

An Intelligent system for identification of Indian Lentil types using Artificial Neural Network (BPNN)

Narendra V G¹ and Muhammad Abdorrazzaghi²

¹(Department of CS and E, Manipal Institute of Technology/Manipal University, India)

²(IAESTE India/MIT Student, Iran)

Abstract: Lentils are designated into two classes, Red Lentils and Lentils other than red. The method of determining the class of a lentil is by seed coat color. Red lentils may be confirmed by the cotyledon color. Lentil varieties may have a wide range of seed coat colors from green, red, speckled green, black and tan. The cotyledon color may be red, yellow or green. The size and color of each Indian Lentil type (i.e. Red, Green and Yellow) is determined to be large or Medium or small, then size and color becomes part of the grade name. An Intelligent system is used to identify the type of Indian lentils from bulk samples. The samples were acquired from the proposed image acquisition system with image resolution 640x480. The proposed system facilitating kernels size and color measurement using image processing techniques. These Lentil size measurements, when combined with color attributes of the sample, classify three lentil varieties commonly grown in India with an accuracy approaching 97.08%.

Key Words: Size and Color measurement, Lentil, Image Processing, Artificial Neural Network

I. Introduction

Lentil was one of the earliest cultivated crops in the world with archeological evidence from the early Stone Age. India leads the world in lentil production, and most of the lentils are consumed by India and its neighbor countries. Historically, lentil was widely used in India, Southwest Asia, and the Mediterranean areas in the form of split lentil (dhal) and it is still an important source of dietary protein in these areas.

Lentils are designated into two classes, Red Lentils and Lentils other than red. The method of determining the class of a lentil is by seed coat color. Red lentils may be confirmed by the cotyledon color. Lentil varieties may have a wide range of seed coat colors from green, red, speckled green, black and tan. The cotyledon color may be red, yellow or green. There are more than a dozen types of lentils grown in India. Many people use the word pulse to represent lentils with small seeds as opposed to the relatively large seeds of beans and chickpeas.

In India, large red lentils are preferred by consumers and are used primarily for splitting. Yellow lentils are preferred in southern India where they act as a substitute when *tur* is priced too high or most consumers. Small red lentils are consumed in lesser amounts. In Southern India, split pigeon peas (toor dal) also called Togari bele in Kannada and Tuvaram paruppu in Tamil are one of the most popular pulses, being an important source of protein in a mostly vegetarian diet. In regions where it grows, fresh young pods are eaten as a vegetable in dishes such as sambar.

Lentils have emerged as a specialty crop in the prairies due to their high value. Consequently, the area under lentil production has increased many fold over the last few years. India now exports about 2.0 million tons of lentils annually to neighbor countries like Pakistan, Bangla and Srilanka and other many more countries in the world. Visual appeal directly influences consumer acceptance and hence value of the lentil product. The main visual factor is the size and color of the grain, which is directly related to its market value. Errors in color assessment can lead to incorrect type assignments, especially close to the type boundaries, causing a further perception that visual grading is incorrect. As type is related to value, growers may not be paid what their crop is actually worth [1].

For grading or sorting lentils, two systems have been reported: one developed at the University of Saskatchewan and the other developed by a seed sorting company in Rocanville, Saskatchewan Both of these systems, however, generally rely upon expensive mechanical hardware or difficult to maintain color camera technology, which limits their affordability, portability, consistency, and reproducibility between systems and the ability to handle small samples. [1]

Machine vision systems provide an alternative to manual inspection of Lentil samples for determining characteristic kernel properties. During Lentil handling operations, information on Lentil type and Lentil quality is required at several stages before the next stage of storage and handling. In the present Lentil-handling system, Lentil type and quality are rapidly assessed by visual inspection. This evaluation process is, however, subjective and is limited by the experience and expertise of the individual. The decision making capabilities of a Lentil inspector can be seriously affected by his/her physical condition, such as fatigue and eyesight, mental state

caused by biases and work pressure, working conditions such as improper lighting, and other environmental conditions. Determining the potential of morphological features to classify different Lentil species, classes, varieties, damaged Lentil, and impurities using a statistical pattern-recognition techniques for classification purposes. Efforts also have been made to integrate all these features in terms of a single classification vector for Lentil kernel identification. [2]

The objectives of this research were to develop and optimize a technique for discrimination of various types of Lentils by extracting the morphological features using images of an individual e kernel as Large, Medium and Small. Finally compare the classification accuracies using back propagation neural network and support vector machine classifiers.

