



QUALITY ANALYSIS AND GRADING OF SOYBEAN USING MACHINE LEARNING

Mr. Vaibhav S. Yende
M.E. Digital Electronics
PRMIT&R Badnera
Amravati, India
vaibhavayende@gmail.com

Prof. S.V. Pattalwar
Associate Professor
PRMIT&R Badnera
Amravati, India
svpattalwar@mitra.ac.in

Dr. S.M. Deshmukh
Head of Department
PRMIT&R Badnera
Amravati, India
smdeshmukh@mitra.ac.in

Abstract-

The use of good quality seed is very important for the better production of a good quality crop and is essential for export in markets. Quality control process is very important in food industry, based on quality of food products are classified and graded into different grades. Soybean is primarily graded based on its grain shape, colour, size and texture. This paper attempts to automate the grading process by using image processing and machine learning techniques. Soybean's grade is affected by damaging, decolourization, infection by insects, immaturity and shrivels, splitting, breaking, cracks, inorganic and organic foreign matter present in the sample. One of the objectives of this paper is to study the effect of these parameters on shape, colour, size and texture of the soybean image. In the present soybean-handling scenario, type and quality are identified manually by visual inspection which is tedious and not accurate. There is need for the growth of accurate, fast and objective system for quality determination of food grains. This paper is automate the system for grading of soybean by extracting morphological features as attributes for classification using image processing techniques and artificial neural network. The classification is done on the basis of features extracted from the segmented images. Simple averaging is used for the combination of results from each classifier. The data is divided into the training and testing data. The data classified as training data is then used for training the neural network. Remaining data is used for the testing purpose. The training and testing dataset involves both normal as well as abnormal cases. Then classification is done into normal and abnormal. It provides better results compared to single neural network that has accuracy of about 90%. This method requires minimum time and it is low in cost.

Keywords: ANN, Grading, Image Processing, Machine Learning, Soybean, Morphological, Seed features

INTRODUCTION

Soybean contributes significantly to the Indian edible oil seed. Currently soybean contributes 43 % to the total oilseeds and 25% to the total oil production in the country. Currently, India's rank is fourth in respect to production of soybean in the world. The crop helps earn valuable foreign exchange (Rs. 62000 millions in 2013-14) by way of soya meal exports. Soybean has largely been responsible in increasing farmer's economic status in many pockets of the country. It usually earns higher income to the farmers owing to the huge export market for soybean de-oiled cake. In contrast production of Soybean in India has increased at a CAGR of 9.60 per cent from 6.87 million tonnes in 2004-05 to 15.68 million tonnes in 2012-13. On the other way Soybean meal consumption has also increased at a CAGR of 10.82 percent over the last eleven years from 1365 thousand million tonnes in 2004-05 to 4225 thousand million tonnes in 2014- 15. Therefore to increasing demand it is imperative to increase the productivity level of Soybean in the country. Production of soybean in India is ruled by Maharashtra and Madhya Pradesh which contribute 89 per cent of the total production. Rajasthan, Andhra Pradesh, Karnataka, Chhattisgarh and Gujarat contribute the remaining 11 per cent production. Because of this global demand of soybean market increasing day by day we need to improved quality of soybean by using image processing and machine learning for better result than that of the manual inspection. In the present

soybean-handling scenario, type and quality are identified manually by visual inspection which is tedious and not accurate. There is need for the growth of accurate, fast and objective system for quality determination of food grains. This paper is automate the system for grading of soybean by extracting morphological features as attributes for classification using image processing techniques and artificial neural network for machine learning. This method requires minimum time and it is low in cost.

