

Original Research Article

Classification of Soya Beans based Image Processing Techniques and Artificial Neural Network

Abstract

The primary purpose of using technology in agriculture include increase in the quality and quantity of crops produced, minimising cost of farming, providing suggestions for prompt action among others. Traditionally, to know the state of soya beans, farmers rely in observing change in color of soya beans leaves from one color to another so as to provide appropriate action to soya beans, this process cannot be fully reliable as color is subjective to human naked eyes, and failure to act when there is changes in the state of the soya beans especially when affected by diseases can reduce the expected yield to harvest. The research aimed at classifying soya beans into various categories such as healthy, unhealthy/disease, ripe not ready for harvest and ripe ready for harvest so that prompt action can be taken. The research work has employed the use of color and texture features of leaves through image processing techniques in the pre-processing phase and artificial neural network for the classification with the aid MATLAB. An accuracy of 95.7% was obtained in the classification of the various categories of soya beans leaves.

Keywords: Image Processing, Artificial Neural Network, Soya beans

Introduction

Advancement in technology has provided useful improvements to our daily activities, of which agriculture is not an exception. The use of technology in agriculture has

24 paved ways for new farming techniques across the globe and the benefits cannot be
25 overemphasized. These benefits include increase in the quality and quantity of crops
26 produced, minimising cost of farming, providing suggestions for prompt action among
27 others. Fakhri et al. (2012) stated that one of the techniques used in improving
28 agricultural activities is *image processing*. Rafael & Richards (2008) noted that image
29 processing does not only deal with taking images as its input and producing them as
30 output, but also incorporates the processes that extract image features/attributes used
31 for object recognition. Thus, digital image processing implies the use of computer to
32 process images. The use of image processing is in wide range; from image analysis to
33 computer vision. Image processing systems have been used in agricultural activities,
34 which include diseases identification, weed detection, fruit grading etc. The
35 application of image processing to agriculture involves the extraction of features to
36 perform the needed task. Artificial Neural Network (ANN) is a computational model
37 of the brain that mimics the nervous system. ANN can be supervised learning
38 (requiring training and the output are known) or unsupervised learning (where output
39 is not known and does not require training). Soya bean is cash crop, source of protein
40 and among the major industrial and food crops grown in every continent because of its
41 nutritive and economic values (Dugje et al., 2009). Soya bean exhibit different states
42 which may require different actions. For instance, when it's healthy, the soya bean
43 leaves are purely green in color. When it's affected by disease, the leave possesses a
44 number of colors (such as black, brown, white spot) mix with green. When the soya
45 bean is ripe but not ready for harvest, the color of the leaves is purely yellow. And
46 when soya bean is ripe, the color of the leaves and seed pods are purely brown. The

47 description of soya bean states necessitates a means of classifying soya bean using
48 color and texture of the leaves so that prompt action can be taken. The remaining
49 section discusses the overview of related work, the methodology, result & discussion,
50 conclusion & future work and references.

51 **Overview of Related Work**

52 To detect the ripeness/grade the quality of fruit and to detect disease that affect plants
53 using either the leaves of the fruits, a number of approaches have been proposed using
54 different techniques for the classification as discussed in the following.

55 Sweths & Arshvant (2016) proposed a novel approach to detect the severity of a
56 disease using three parameters (Infection per Region IPR, Disease Severity Index DSI
57 and Disease Level Parameter DLP) and classify the disease of soya beans using the
58 leaves of the plant. The approach involves image acquisition, pre-processing
59 (Enhancement and Segmentation) and Classification based on neural networks.
60 Textures and color features were extracted which serves as an input and the output are
61 the different soya beans diseases. An accuracy of more than 90% was achieved in
62 classifying soya beans foliar diseases. Similarly, Shahzed (2015) proposed a method
63 for classifying fruits as good or infected with apple as a case study using artificial
64 neural networks. The approach used the color features by extracting the red, green and
65 blue components and convert them to grey scale (i.e. Hue, saturation and Intensity,
66 HSI) images. The approach achieved an accuracy of 90% in classifying the fruit. This
67 approach is best suited for fruits with a dominant color such as apple, Mango, etc. but
68 weak in classifying fruits with mixed color such as banana.