II. Concepts

2.1. Lentil samples

For the crop year 2012 of Red, Green and Yellow Lentil samples (see fig. 1) were received from the University of Agricultural Sciences (UAS), Dharwad (Karnataka). These samples were from storage and at least 90 days old. When a grade did not appear to match the visual appearance of the sample, senior inspectors re-inspected the samples and assigned a grade. Samples were about 1 Kg in size.

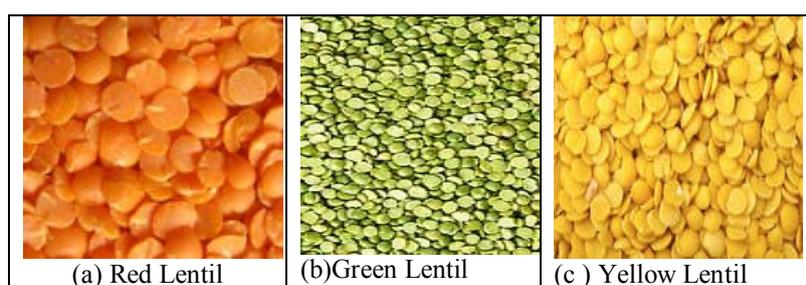


Fig. 1 Types of Indian Lentils

1.2. Dimension Calculation

Unless some epical techniques are used for image acquisition the 3-D information of the image is lost. Hence, only 2-D information is retained in digital images. Hence, this research makes an attempt to calculate the morphological features of an individual Lentil. (Ref. Table 01).

Table 1. The morphological features used in this study.

No	Name	Formula	Description
1	Area	Ap	The actual area based on the number of pixels in the region
2	Length	Sr	The length of the rectangle bounding the Lentil kernel
3	Width	W	The width of the rectangle bounding the Lentil Kernel
4	Major axis length	LM	The length of the major axis is the distance between the end points of the longest line that could be drawn through the object
5	Minor axis length	Lm	The length of the minor axis is the distance between the end points of the longest line that could be perpendicular to the major axis
6	Extent	$E = (Sr \times W) / Ap$	The proportion of the pixels in the bounding box that are also in the Lentil Kernel region (computed as the area divided by area of the bounding box)

2.3. Feret's Diameter

Feret's Diameter (see fig. 2) is also called as the caliper length. It is the longest distance between any two points in the along the boundary of the region of interest. [3] All the Lentil kernels of different types are look like elliptical shape. So we considered major and minor axis as actual length and width of each Lentil kernel type. Then a and b are major axis/2 and minor axis/2 respectively.

Feret's Diameter=Length of Major Axis-Length of Minor Axis.

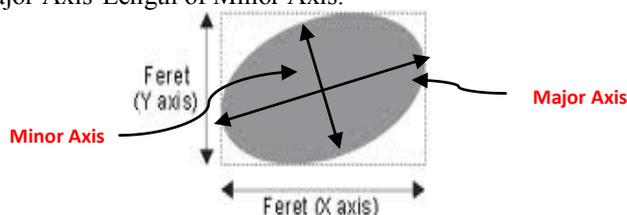


Fig. 2: Illustrates the Feret's Diameter of elliptical shape

2.4. Otsu Method

This method uses the basic clustering idea of using threshold value that minimizes the weighted within-class variance turns out to be the same as maximizing the in between class variance. It works very efficient and fast once the histogram is computed as the method operates directly on the gray level histogram.

III. Methodology

Computer vision system includes methods for acquiring, processing, analysing and understanding images from the real world produce numerical or symbolic information to aid in decision making. The Figure 2 shows the various levels of Image Processing.