LITERATURE REVIEW

Md Abdul Momin, Kazuya Yamamoto, Munenori Miyamoto and Naoshi Kondo[1] proposed the Machine vision based soybean quality evaluation objective of their research was to develop a machine-vision based proof of concept of a grain quality monitoring system, targeted at application on a combine harvester. The machine vision subsystem of the grain quality monitoring concept comprised a digital camera, in combination with back, front and structured lighting. It used a container in which soybean samples were presented with common MOGs such as broken seeds, defect seeds, as well as leaf and pod material. Because of its overall popularity, a Fukuyutaka soybean variety was used for their work, which was harvested in December 2011 in the Mie prefecture, Japan. Soybean samples obtained from a combine harvester tank, and a moisture of content of approximately 11%. The samples were kept in closed plastic bags and stored in a refrigerated chamber at approximately 12 C. From the collected samples, 500 g material samples were chosen at random. The aim was to detect common anomalies from clean undamaged soybeans, also known as dockage fractions, allowing targeted adjustment of combine harvester subsystems to optimize grain quality in the tank. A web camera with back and front lighting was used to evaluate the accuracy of detection of (1) normal (undamaged) beans, (2) split beans, (3) contaminated beans (4) defect beans and (5) stem/pod material. 50 sets of data were collected where a fixed number of these objects were randomly arranged in an imaging cell with a size of 50 50 mm. A digital image processing algorithm was able to distinguish each targeted dockage fraction present in the soybean samples based on front lit and back lit images.

Salome Hema Chitra, S. Suguna and S. Naganandini Sujatha[2] have proposed A Survey on Image Analysis Techniques in Agricultural Product In this work, they implemented a five processing module for seed identification. Seed sample images been taken for the basis of image acquisition and then a seed image is pre-processed by removal of noise and image enhancement technique. edge detection is used for Enhanced image and segmentation was done. From the segmented image extract the features like colour, shape and texture for normal and defected seed which may help to identify the seed by image analysis techniques.

Timothy J. Herrman, Extension State Leader Grain Science and Industry Carl Reed,[3] Extension Specialist, Grain Storage Grain Science and Industry from Kansas State University Agricultural Experiment Station and Cooperative Extension Service have proposed soybean grading procedure on the basis of 1. Classes as yellow soybeans and mixed soybeans. 2. Damaged kernels 3. Foreign material 4. Heat-damaged kernels 5. Soybeans of other colours they Examine the sample for heating, odour, animal filth, castor beans, crotalaria seeds, garlic, glass, insect infestation, purple mottled and stained, smut, stones, unknown foreign substances, and other unusual conditions. Divide out a indicative portion from the sample and determine its moisture content. Determine the test weight per bushel of the sample. When deemed necessary, divide out representative portions and determine the percentage of class, damaged kernels, heat-damaged kernels, foreign material, oil, protein, soybeans of other colours, and splits. All these procedure is done by manual inspection.

D. Wang, M. S. Ram, F. E. Dowell [4] have proposed Classification of damaged soybean seeds using near infrared spectroscopy. They were classify sound and damaged soybean seeds and discriminate among various types of damage using NIR spectroscopy.

Sudhir Gupta, Smitha Patil, Asha Kulkarni and Zoheb Hukkeri[5] proposed the Classification of Soybean seeds by Color Image Analysis. Color features in the RGB (red-green-blue) colorspace are computed. A feed-forward neural network is trained to classify sample soybean seeds. Image features of the soybean seeds were extracted to characterize the physical quality attributes of soybeans. A number of color features were computed and tested. They included the means and standard deviations of R, G, and B (red, green, and blue); the means of H, S, and I (hue, saturation, and intensity); excess red (2R-G-B), excess green (2G-R-B), and excess blue (2B-R-G). The excess colors correspond more closely to the way humans perceive colors than the RGB representation. Algorithms were developed in Windows environment using Matlab 7 programming language to extract color features of individual soybean seeds. From the red, green, and blue color bands of an image, hue, intensity and saturation were calculated.

Irfan S. Ahmad, Research Associate, John F. Reid, Professor, Marvin R. Paulsen, Professor [6], Department of Agricultural Engineering, and James B. Sinclair, Professor, Department of Crop Sciences, University of Illinois

at Urbana-Champaign, proposed Color Classifier for Symptomatic Soybean Seeds Using Image Processing Symptoms associated with fungal damage, viral diseases, and immature soybean (*Glycine max*) seeds were characterized using image processing techniques. A Red (R), Green (G), Blue (B) color feature-based multivariate decision model differentiated between asymptomatic and symptomatic seeds for inspection and grading. The color analysis showed distinct color differences between the asymptomatic and symptomatic seeds. A model comprising of six color features including minimums, averages, and variances for Red, Green and Blue pixel values was developed for describing the seed symptoms. The color analysis showed that color alone did not satisfactorily describe some of the differences between symptoms.

