

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

Abstract

The benefits of using technology in agriculture cannot be overemphasised because of its impact that results in an increase in the quality and quantity of crops produced, minimising cost of farming, and providing suggestions for prompt action among others. Traditionally, to know the state of soya beans, farmers rely on observation to note the change in colour of the leaves so as to provide appropriate action to the crop. This process cannot be fully reliable as colour is subjective to human impression; and failure to act when there are changes in the state of the soya beans especially when affected by diseases can reduce the expected yield. The goal of this study is to classify soya beans leaves 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 work has employed the use of colour 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% is obtained in the classification of the various categories of soya beans leaves.

Keywords: Image Processing, Artificial Neural Network, Soya beans, MATLAB, Machine learning, object recognition.

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 paved the way for new farming techniques across the globe and the benefits cannot be overemphasized. These benefits include increase in the quality and quantity of crops produced, minimising cost of farming, providing suggestions for prompt action among others. Fakhri et al. (2012) state that one of the techniques used in improving agricultural activities is *image processing*. Rafael & Richards (2008) note that image processing does not only deal with taking images as its input and producing them as output, but also incorporates the processes that extract image features/attributes used for object recognition. Thus, digital image processing implies the use of computer to process images. The use of image processing is in wide range; from image analysis to computer vision. Image processing systems have been used in agricultural activities, which include diseases identification, weed detection, fruit grading and so on. The application of image processing to agriculture involves the extraction of features to perform the needed task. Artificial Neural Network (ANN) is a computational model of the brain that mimics the nervous system. ANN can be supervised learning (requiring training and the output are known) or unsupervised learning (where output is not known and does not require training). Soya beans is cash crop, source of protein and among the major industrial and food crops grown in every continent because of its nutritive and economic values (Dugje et al., 2009). Soya beans exhibit different states which may require different actions. For instance, when it is healthy, its leaves are purely green, but when it is affected by disease, the leave possesses a number of colours (such as black, brown, white spot) mix with green. However, when soya bean is ripe but not ready for harvest, the colour of the leaves is purely yellow. When soya bean is ripe and ready for harvest, the colour of the leaves and seed pods are purely brown. The description of the various states of soya bean necessitates a means of classifying the crop using colour and texture of the leaves so that prompt action can be taken. The remaining section discusses the overview of related works, the methodology, result and discussion, conclusion and future work and references.

54 **Review of Related Works**

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

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

72 Furthermore, Jasmeet & Er (2016) propose a method of detecting plant diseases of
73 grape and wheat using support vector machine. Features extracted are colour, texture
74 and shape. To enhance the accuracy of detection, three algorithms are applied in the
75 proposed approach namely; back propagation algorithm for training the system,
76 principal component analysis is used to reduce the data/features and singular value

77 decomposition was used to enhance the accuracy of detection. A good recognition
78 result is obtained but with low accuracy. In addition, Priyanka et al. (2014) proposed a
79 method that detects and classifies plant diseases on real time. The approach consists of
80 two parts; the hardware (robotic which captures the images and gives out output) and
81 software where the image analysis and classification are conducted with the aid of
82 support vector machine. Mango leaves are used to ascertain the effectiveness of the
83 developed method and a minimum of 71% accuracy is obtained. Moreover,
84 Jiangsheng et al. (2015) proposed a novel method for extracting the portion of affected
85 regions of soya beans leaf based on modified salient regions. The method uses low
86 level features of luminance and colour features together with multi-scale analysis to
87 determine saliency maps in images and applied K-means clustering for segmentation.
88 The results obtained shows that the method has successfully detected the diseases on
89 soya beans leaves in the presence of excessive background. Additionally, Shiv et al.
90 (2013) propose a method of defect detection using K-means clustering on fruits
91 (Using apple as a case study). The captured image is converted to L*a*b colour
92 components and colour *a and *b are segmented into various segment. The approach
93 successfully detected defects from the input images. The choice on the number of
94 clusters depends on the severity of the defects as indicated in the approach; the larger
95 the area of defects, the smaller the number of clusters needed to detect defects.
96 Babawuro (2013) implements a novel method of classifying high resolution satellite
97 images using K-means clustering. The technique involves segmenting the image and
98 classifying the image based on objects methods which reduces the problem of pixel-
99 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 consists of four stages as shown in figure 1.

