumbo | 342 | 9.0 | 9.4 | 0.75 |
| 3 | orange | spanish_jumbo | 356 | 9.2 | 9.2 | 0.75 |
| 3 | orange | spanish_jumbo | 362 | 9.6 | 9.2 | 0.74 |
| 3 | orange | selected_seconds | 204 | 7.5 | 9.2 | 0.77 |
| 3 | orange | selected_seconds | 140 | 6.7 | 7.1 | 0.72 |
| 3 | orange | selected_seconds | 160 | 7.0 | 7.4 | 0.81 |
| 3 | orange | selected_seconds | 158 | 7.1 | 7.5 | 0.79 |
| 3 | orange | selected_seconds | 210 | 7.8 | 8.0 | 0.82 |
| 3 | orange | selected_seconds | 164 | 7.2 | 7.0 | 0.80 |
| 3 | orange | turkey_navel | 190 | 7.5 | 8.1 | 0.74 |
| 3 | orange | turkey_navel | 142 | 7.6 | 7.8 | 0.75 |
| 3 | orange | turkey_navel | 150 | 7.1 | 7.9 | 0.75 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | orange | turkey_navel | 180 | 7.6 | 8.2 | 0.79 |
| 3 | orange | turkey_navel | 154 | 7.2 | 7.2 | 0.82 |
| 4 | lemon | spanish_belsan | 194 | 7.2 | 10.3 | 0.70 |
| 4 | lemon | spanish_belsan | 200 | 7.3 | 10.5 | 0.72 |
| 4 | lemon | spanish_belsan | 186 | 7.2 | 9.2 | 0.72 |
| 4 | lemon | spanish_belsan | 216 | 7.3 | 10.2 | 0.71 |
| 4 | lemon | spanish_belsan | 196 | 7.3 | 9.7 | 0.72 |
| 4 | lemon | spanish_belsan | 174 | 7.3 | 10.1 | 0.72 |
| 4 | lemon | unknown 132 | 5.8 | 8.7 | 0.73 | |
| 4 | lemon | unknown 130 | 6.0 | 8.2 | 0.71 | |
| 4 | lemon | unknown 116 | 6.0 | 7.5 | 0.72 | |
| 4 | lemon | unknown 118 | 5.9 | 8.0 | 0.72 | |
| 4 | lemon | unknown 120 | 6.0 | 8.4 | 0.74 | |
| 4 | lemon | unknown 116 | 6.1 | 8.5 | 0.71 | |
| 4 | lemon | unknown 116 | 6.3 | 7.7 | 0.72 | |
| 4 | lemon | unknown 116 | 5.9 | 8.1 | 0.73 | |
| 4 | lemon | unknown 152 | 6.5 | 8.5 | 0.72 | |
| 4 | lemon | unknown 118 | 6.1 | 8.1 | 0.70 | |

## 5.2 PYTHON SOURCE CODES

```
import pandas as pd
fruits=pd.read_table('fruit_data_with_colors.txt')
fruits.head()
print(fruits.shape)
print(fruits.keys())
print(fruits['fruit_name'].unique())
print(fruits.groupby('fruit_name').size())
import seaborn as sns
import matplotlib.pyplot as plt
fig,ax=plt.subplots()
fig.set_size_inches(4,4)
sns.countplot(fruits['fruit_name'],label="Count")
plt.show()
print(fruits.describe())
print(fruits.fruit_label)
fruits.drop('fruit_label', axis=1).plot(kind='box', subplots=True, layout=(2,2), sharex=False,
sharey=False, figsize=(9,9),
title='Box Plot for each input variable')
plt.savefig('fruits_box')
plt.show()
feature_names = ['mass', 'width', 'height', 'color_score']
X = fruits[feature_names]
y = fruits['fruit_label']
print(X.head())
print(y.head())
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
```

```
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

print(X_train.shape)
print(y_train.shape)

from sklearn.tree import DecisionTreeClassifier
clf = DecisionTreeClassifier().fit(X_train, y_train)
print('Accuracy of Decision Tree classifier on training set: {:.2f}'
     .format(clf.score(X_train, y_train)))
print('Accuracy of Decision Tree classifier on test set: {:.2f}'
     .format(clf.score(X_test, y_test)))

from sklearn.naive_bayes import GaussianNB
nb= GaussianNB()
nb.fit(X_train, y_train)
print('Accuracy of Naive Bayesian classifier on training set: {:.2f}'
     .format(nb.score(X_train, y_train)))
print('Accuracy of Naive Bayesian classifier on test set: {:.2f}'
     .format(nb.score(X_test, y_test)))

from sklearn.svm import SVC
svm = SVC()
svm.fit(X_train, y_train)
print('Accuracy of SVM classifier on training set: {:.2f}'
     .format(svm.score(X_train, y_train)))
print('Accuracy of SVM classifier on test set: {:.2f}'
     .format(svm.score(X_test, y_test)))
```

## 5.3 SCREENSHOTS

| | fruit_label | fruit_name | fruit_subtype | mass | width | height | color_score |
|---|---|---|---|---|---|---|---|
| 0 | 1 | apple | granny_smith | 192 | 8.4 | 7.3 | 0.55 |
| 1 | 1 | apple | granny_smith | 180 | 8.0 | 6.8 | 0.59 |
| 2 | 1 | apple | granny_smith | 176 | 7.4 | 7.2 | 0.60 |
| 3 | 2 | mandarin | mandarin | 86 | 6.2 | 4.7 | 0.80 |
| 4 | 2 | mandarin | mandarin | 84 | 6.0 | 4.6 | 0.79 |