69 Furthermore, Jasmeet & Er (2016) proposed a method of detecting plant diseases of
70 grape and wheat plant using support vector machine. Features extracted were color,
71 textures and shape. To enhance the accuracy of detection, three algorithms were
72 applied in the proposed approach namely; back propagation algorithm for training the
73 system, principal component analysis was used to reduce the data/features and
74 singular value decomposition was used to enhance the accuracy of detection. A good
75 recognition result was obtained but with low accuracy. In addition, Priyanka et al.
76 (2014) proposed a method that detects and classifies plant diseases on real time. The
77 approach consists of two parts; the hardware (robotic which captures the images and
78 gives out output) and software where the image analysis and classification were
79 conducted with the aid of support vector machine. Mango leaves were used to
80 ascertain the effectiveness of the developed method and a minimum of 71% accuracy
81 was obtained. Moreover, Jiangsheng et al. (2015) proposed a novel method for
82 extracting the portion of affected regions of soya beans leaf based on modified salient
83 regions. The method uses low level features of luminance and color features together
84 with multi-scale analysis to determine saliency maps in images and applied K-means
85 clustering for segmentation. The results obtained showed that the method had
86 successfully detected the diseases on soya beans leaves in the presence of excessive
87 background. Additionally, Shiv et al. (2013) proposed a method of defect detection
88 using K-means clustering on fruits (Using apple as a case study). The captured image
89 was converted to L^*a^*b color components and color a^* and b^* were segmented into
90 various segment. The approach successfully detected defects from the input images.
91 The choice on the number of clusters depends on the severity of the defects as

indicated in the approach; the larger the area of defects, the smaller the number of clusters needed to detect defects. Babawuro (2013) implemented a novel method of classifying high resolution satellite images using K-means clustering. The technique involves segmenting the image and classifying the image based on objects methods which reduces the problem of pixel-by-pixel method of classification. The technique has yielded a better clarity and higher accuracy of 88.89% by classifying the satellite image into various categories (such as farmland and bare land).

Methodology

The methodology adopted in the research work consists of four stages as shown in figure 1.