IMPORTANCE OF QUALITY SEED

In the past few years, the soybean crop has become very important for farmers due to its comparatively better market value and its short duration, which ensures farmers quick cash returns. A recent report by the government Seed Testing Laboratory at Nagpur says that the germination rate of seed is about 75%. With some samples showing a germination rate as low as 40 % It is being very important to use quality of seed for farming. By using quality seed can increase your yield between 10 and 20 percent. Using quality of seed we will save money as inputs to the crops and will get higher production.

EXPERIMENTAL SETUP AND ARCHITECTURE

I have carried out different experiments with the framework depicted in Fig.1. With provided solution, Soybean seed samples are taken and put into tray and image is captured under constant illumination conditions. The images are acquired with a color Digital Camera that used to capture images of soybean grain samples. The images are save in Jpeg format and then it use for image processing. Captured image is then sent to processor for further processing. The Morphological features are extracted for quality measurement based on different grain varieties.

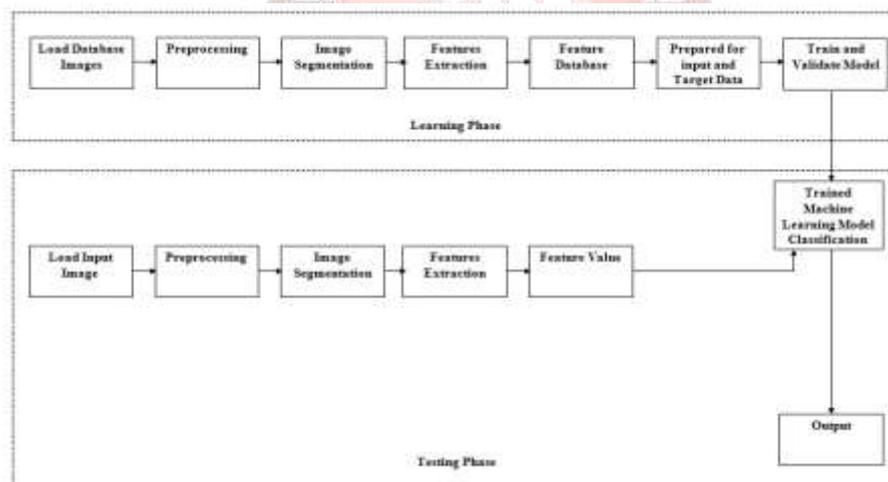


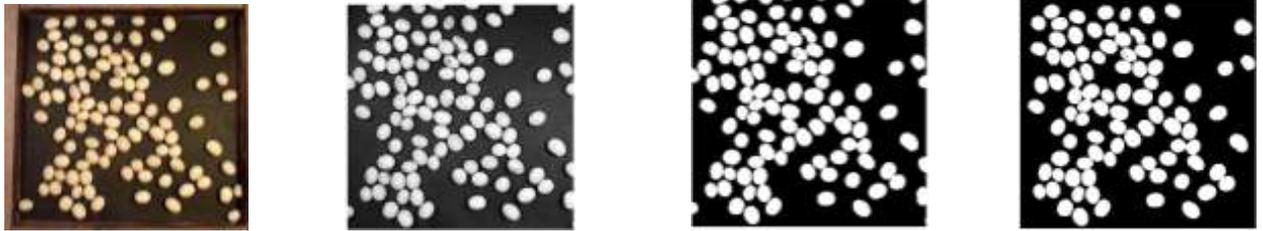
Figure 1:- Block Diagram

The main components of the architecture are as given below:

1. Image Acquisition
2. Image Pre-Processing
3. Image Segmentation
4. Feature Extraction
5. Seeds classification
6. Report generation

1. Image Acquisition: - Image acquisition is the creation of digitally encoded representation of the visual characteristics of an object, such as physical scene or the interior structure of an object. The term is also including the processing, compression, storage, printing and display of such image. Food grain images acquisition is considered as the most critical step of the grain recognition system, as it determines the final grain image quality, which has drastic effects on overall system performance. The images are acquired with a color Digital Camera that used to capture images of soybean grain samples. The images are save in jpeg format and then it use for image processing. I have acquired Soybean grain images from Intex IT-306 WC PC WEBCAM with black background under best illumination scenario. The saved images in Jpeg format are then used for image processing.