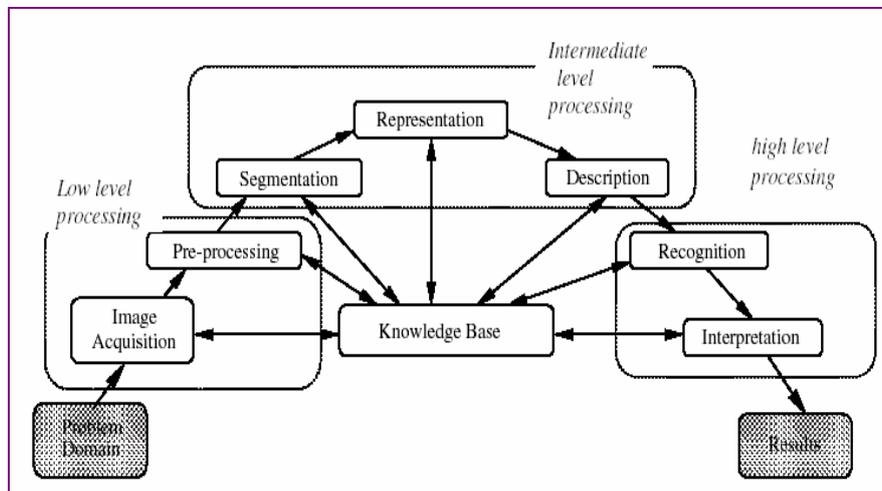


Fig. 2: The various levels of Image Processing [3]

3.1. Image Acquisition

The 640x480 size of images in RGB color space (*R* - red, *G* - green and *B* - blue) Lentil kernels of different types (Red, Green and Yellow) were acquired by using webcam (iBall), in a room with a proper control of lighting intensity(Ref. Table 2). Distances between webcam and Lentil kernels were fixed at 10 Centimeters for all the samples. Figure 3 shows the experimental setup for image acquisition.

Table 2 Lentil types and number of samples acquired for type identification

Sl. No.	Lentil type	Size	Total Samples	Training set	Testing Set
01	Red Lentil	Large	300	200	100
		Medium	250	150	100
		Small	200	150	50
02	Green Lentil	Large	300	200	100
		Medium	250	150	100
		Small	200	150	50
03	Yellow Lentil	Large	300	200	100
		Medium	250	150	100
		Small	200	150	50
Total Samples			2250	1500	750



Fig. 3: Setup for Image Acquisition

3.2. Image Pre-processing

The mean filter has been applied on the acquired image to decrease the noise that has crept in during the digitization/acquisition process. The mean filtering is simply to replace each pixel value in an image with the mean ('average') value of its neighbors, including itself. This has the effect of eliminating pixel values which are

unrepresentative of their surroundings. Mean filtering is usually thought of as a convolution filter. Like other convolutions it is based around a kernel, which represents the shape and size of the neighborhood to be sampled when calculating the mean. Often a 3×3 square kernel is used, although larger kernels (e.g. 5×5 squares) can be used for more severe smoothing. The use of mean facilitates the identification of an optimal threshold for the image during segmentation.

3.3. Segmentation and Representation

Segmentation subdivides an image into meaningful non-overlapping region that is used for further analysis. It is assumed that the regions obtained correspond to the physical parts or objects of a scene (3-D) represented by the image (2-D). Image representation involves representing the segmented image as a boundary or a region. Boundary representation is suitable for analysis of size and shape features. In this context, segmenting an image consists of separating the Lentil kernel from the background of the image. The segmentation carried out by using Otsu's automatic threshold algorithm.

The Otsu's algorithm maximizes the variance between the object and background classes. Pixels that have the values less than the global threshold are classified as Lentil Kernel and the others are classified as background (resulting in a binary image).

3.4. Image Description (measurement) deals with the extraction of quantitative information from the previously segmented image regions. After getting the boundary of the selected region, properties such as area, perimeter, equivalent diameter, major and minor axis length can be determined. The area is the actual number of pixels in the labeled region. The major and minor axis length can represent the actual length and width of a Lentil kernel. This measurement determines the size of a Lentil kernel. The major axis length of the Lentil kernel is the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the regions, while the minor axis length is the minor axis of the ellipse. However, the property of major and minor axis length are only supported for a 2-D image input. The pixel count of the processed image depends on the distance between the camera and the object when the picture is taken. The smaller the distance, the larger the pixel counts. [1]

In the research the morphological features of an individual Lentil Kernel of type Red, Green and Yellow were evaluated as Large, Medium and Small. The Length (l) and Width (w) of an individual Lentil Kernel play vital role in deciding the other measuring of morphological features.