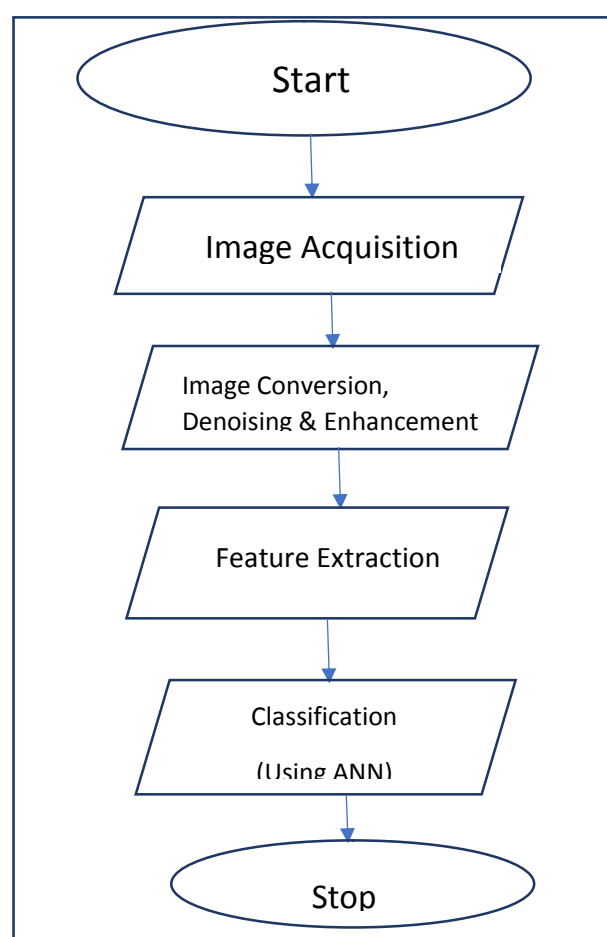


Figure 1: Flowchart of the proposed Methodology

A. Image Acquisition: The images of soya beans leave of various categories ranging from ripe leaf (but not ready for harvest, usually yellow in color), unripe (purely green), unhealthy leaf (that is diseased leaf usually in mix color of green, brown and grey) to fully ripe leaf (purely brown in color) were captured using a digital camera (Sony with 20.1 mega pixel and 5x optical zoom). The images were captured from a farm located 2 kilometres away from Kaduna, Kaduna State at around 11.00am. The images were captured under uncontrolled environment; this is to obtain their natural color without introducing any variation in intensity/brightness to the images. For each category of leaf as mentioned above, fifty samples of each category were captured. The captured leaves are 1920x1080 pixel (width and height) and of size ranging from 600 to 764 kilobytes but were later rescale using Microsoft paint to 304x237 and rename in ascending order for easier referencing during processing.



Figure 2 (a) healthy leaf of soya bean



Figure 2 (b) ripe leaf of soya bean



Figure 2 (c) ripe not ready leaf of soya bean



Figure 2 (d) diseased leaf of soya bean

The figures above are various categories of soya beans leaves, figure 2(a) is healthy leaf, figure 2(b) is a ripe leaf, figure 2(c) ripe but not ready for harvest and figure 2(d) is one type of disease/unhealthy leaf.

B. Image Conversion, De-noising and Enhancement: The captured images in their red, green, blue (RGB) color model were converted to hue, saturation, value (HSI) color model because according to Rafael & Richard (2008), HSI model is an ideal tool used in image processing and when humans view a color, they describe it in its color form, thus, the model becomes an intuitive form of description for humans. Noise refers to variation in intensity or brightness in an image (Rafeal et al. 2008). To de-noise and enhance the images, a 3x3 averaging filter (also known as a sub image, mask, kernel or window) was applied to intensity component of HSI model so as to preserve the original color of the image without distorting the hue and saturation component. This is to ensure that desirable features such as image edges are not blurred. The algorithm used is described below:

Algorithm: Image de-noising and Enhancement

Input: RGB image

Output: De-noised and enhanced image

Begin/Start

1.0 Read the RGB image

2.0 Convert the RGB image to image in the HSI color model

3.0 Extract the components (HSI) of the converted image

168 4.0 Apply an averaging filter to the value component of step 3.0
 169 5.0 Combine the hue & saturation component with the filtered value
 170 component of step 4.0
 171 6.0 Convert the image in step 5.0 to RGB image
 172 7.0 Store the result
 173 End/Stop

174 **C. Feature Extraction:** According Garima (2014), feature extraction involves the
 175 manipulation and transformation of input data into a set of features that
 176 uniquely identifies or represents an object (image). These features are color,
 177 shape and texture. Color and texture features were extracted because according
 178 to Pratap et al. (2014) more accuracy is obtained when color and texture are
 179 combine to classify fruits or detect diseases. To extract color features, the mean
 180 and standard deviation of each HSI component were extracted for easier
 181 analysis. The algorithm used for color extraction is shown below

182 *Algorithm: Color extraction*
 183 *Input: RGB image*
 184 *Output: Color features (HSI Components)*
 185 *Begin*
 186 1. Read the RGB image
 187 2. Convert the RGB image to HSV image
 188 3. Extract the hue, saturation and intensity component from RGB image.
 189 4. For each of the components, computer the mean and standard deviation
 190 5. Store the result of step 4.
 191 *End.*

192 As for the texture features, gray level co-occurrence matrix (GLCM) was used
 193 to extract the statistical parameters such as energy, contrast, homogeneity and
 194 correlation. The algorithm texture feature extraction is detailed below.