2. Image Pre-Processing:- Median filtering is a nonlinear filtering method useful in reducing impulsive or salt and pepper noise. It is also useful in preserving edges in an image and also reducing the random noise. Impulsive or salt and pepper noise can occur in image due to a random bit error in a communication channel. The Median filter substitutes a pixel by the Median of all pixels in the neighbourhood particularly. The



Binary morphological operators are applied on the binarized image. The main purpose of the morphological operators is to eliminate hurdle and noise from the image and also to fill the gaps in the binarized image. After applying the morphological operator, the white pixel regions are removed so that the skull stripped image can be obtained.

3. Image Segmentation:- The Image segmentation technique is actually used to separate the grains components. The field of mathematical morphology contributes to a wide range of operators to image processing. We are only handling binary images. For a binary image, black pixels ("0") are normally taken to represent background regions, while white pixels ("1") denote foreground. The two most basic operations in mathematical morphology are dilation and erosion. These operations can be considered as morphological non-linear filters. The segmented image undergoes a series of morphological operations to detect the exact shape of the object.

4. Feature Extraction: - Object measurement is made on a per-object basis, with one set of measurements for each object detected in a field. Typically of morphological characteristics measurements are of the objects detected, and can include size, shape, position, intensity, color, count etc. The region props function supports several properties that can be used with gray scale images, including 'Area', 'Mean Intensity', 'Min Intensity', and 'Max Intensity' and so many factors. These properties use the original pixel values of the sample object for their calculations.

5. Seeds classification: - The classification is performed using Artificial Neural Network. Artificial neural networks (ANNs) are inspired by biological neural networks. An ANN is an information processing paradigm that is inspired by the way biological nervous system process information. A trained neural network is of as an expert in the category of information it has been given to analyze. Some of the advantages of an ANN are: It has the capacity of adaptive learning. It has the capacity of self-organization. It has the capacity of real-time operations. An ANN consist of three layers the input layer, the hidden layer and the output layer with two modes of operation: the Training mode and the Testing mode. In the training mode the ANN is trained with the input feature values and in testing mode validation of the used algorithm is carried out. Artificial Neural Network (ANN) classifications are used for this purpose. Here, classification is done using the ensemble classification based on the neural network. In this the back propagation algorithm is used as the base classifier using different methods. The classification is done on the basis of features extracted from the segmented images. Simple averaging is used for the combination of results from each classifier. The data is divided into the training and testing data. The data classified as training data is then used for training the neural network. Remaining data is used for the testing purpose. The training and testing dataset involves both normal as well as abnormal cases. Then classification is done into normal and abnormal. It provides better results compared to single neural network that has accuracy of about 90%. The work flow for any problems has seven primary steps. (Data collection in step 1, it is important, generally this step is occurs outside the MATLAB environment). 1. Collect data 2. Create the network 3. Configure the network 4. Initialize the weights and biases 5. Train the network 6. Validate the network 7. Use the network