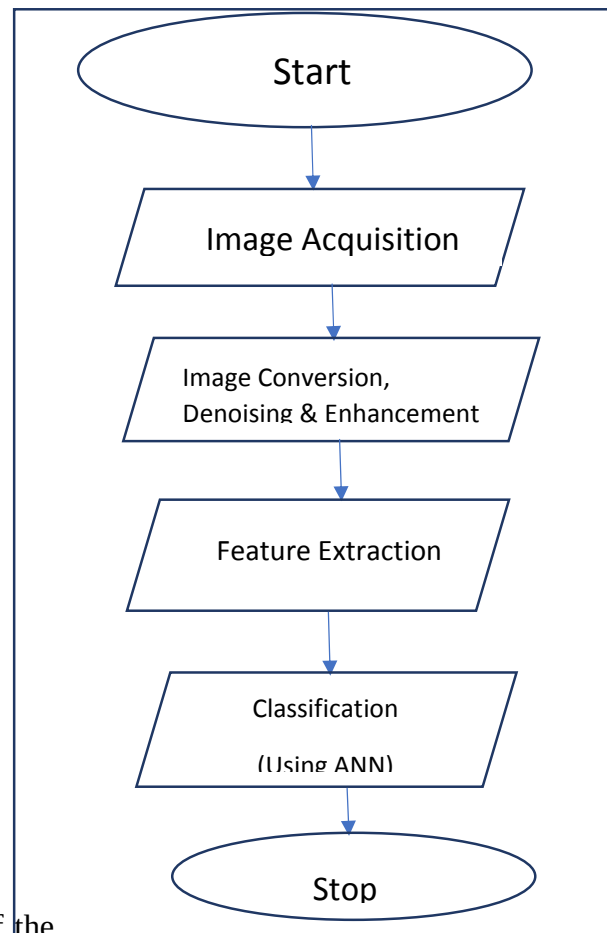


Figure 1: Flowchart of the proposed

Methodology

A. Data Collection: The images of soya beans leaves of various categories ranging from ripe leaf (but not ready for harvest, usually yellow in colour), unripe (purely green), unhealthy leaf (that is diseased leaf usually in mix colour of green, brown and grey) to fully ripe leaf (purely brown in colour) are captured using a digital camera (Sony with 20.1 mega pixel and 5x optical

zoom). The images are captured from a farm located 2 kilometres away from Kaduna, Kaduna State at around 11.00am. The images are captured under uncontrolled environment; this is to obtain their natural colour without introducing any variation in intensity/brightness to the images. For each category of leaf as mentioned above, fifty samples of each category are 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

143

144

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 leave.

B. Image Conversion, De-noising and Enhancement: The captured images in their red, green, blue (RGB) colour model are converted to hue, saturation, value (HSI) colour model because according to Rafael & Richard (2008), HSI model is an ideal tool used in image processing and when humans view a colour, they describe it in its colour 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) is applied to intensity component of HSI model so as to preserve the original colour 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 colour model

3.0 Extract the components (HSI) of the converted image

4.0 Apply an averaging filter to the value component of step 3.0

5.0 Combine the hue & saturation component with the filtered value component of step 4.0

6.0 Convert the image in step 5.0 to RGB image

7.0 Store the result

End/Stop

C. Feature Extraction: According to Garima (2014), feature extraction involves the manipulation and transformation of input data into a set of features that uniquely identifies or represents an object (image). These features are colour, shape and texture. Colour and texture features are extracted because according to Pratap et al. (2014) more accuracy is obtained when colour and texture are

combined to classify fruits or detect diseases. To extract colour features, the mean and standard deviation of each HSI component are extracted for easier analysis. The algorithm used for colour extraction is shown below:

Algorithm: Colour extraction

Input: RGB image

Output: Colour features (HSI Components)

Begin

1. *Read the RGB image*

2. *Convert the RGB image to HSV image*

3. *Extract the hue, saturation and intensity component from RGB image.*

4. *For each of the components, computer the mean and standard deviation*

5. *Store the result of step 4.*

End.