195 *Algorithm: Texture features extraction*
 196 *Input: RGB image*
 197 *Output: Texture Features*
 198 *Start*
 199 1.0Read the enhanced image and convert it to grey
 200 2.0Calculate the GLCM from four directions using the image obtained in step 1.0
 201 3.0Extract the statistical parameters using the value of GLCM obtained in step
 202 2.0
 203 4.0Save the values of those parameters

Stop

D. Image Classification: As stated by Le Hoang et al. (2012), the principal objective of image classification is to predict the class of input image based on its features. To predict the category of input features, there are a number of approaches that can be used such as supervised training (requiring training and the output are known) such as Support Vector Machine (SVM), Artificial Neural Network (ANN), adaptive boost etc. and unsupervised training (doesn't require training) such K-Nearest Neighbor. In this research, ANN was adopted because, it has the ability to classify features it has not train before and it can predict an independent variable into more than two possible states.

Result and Discussion

To test the accuracy of the designed neural network in classifying leaves of soya bean into various categories, first, samples of soya bean used during the training stage were used to test the system, and secondly, those that were not used in the training stage to test the system and the result obtained is shown in the table.

S/N	Category of Leave	Correctly Classified
1	Diseased Leave	Yes
2	Diseased Leave	Yes
3	Diseased Leave	No
4	Diseased Leave	Yes
5	Diseased Leave	Yes
6	Healthy Leave	Yes
7	Healthy Leave	Yes
8	Healthy Leave	Yes

9	Healthy Leave	Yes
10	Healthy Leave	Yes
11	Ripe not ready	Yes
12	Ripe not ready	Yes
13	Ripe not ready	Yes
14	Ripe not ready	Yes
15	Ripe not ready	Yes
16	Ripe Ready	Yes
17	Ripe Ready	Yes
18	Ripe Ready	Yes
19	Ripe Ready	Yes
20	Ripe Ready	Yes

220

Table 1: Testing with Untrained Samples

S/N	Category of Leave	Correctly Classified
1	Diseased Leave	Yes
2	Diseased Leave	Yes
3	Diseased Leave	Yes
4	Diseased Leave	Yes
5	Diseased Leave	Yes
6	Healthy Leave	Yes
7	Healthy Leave	Yes
8	Healthy Leave	Yes
9	Healthy Leave	Yes
10	Healthy Leave	No
11	Ripe not Ready	Yes
12	Ripe not Ready	Yes
13	Ripe not Ready	Yes
14	Ripe not Ready	Yes
15	Ripe not Ready	Yes
16	Ripe Ready	Yes

17	Ripe Ready	Yes
18	Ripe Ready	Yes
19	Ripe Ready	Yes
20	Ripe Ready	Yes

221

222 Table 2: Testing with trained samplese.

223 The result can be further view when we plot the percentage accuracy of the
224 classification as shown in the below figure.