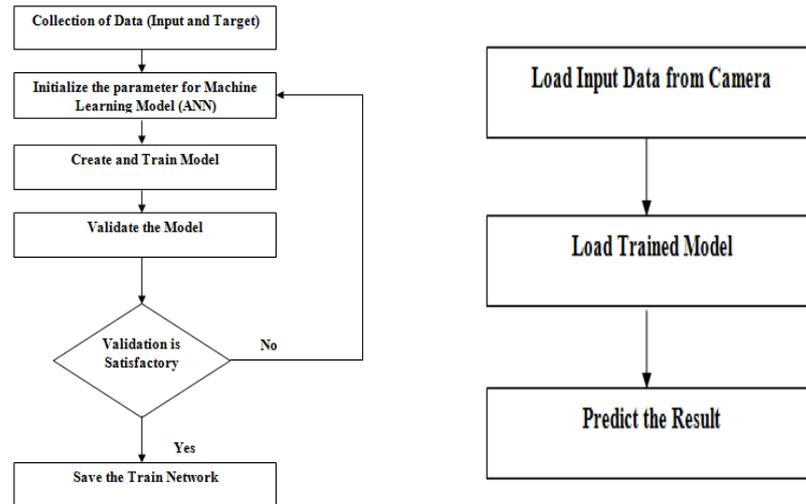


Figure 2:-a) Learning Phase

b) Testing Phase

Algorithm: -

Step 1: Determine the architecture Number of input and output neurons Hidden neurons and layers.

Step 2: Initialize all weights and biases to small random values, typically $\in [-1, 1]$, choose a learning rate η

Step 3: Repeat until termination criteria satisfied Present a training example and propagate it through the network

(Forward pass) Calculate the actual output.

1. Inputs applied
2. Multiplied by weights
3. Summed
4. Quashed by sigmoid activation function
5. Output passed to each neuron in next layer

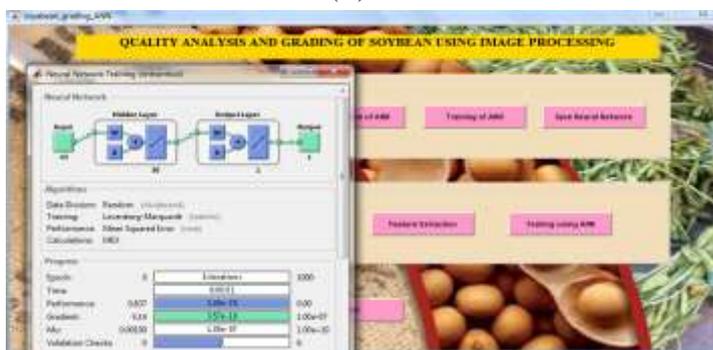
6. Report Generation: - It is very difficult to analyse the sample based on text data, so that graphical analysis is also provided. Different bar charts provides the length and width, size of seed distributions among the sample. It also shows the bar chart for distribution of aspect ratio among grain sample. Size measurement pie charts are provided according to organisation criteria. Colour classification pie charts are generated to identify colour group detail in grain sample. All different pie charts show the % distribution of particular group classification in sample.

Also the report is generated for grading of soybean as Grade 1, Grade 2 and Grade 3 seeds. According to grading of soybean we can generated the market value of soybean.

RESULT



(A)



(B)

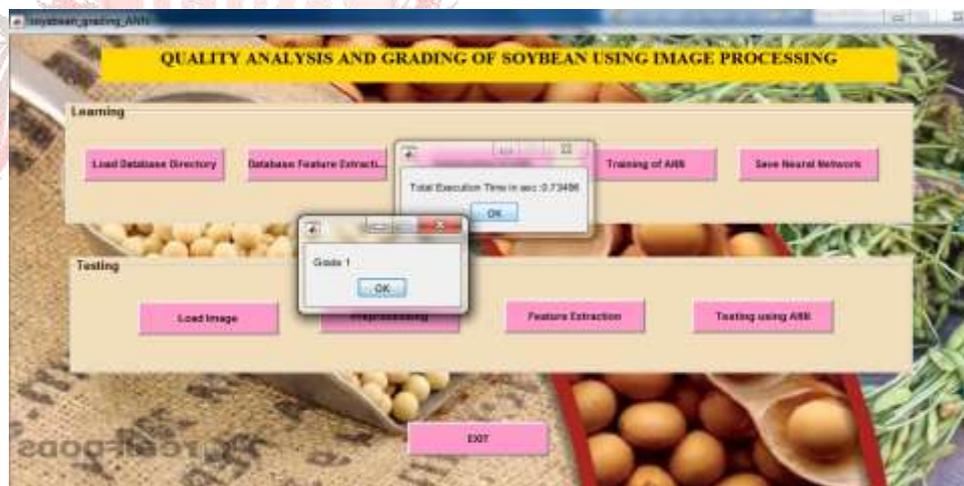


(C)

(D)