3.5. High level processing involves recognition and interpretation, typically using Soft Computing Techniques. These steps provide the information necessary for the process/ machine control for quality sorting and grading. Our work is restricted to estimation of the morphological features of a Lentil kernel using the image processing toolbox of MATLAB.

IV. Proposed Methods

We are proposing the algorithm to find morphological characteristics of Red, Green and Yellow Lentil kernel types as Small Medium and Large.

4.1. Algorithm: Determine the morphological characteristics of a Lentil kernel

Input: Original 24-bit color image of size 640x480.

Output: Morphological features.

Start

Step 1: Read the image of resolution 640x480

Step 2: Convert RGB to grayscale image.

Step 3: Filter the image using Mean filter (Remove the noise)

Step 4: Binarize the image using Otsu Method

Step 5: Find the Boundaries

Step 6: Determine the morphological features of a Lentil kernel of different type from the input image.(Ref. Table 3)

Step 7: Calculate the Minimum, Maximum, Mean(μ) and Standard Deviation(σ) from the extracted morphological features.

$$\mu = \frac{1}{N} \sum_{i=1}^N f_i x_i = \sum_{i=1}^N p_i x_i$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

Stop

Table 3. The morphological features used in this study.

Sl. No.	Name	Formula	Description
1	Area	Ap	The actual area based on the number of pixels in the region
2	Perimeter	P	Computes the perimeter by calculating the distance between each adjoining pair of pixels around the border of the region
3	Major axis length	LM	The length of the major axis is the distance between the end points of the longest line that could be drawn through the object
4	Minor axis length	Lm	The length of the minor axis is the distance between the end points of the longest line that could be perpendicular to the major axis
5	Eccentricity	E	The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length.
6	Equivalent Dimeter	Eq= sqrt(4*Ap/pi)	Scalar that specifies the diameter of a circle with the same area as the region.

4.2. Algorithm: Color Feature Extraction

Input: Original 24-bit color image of size 640x480.

Output: 9 color features.

Start

Step 1: Separate the RGB components from the original 24-bit input color image

Step2: Obtain the Red, Green and Blue components.

Step 3: Compute mean, variance, and range for each Red, Green and Blue components using the equations.(Ref. Table 4)

Stop.

Table 4: Color Features of Red, Green and Yellow Lentils from the image

Sl No.	Feature	Formula
1	Red Mean	$Mean \mu = \sum_x x \sum_y P(x,y)$
2	Blue Mean	
3	Green Mean	
4	Red Variance	$Variance V = \sum_{x,y} (x - \mu)^2 P(x,y)$
5	Green Variance	
6	Blue Variance	
7	Red Range	$Range = Max(P(x,y)) - Min(P(x,y))$
8	Green Range	
9	Blue Range	

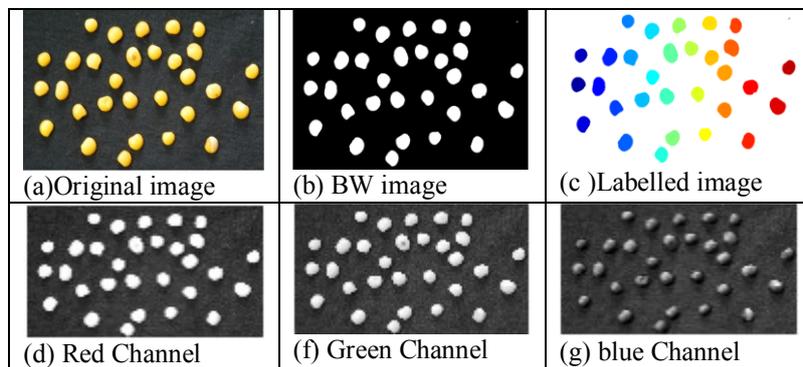


Fig. 5 The Output of Algorithm 1 and 2

Table 5. Morphological features of Red Lentil(Large)

Features	Min	Max	Mean	Std.Deviation
Area	9486	16585	12232.55	2831.965
Perimeter	46.87006	487.1442	412.3628	74.34578
Major axis Length-Minor axis Length	4.52897	32.9334	19.05621	7.386151
Eccentricity	0.267332	0.952297	0.511712	0.116077
EquivDiameter	11.72646	145.3158	122.8473	22.35211

