Automation of Classification and Grading of Rice Grains

K A Spoorthi[1] , Damini P Shenoy[2], Geethanjali N[3], Mainashree M K[4], Sandeep R[5]
[12345]*Department of ECE, Vidyavardhaka College of Engineering, Mysore-570002*

Abstract– Rice is the agricultural commodity with the third-highest worldwide production. As a cereal grain, it is the most widely consumed staple food for a large part of the world's human population, especially in Asia. There are different types of rice with different qualities to suit different consumer preferences. In general, many countries quantify rice into four main categories which includes milling quality, cooking, eating and processing quality, nutritional quality and specific standards for cleanliness, soundness and purity. The quality grade testing of rice grains is growing rapidly among the consumers in fulfilling the above mentioned requirements. Quality factors relate to grain length, stickiness, aroma, texture, flavor and nutritional content. The varietals purity is one of the factor whose inspection is difficult and more complicated. The manual grading of rice consumes more time and is influenced by human factors and working conditions such as eyesight, mental state, improper lighting, and climate. Thus a model of quality grade testing and identification is proposed based on features such as the color, area, major axis and minor axis using SVM algorithm.

Keywords- Image processing, Thresholding, SVM algorithm.

I. INTRODUCTION

Rice (Oryza Sativa L.) is one of the most important food crops as it is consumed by half of the world's population. It is a widely used agricultural commodity. Production of rice per annum is in hundreds of millions of tons. Rice grains are generally differentiated based on size, color and shape. Product quality is the driving force for every enterprise, which is an important factor to keep an impregnable position in the modern global competitive environment. Product quality inspection or monitoring is basically done by the performance tests of products as well as appearance assessments, to avoid the possible defects in products and ensure customer satisfaction. Deceptive naming of one variety of rice grains as another is a major issue in the food industry. The International Rice Research Institute (IRRI) and British Retail Consortium (BRC) consider the appearance of rice grains being unfavorable to consumers. These organizations have established standards and regulations for evaluating varieties of rice grains. These standards also grant special status to particular rice varieties such as basmati due to their high trade value. Some physical parameters like aspect ratio and minimum average length are specified clearly in the standards. India is one of the largest producers of rice in the world after china; India is imposing and administering the use of rice grain standards to assure that producers get maximum price for their product. The quality analysis of rice grains is still carried out manually by trained professionals. This strategy suffers from manydrawbacks like, it is subject driven and is influenced by human factors and working conditions. Human perception can easily be influenced by external factors like physical conditions such as fatigue, eyesight, mental state, improper lighting and climate that results in inconsistent results. The rate of cleaning and recovery of damaged rice is limited. For the given reasons, these tasksrequire machine work and develop imaging systems that can be beneficial for quality inspection. The computer vision based methods or image processing can be used to overcome above mentioned drawbacks.

The objective of this study is to develop a simple automated process to deliver high accuracy results for quality testing, grading and classification of rice grains.

II. MATHEMATICAL FRAMEWORK

Support Vector Machine (SVM) is an effective technique for classifying high dimensional data. Unlike the nearest neighbour classifier, SVM learns the optimal hyper plane that separates training examples from different classes by maximizing the classification margin. It is also applicable to data sets with a non-linear decision surfaces by employing a technique known as the kernel trick, which projects the input data to a higher dimensional feature space, where a linear separating hyperplane can be found. SVM avoids the costly similarity computation in high-dimensional feature space by using a surrogate kernel function. It is known that SVM are capable of effectively processing feature vectors of some 10,000 dimensions, given that these are sparse. Several authors have shown, that SVM provides a fast and effective means for learning text classifiers from examples. Documents of a given topic could be identified with high accuracy.

SVM's are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. In addition to performing linear classification, SVM's can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. When data is unlabelled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups. The support-vector clustering algorithm, created by HavaSiegelmann and Vladimir Vapnik, applies the statistics of support vectors, developed in the SVM algorithm, to categorize unlabeled data, and is one of the most widely used clustering algorithms in industrial applications.