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

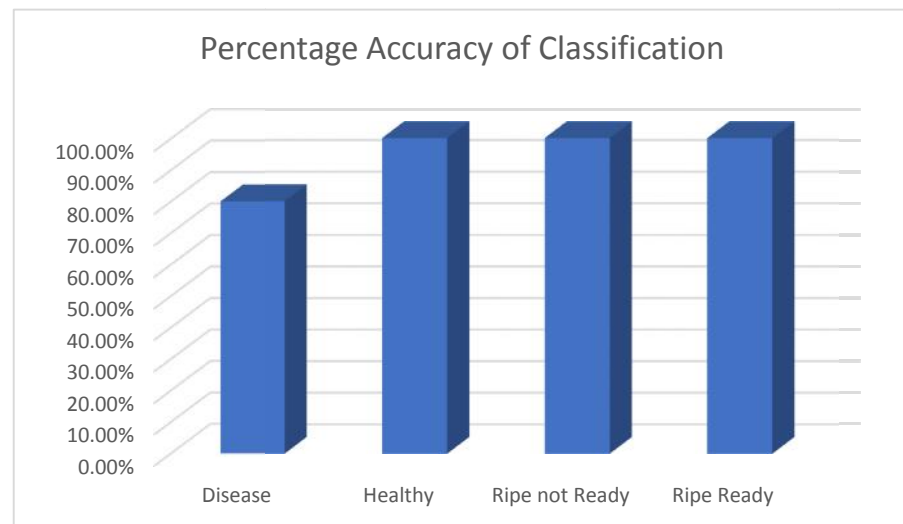


Figure 3: Percentage accuracy of classification with untrained samples

246 Conclusion and Future Work

247 The proposed approach for classifying various categories of soya bean leaves has been
248 presented. The approach uses Artificial Neural Network for the classification process
249 and image processing techniques was used to denoise, enhance and extract the
250 features (color and texture) from captured images. Based on the experiment carried
251 out, an accuracy of 95.7% was obtained. However, despite the accuracy obtained in

classifying the various categories of soya bean using the mentioned features, more ways are needed that will further differentiate between the categories especially between healthy/unripe and disease leave, and between ripe ready for harvest and not ripe ready.

References

- A Fakhri, A. N., M, N. A., & A, R. M. (2012). A study of image processing in Agriculture Application Under High Performance Computing Environment. *Journal of Computer Science & Telecommunication*, 3(8), 16.
- Babawuro, U. (2013). Satellite Imagery Land Cover Classification using K-Means Clustering. *Elixir International Journal, Computer Science and Engineering*, 63, 18671-18675.
- Dugje, I., Omoigui, L., F, E., & R, B. (2009, May 13). *International Institute of Tropical Agriculture (IITA)*. Retrieved October 27, 2016, from http://www.iita.org/c/document_library/get_file?p_l_id=25417&folderId=25529&name=DLFE-192.pdf
- Garima, T. (2014). Review on Color and Texture Feature Extraction Techniques. *Internal journal of Enhanced Research in Management and Computer Applications*, 3(5), 77-81.
- Jasmeet, K., & Er, R. K. (2016). Plant Disease Detection Using SVM Algorithm and Neural Network. *International Journal of Innovative Research in Computer and Communication Engineering*, 4(6).
- Jiangsheng, G., Li, H., Qing, Z., & Xiaolan, B. (2015). A new method of Soyabeans leaf Disease Detection based on modified Regions. *International Journal of Multimedia and Ubiquitous Engineering*, 10(6), 45-52.
- Le Hoang, T., Tran, S. H., & Nguyen, T. T. (2012). Image Classification using Support Vector Machine and Artificial Neural Networks. *International Journal of Information Technology and Computer Science*, 5, 32-38.
- Pratap, B., Navneet, A., Sunil, J., & Suruti, G. (2014). Development of ANN Based Efficient Fruit Recognition Technique. *Global Journal of Computer Science and Technology: Software and Data Engineering*, 14(5), 5-18.
- Priyanka, P., Kanwal, R., Sanam, S., & S, T. K. (2014). Machine vision Guided Sysytem for Classification and Detection of Plant Diseases Using Support Vector Machine. *International Journal of Computer Communication and Computer Engineering.*, 5(4).
- Rafael, C. G., & Richard, E. W. (2008). *Digital Image Processing* (2nd ed.). New Delhi: Prentice-Hall of India Private Limited.
- Shahzed Preet Singh, S. B. (March 2015). Classification of Apples Using Neural Network. *International Journal of Science, Technology and Mangement*, 4(1).

287 Shiv, R. D., Pushakar, D., Nishant, S., & jay, P. (2013). Infected Fruit Part Detection Using K-Means
 288 Clustering Segmentation Technique. *International Journal of Artificial Intelligence and*
 289 *Interactive Multimedia*, 2(2).

290 Sweths , V. S., & Arshvant, R. (2016). Classification and Severity Detection of Soyabeans Foliar
 291 Diseases Using Artificial Neural Network. *International Journal of Advanced Research in*
 292 *Management, Architecture, Technology and Engineering*, 2(7).

293

294