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

Algorithm: Texture features extraction

Input: RGB image

Output: Texture Features

Start

1.0*Read the enhanced image and convert it to grey*

2.0*Calculate the GLCM from four directions using the image obtained in step 1.0*

3.0*Extract the statistical parameters using the value of GLCM obtained in step*

2.0

4.0*Save 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 and so on, and unsupervised training (doesn't require training) such K-Nearest Neighbor. In this research, ANN is

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.

Overview of Artificial Neural Network

The structure of ANN according to Kriesel (2009) consists of three (3) layers; the input layer consists of neurons through which information is fed into the network, the hidden layer where information is processed by the neurons and the output layer which consist of one or more neurons for result. The process of training a neural network involves adjusting or turning the weights between neurons so that the output of a network matches the target class. This training process is shown in Figure 3.3 (MathWorks, 2017).

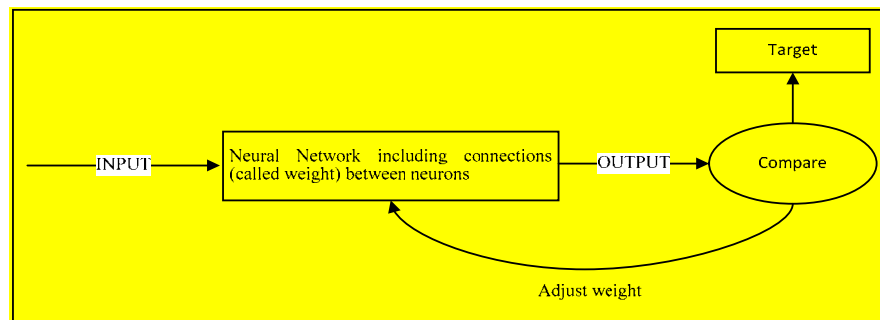


Figure 3.3: Neural network training workflow

MathWorks (2017) stated that artificial neural network can be feedforward neural network (where the flow of information is unidirectional, mostly used in object recognition or classification) or feedback neural network (where the flow of information is bidirectional, mostly used in content addressable memory). The architecture of neural network can be single layer consisting of single hidden layer or it can be multilayer neural network with multiple hidden layers.

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 are

used to test the system, and secondly, those that are not used in the training stage to test the system and the result obtained is shown in the table.

S/N	Samples Not Used for Training		Samples Used for Training	
	Category of Leave	Correctly Classified	Category of Leave	Correctly Classified
1	Diseased Leave	Yes	Diseased Leave	Yes
2	Diseased Leave	Yes	Diseased Leave	Yes
3	Diseased Leave	No	Diseased Leave	Yes
4	Diseased Leave	Yes	Diseased Leave	Yes
5	Diseased Leave	Yes	Diseased Leave	Yes
6	Healthy Leave	Yes	Healthy Leave	Yes
7	Healthy Leave	Yes	Healthy Leave	Yes
8	Healthy Leave	Yes	Healthy Leave	Yes
9	Healthy Leave	Yes	Healthy Leave	Yes
10	Healthy Leave	Yes	Healthy Leave	No
11	Ripe not ready	Yes	Ripe not Ready	Yes
12	Ripe not ready	Yes	Ripe not Ready	Yes
13	Ripe not ready	Yes	Ripe not Ready	Yes
14	Ripe not ready	Yes	Ripe not Ready	Yes
15	Ripe not ready	Yes	Ripe not Ready	Yes
16	Ripe Ready	Yes	Ripe Ready	Yes
17	Ripe Ready	Yes	Ripe Ready	Yes
18	Ripe Ready	Yes	Ripe Ready	Yes
19	Ripe Ready	Yes	Ripe Ready	Yes
20	Ripe Ready	Yes	Ripe Ready	Yes