- LINEAR SVM

Given the training dataset of n points of the form$(\vec{x}_1, y_1)\ldots(\vec{x}_n, y_n)$ where the y_i are either 1 or -1, each indicating the class to which the point$\vec{x}_i$ belongs. Each $\vec{x}_i$is a p-dimensional real vector. We want to find the "maximum-margin hyperplane" that divides the group of points$\vec{x}_i$ forwhich y_i=1 from the group of points for which y_i=-1, which is defined so that the distance between the hyperplane and the nearest point $\vec{x}_i$from either group is maximized.

Any hyperplane can be written as the set of points $\vec{x}$satisfying

$$\vec{w}.\vec{x} - b = 0$$

where$\vec{w}$is the (not necessarily normalized) normal vector to the hyperplane. This is much like Hesse normal form, except that $\vec{w}$is not necessarily a unit vector. The parameter $\frac{b}{\|\vec{w}\|}$determines the offset of the hyperplane from the origin along the normal vector$\vec{w}$.

- Hard-margin

If the training data is linearly separable, we can select two parallel hyperplanes that separate the two classes of data, so that the distance between them is as large as possible. The region bounded by these two hyperplanes is called the "margin", and the maximum-margin hyperplane is the hyperplane that lies halfway between them. With a normalized or standardized dataset, these hyperplanes can be described by the equations.

$\vec{w}.\vec{x} - b = 1$ (Anything on or above this boundary is of one class, with label 1)

$\vec{w}.\vec{x} - b = -1$ (Anything on or below this boundary is of the other class, with label -1)

Geometrically, the distance between these two hyperplanes is$\frac{2}{\|\vec{w}\|}$, so to maximize the distance between the planes we want to minimize $\|\vec{□}\|$.The distance is computed using the distance from a point to a plane equation. We also have to prevent data points from falling into the margin, we add the following constraint: for eachieither:

$\vec{□} \cdot \vec{□}_i - □ \geq 1$, if$y_i$=1, or

$\vec{□} \cdot \vec{□}_i - □ \leq -1$, if$y_i$=-1

These constraints state that each data point must lie on the correct side of the margin.

This can be rewritten as

$\vec{□}_i\,(\vec{□} \cdot \vec{□}_i - □) \geq 1$, for all$1 \leq □ \leq □$(1)

We can put this together to get the optimization problem:

"Minimize $\|\vec{□}\|$subject to$□_i$ $(\vec{□} \cdot \vec{□}_i - □) \geq 1$ for $□ = 1, \ldots\ldots, □$"

The $\vec{□}$ and □ that solve this problem determine our classifier,

$$\vec{□} \mapsto □□□(\vec{□} \cdot \vec{□} - □).$$

An important consequence of this geometric description is that the max-margin hyperplane is completely determined by those $\vec{x}_i$that lie nearest to it. These $\vec{x}_i$are called support vectors.

- Soft-margin

To extend SVM to cases in which the data are not linearly separable, we introduce the hinge loss function,

$□□□(0, 1 - y_i\,(\vec{□} \cdot \vec{□}_i - □))$.

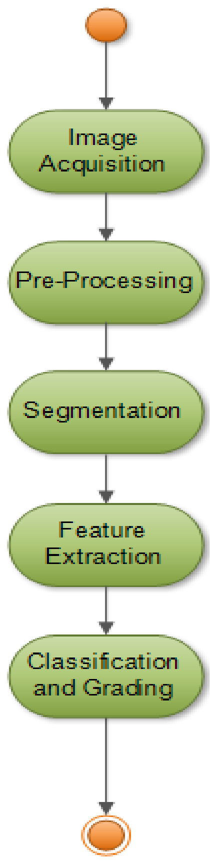
Note that y_i is the $□ - □□$ target (i.e., in this case, 1 or -1), and $\vec{□} \cdot \vec{□}_i - □$is the current output.

This function is zero if the constraint in equation (1) is satisfied, in other words, if$\vec{□}_i$lies on the correct side of the margin. For data on the wrong side of the margin, the function's value is proportional to the distance from the margin [[23]].

III. PROPOSED SYSTEM

The main purpose of this project is to provide some new ideas and develop a proper sequence of image processing methods for evaluating rice appearance quality. The potency and accuracy of inspection have been improved through these methods.