Fig 5.3.1 Data rows of head ()

| | fruit_label | mass | width | height | color_score |
|---|---|---|---|---|---|
| count | 59.000000 | 59.000000 | 59.000000 | 59.000000 | 59.000000 |
| mean | 2.542373 | 163.118644 | 7.105085 | 7.693220 | 0.762881 |
| std | 1.208048 | 55.018832 | 0.816938 | 1.361017 | 0.076857 |
| min | 1.000000 | 76.000000 | 5.800000 | 4.000000 | 0.550000 |
| 25% | 1.000000 | 140.000000 | 6.600000 | 7.200000 | 0.720000 |
| 50% | 3.000000 | 158.000000 | 7.200000 | 7.600000 | 0.750000 |
| 75% | 4.000000 | 177.000000 | 7.500000 | 8.200000 | 0.810000 |
| max | 4.000000 | 362.000000 | 9.600000 | 10.500000 | 0.930000 |

Fig 5.3.2 Describe () method

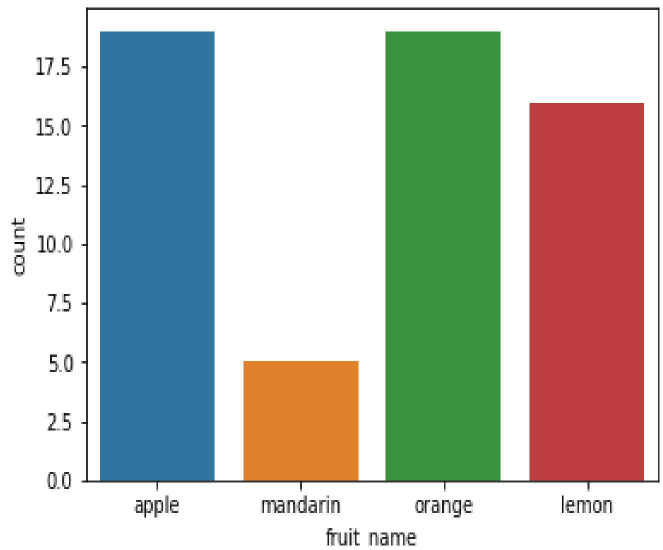


Fig 5.3.3 Plot diagram

```
(44, 4)
(44,)
Accuracy of Decision Tree classifier on training set: 1.00
Accuracy of Decision Tree classifier on test set: 0.73
Accuracy of Naive Bayesian classifier on training set: 0.86
Accuracy of Naive Bayesian classifier on test set: 0.67
Accuracy of SVM classifier on training set: 0.61
Accuracy of SVM classifier on test set: 0.33
```

Fig 5.3.4 Predicted output

# Chapter 6

# CONCLUSION

Other fruit recognition techniques that were introduced earlier, rarely considered the three characteristics to properly classify the fruit. From the experimental result it has been observed that extraction of color, texture and shape features for fruit recognition significantly improve the efficiency of the algorithm. Thus in this work an attempt has been taken to optimize the recognition technique accounting the above mentioned visible features. Texture property extracted from a Log Gabor filtered image is infrequently and hardly used for fruit recognition, which has been considered in this work. A randomly chosen data set of fruit is applied to the Artificial Neural Network classifier for testing. According to the result, most of the cases it has been accomplished without any confusion. This work can be enhanced through the inclusion of some new features so that more accuracy can be achieved. The main constraint of the work is single fruit image with white background has been considered for training as well as for testing. In the future work sincere attempt may be taken to overcome the constraint.

# CHAPTER 7

# REFERENCES

[1] Ryszard S. Choras, “Image Feature Extraction Techniques and their applications for CBIR and Biometrices Systems”, International Journal of Biology and Biomedical Engineering, Issue 1 vol. 1, 2007.

[2] L. Yang, J. Dickinson, Q. M. J. Wu, S. Lang, “A Fruit Recognition Method for Automatic Harvesting”, http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp? punumber=4430702, 4-6 Dec. 2007, Pg. 152 – 157.

[3] Hetal N. Patel, Dr. R. K. Jain, Dr. M. V. Joshi, “Fruit Detection using Improved Multiple Features based Algorithm”, International Journal of computer Applications (0975-8887), Volume 13- No.2, January 2011.

[4] P. Mohanaiah, P. Sathyanarayana, L. GuruKumar, “Image Texture Feature Extraction Using GLCM Approach”, International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013 1 ISSN 2250-3153.

[5] Costa, A. F., G. E. Humpire-Mamani, A. J. M. Traina. 2012. "An Efficient Algorithm for Fractal Analysis of Textures." In SIBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images), 39-46, Ouro Preto, Brazil.

[6] Kamarul Hawari Ghazali, Mohd Fais Mansor, Mohd. Marzuki Mustafa, Aini Hussain, “Feature Extraction Technique using Discrete Wavelet Transform for Image Classification”, The 5th Student Conference on Research and Development – SCOReD 2007 11-12 December 2007, Malaysia.

[7] Dierk Schleicher, “Hausdorff Dimension, Its Properties, and Its Surprises”, THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 114 June–July 2007