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

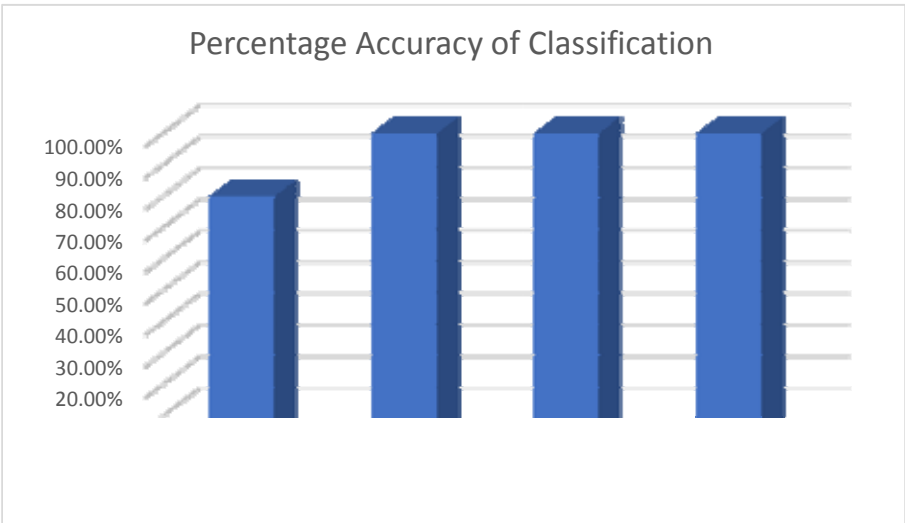


Figure 3: Percentage accuracy of classification with untrained samples

Comparison with other works

This research focuses on classifying the various categories of soya beans leaves so that prompt action can be taken. The research differs from the works of Sweths & Arshvant (2016), Jasmeet et al. (2016), Somwarshi & Humbe (2015) and Priyanka (2014) where the main objective in their work is to detect diseases that affect a plant or various plants, although we used same classifier. Similarly, the research differs from the works of Shahzed (2015), Pratap et al. (2014) and Norasyikin et al. (2012) where the objective is to detect the ripeness or grade the quality of fruits using different categories of fruits. This research presents a another means of preventing loss of agricultural produce (especially those crops where their status can be detected through their leaves) since not only do diseases or quantity of chemicals applied to plants lead to massive loss of crops, other factors such as failure to take appropriate action (such failure to harvest when ripe or disease) lead to massive loss of crops as shown in this research.

Conclusion and Future Work

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

Ethical approval and consent is not applicable

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 Colour 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., & Xiaoan, 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.
- Kriesel, D. (2009, 10 11). *Dkriesel.com*. Retrieved November 22, 2016, from http://www.dkriesel.com/en/science/neural_networks
- 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.
- Norasyikin, F. et al., 2012. Intelligent Colour Vision System for Ripeness Classification. *Sensors (Open Access)*, Volume 12, pp. 14179-14195.
- 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 Sysstem 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.

- 307 Somwarshi, D. & Humbe, V., 2015. Soyabean Rust Severity Measurement Using Image Processing
308 Technique. *Advances in Computation Research* , 7(1), pp. 265-267.
- 309 Shahzed Preet Singh, S. B. (March 2015). Classification of Apples Using Neural Network. *International*
310 *Journal of Science, Technology and Mangement*, 4(1).
- 311 Shiv, R. D., Pushakar, D., Nishant, S., & jay, P. (2013). Infected Fruit Part Detection Using K-Means
312 Clustering Segmentation Technique. *International Journal of Artificial Intelligence and*
313 *Interactive Multimedia*, 2(2).
- 314 Sweths , V. S., & Arshvant, R. (2016). Classification and Severity Detection of Soyabeans Foliar
315 Diseases Using Artificial Neural Network. *International Journal of Advanced Research in*
316 *Management, Architecture, Technology and Engineering*, 2(7).