4.3. Classification

Artificial Neural Network: The multilayer feed forward network model with back propagation algorithm for training is employed for classification. The number of input neuron in the input (I) layer is equal to the number extracted morphological and color features. The number of neurons in the output layer is equal to the Lentil types (i.e. Red, Green and Yellow) as Large Medium and Small of samples considered. In case of identification based on morphological features, the number of neurons is 6, based on color features the number of neurons is 9 and 15 when features are combined. In all the cases, the output(O) layer has 9 nodes. The network recognizes a pattern vector P belonging to class O_i if the i^{th} output of the network is high while all other outputs are low. The training samples are chosen such that every possible pattern of the image is reflected. Inspite of training, if unexpected input is encountered the recognition accuracy suffers. Regarding precision computation complexity, we have not concentrated but used ANN Tool Box available on MATLAB 2010b. [5]

The number of nodes in the hidden layer is calculated using the formula

$$\text{No. of Hidden layer Neurons (n)} = ((I+O)/2) + y^{1/2}$$

Where I = No. of input features

O = No. of output outputs and

y = No. of input patterns in the training set.

Table 6: Output of the ANN using feed forward back Propagation algorithm

Sl. No.	Lentil type	Size	Accuracy
01	Red Lentil	Large	97.45%
		Medium	98.12%
		Small	95.65%
02	Green Lentil	Large	96.58%
		Medium	95.28%
		Small	97.89%
03	Yellow Lentil	Large	96.67%
		Medium	98.12%
		Small	97.88%
Average accuracy			97.08%

V. Results and Discussion

The selection of parameters for the discrimination of features within the image is based upon experience and co-operation and advice of Professors at University Agricultural Sciences(UAS) Dharwad, India. The objective is to have a standard and repeatable set of parameters so that lentil types can be consistently identified by Intelligent System. Samples from successive crop years, graded by the same inspectors each year did demonstrate consistency in Lentil type identification.

Morphological operations successfully separated the type of Lentil boundaries for Lentils kernel size measurements from the bulk sample images (Fig. 5). Successive erosion and ultimate dilation operations worked for large, Medium and small kernels of each Lentils type. The condition of perimeter of objects excluded most of the unwanted objects from entering the data set, i.e., Lentils kernel at the image boundary, Lentils kernel with joined boundaries, and Lentils kernel positioned incorrectly were not measured. A few apparently non-circular blobs present in the processed image though undesirable are inevitable. While setting a tighter condition of perimeter could remove these ‘undesirable’ objects, this can also remove some desirable blobs as Lentils themselves are not perfectly elliptical. Figure represent the final processed images used for size and color measurements. Success in separating Lentils kernel boundaries kept the Intelligent system simple and cost effective. It allowed for size and color measurements from the bulk sample images eliminating the need for a Lentils kernel presentation or separation device that would add to the cost and complexity of the Intelligent System.

Analyses the individual image of each Lentils type features indicated that the mean and standard deviation Lentils kernel size (area, perimeter, major and minor axis etc.) and the color (mean, variance and range) of each Lentils type in the image when combined into a artificial neural network classifier, can be used to identify all three types of lentils as Large, Medium and Small.

When using size and color features as the input variables, the artificial neural network classifier identified all three Lentils type with an overall accuracy approaching 97.08%. (Ref. Table 6)

VI. Conclusions

The following conclusions were drawn from this study.

1. Morphological erosion followed by ultimate dilation can effectively separate seed boundaries to facilitate Lentils kernel size measurements on bulk lentil samples.

2. The Lentil kernel morphological features and the overall color of a sample can accurately predict lentil type using an artificial neural network classifier.

References

- [1] M.A. Shahin and S.J. Symons, A machine vision system for grading lentils Canadian Biosystems Engineering Volume 43 2001
- [2] N.S. Visen¹, D.S. Jayas, J. Paliwall and N.D.G. White, Comparison of two neural network architectures for classification of singulated cereal grains, Canadian Biosystems Engineering Volume 46 2004
- [3] Da-Wen-Sun "Computer Vision Technology for Food Quality Evaluation", Food Science and Technology, International series, Elsevier Inc., 2008.
- [4] B S Anami and D G Savakar Improved Method for Identification and Classification of Foreign Bodies mixed Food Grains Image analysis, ICGST-AIML Journal, ISSN 1687-4846, Vol. 9(1), 2009