Output Class \ Target Class	1	2	3	4
1	10 91.6%	1 8.3%	0 0%	0 0%
2	0 0%	10 100%	0 0%	0 0%
3	0 0%	0 0%	10 100%	0 0%
4	0 0%	0 0%	0 0%	10 100%

(E)



(F)

Figure 3:- Result of System A) Front end of system B) Load database for training of NN C) Training of NN D) Input image for testing E) Confusion Matrix F) Testing input image with ANN showing results as Grade 1

CONCLUSION

This paper will uses colour and geometrical features as attributes for classification using image processing techniques and artificial neural network for machine learning. This method requires minimum time and it is low in cost. As per the literature survey many authors grades the soybean by physical separation by the human and

image processing tools. It has been proved in this paper not only grading of soybean but also separate the good quality of soybean depends upon morphological features by using machine learning. Various samples are considered of soybean seeds, Morphological and colour features are extracted from the images and are stored in the database. A feed forward back propagation neural network is predominantly a classifier that maps any input pattern to a number of classifications with samples. From confusion matrix also we proved the 98.9% accuracy of system. The expected output is much better than the previous methods.

REFERENCES

- Md Abdul Momin, Kazuya Yamamoto, Munenori Miyamoto and Naoshi Kondo "Machine vision based soybean quality evaluation" M.A. Momin et al. / Computers and Electronics in Agriculture Elsevier Volume 140, August 2017, Pages 452-460
- Salome Hema Chitra,S.Suguna and S.Naganandini Sujatha, "A Survey on Image Analysis Techniques in Agricultural Product", Indian Journal of Science and Technology, Vol 9(12), DOI:10.17485/ijst/2016/v9i12/77171, March 2016
- Timothy J. Herrman, Extension State Leader Grain Science and Industry Carl Reed "Soybean Grading Procedures" Publications from Kansas State University are available on the World Wide Web at: <http://www.oznet.ksu.edu>
- D. Wang, M. S. Ram, F. E. Dowell, "Classification of damaged soybean seeds using near infrared spectroscopy" Vol. 45(6): 1943-1948 2002 American Society of Agricultural Engineers ISSN 0001-2351
- Sudhir Gupta, Smitha Patil, Asha Kulkarni and Zoheb Hukkeri "Classification of Soybean seeds by Color Image Analysis" 5th International Conference of the Asian Federation For Information Technology in Agriculture (AFITA), Bangalore, 9-11 November 2006
- Irfan S. Ahmad, Research Associate, John F. Reid, Professor, Marvin R. Paulsen, Professor "Color classifier for symptomatic soybean seeds using image processing." APS Publication, Plant Dis. 83:320-327.
- Casady, W. W., Paulsen, M. R., Reid, J. F., and Sinclair, J. B. 1992. A trainable algorithm for inspection of soybean seed quality. Trans. ASAE 35:2027-2034.
- Ms.Sushma,Mr.Krishan Kumar,Mr.Rajender Kumar quality testing and grading of food grains using digital image processing IJEEEVol No.8 Issue No. 02 July-December 2016
- AmanChadha, SushmitMallik, RavdeepJohar, "Comparative Study and Optimization of Morphological Feature-Extraction Techniques for grains Image Retrieval" International Journal of Computer Applications (0975 -) Volume 52- No.20,pp.0975-0979, August 2012
- Deepika Sharma,Sharad D. Sawant Department of Electronics and Telecommunication, NBN Sinhgad School of Engineering, Savitribai Phule Pune University, Pune, India "Grain quality detection by using image processing for public distribution" Publication Year: 2017, Page(s): 1118 - 1122,2017 IEEE International Conference on Intelligent Computing and Control Systems (ICICCS).
- Mayur P. Raj Research Scholar, Rai University, Ahmedabad, Gujarat, India Dr. Priya R. Swaminarayan Professor, ISTAR, Anand, Gujarat, India "Applications of Image Processing for Grading Agriculture products" International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 3 1194 - 1201.
- "Digital image processing using MATLAB " Third edition" by Rafael C. Gonzalez.
- "Computerized Control Systems in the Food Industry" by Gauri S. Mittal.

E-ISSN NO:2349-0721