The objective of this work is to develop an algorithm to grade different varieties of rice grains into grade 1, grade 2 and grade 3 with help of SVM by distinguishing between healthy and unhealthy rice by rough rice (head Rice, broken, brewers, chalky, long, short, slender, bold and round).



- ALGORITHM

Step1: Image acquisition using high pixel camera.
Step2: Captured image is pre-processed and RGB image is converted to gray scale image.
Step3: Enhancing the quality of the image to improve its visible impacts.
Step4: Threshold based segmentation is performed to obtain binary image.
Step5: The geometric features such as major axis, minor axis and area of all the individual grains are extracted.
Step6: SVM classifier is used to recognize different varieties of rice grains.
Step7: Analysis on the quality using the average values of the features extracted are performed.

- IMAGE ACQUISITION

Rice grain image 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 rice grain images are acquired with a mobile camera. The considered camera includes optical image stabilization and video stabilization, a quardLED true tone flash and the ability to shoot 63 Mega pixel panorama photos. The phone features 5x digital zoom and has dual camera 12 Mega pixel system allowing for 2x optical zoom and 10x software zoom. To collect data a camera has been placed at a location situated with a plane normal to the object's path. A black background was used.

- PRE-PROCESSING

a. GRAY-LEVEL CONVERSION**:**

Each pixel in the image is specified by three values and each value is assigned as red, blue, and green. The array of class single or double pixel specify intensity values. So it converts RGB to Gray scale image. For single or double arrays, values range from [0, 1]. For unsigned integer 8-bit, values range from [0, 255]. For unsigned integer 16-bit, values range from [0, 65535]. In this work, the image has been taken from the RGB color in jpeg format. It has high pixel rate, due to which error occurrence has been reduced. Gray scale image is an image that each pixel of rice holds a single sample, the rice intensity information, also known as black-and-white image. The intensity is calculated by common formula: 30% of Red + 59% of Green + 11% of Blue. After processing the gray scale level for image, it has only black-and-white pixels. It varies from black at the weakest intensity to white at the strongest intensity.

It converts the gray scale to a binary. The output image replaces all pixels in the input image with luminance greater than level with the value 1 (white) and replaces all other pixels with the value 0 (black). It specifies the level in the range [0, 1], regardless of the class of the input image. The function gray thresh can be used to compute the level argument automatically in order to separate an object in the image from the background. The color of the object (usually white) is referred to as the foreground color. The rest (usually black) is referred to as the background color.

b. RESIZE:

Image resizing is necessary when you need to increase or decrease the total number of pixels, whereas remapping can occur when you are correcting for lens distortion or rotating an image. Zooming refers to the increasing of quantity of pixels, so that when an image is zoomed, more detailed image is obtained.

- IMAGE ENHANCEMENT

Image enhancement techniques have been widely used in many applications of image processing where the subjective quality of images is important for human interpretation. Contrast is an important factor in any subjective evaluation of image quality. Contrast is created by the difference in luminance reflected from two adjacent surfaces. In other words, contrast is the difference in visual properties that makes an object distinguishable from other objects and the background. In visual perception, contrast is determined by the difference in the colour and brightness of the object with other objects. Our visual system is more sensitive to contrast than absolute luminance; therefore, we can perceive the world similarly regardless of the considerable changes in illumination conditions.

a. Median Filtering

The purpose of smoothing is to reduce noise and improve the visual quality of the image. Often, smoothing is referred to as filtering. In this work, median filter is used.

A median filter is a non-linear digital filter which is able to preserve sharp signal changes and is very effective in removing impulse noise (or salt and pepper noise). An impulse noise has a gray level with higher or lower value that is different from the neighbourhood point. Linear filtering technique is known for signalling and for being particularly effective in removing impulse noise of rice. Thus median filters have advantages over linear filters for this type of noise. Therefore median filter is widely used in digital signal and image/video processing applications window of odd size (e.g. $3x3$ windows) over an image. At each window position, the sampled values of signal or rice images are stored and the median value of the samples are replaced.

- SEGMENTATION

The next stage is image segmentation which is the process of partitions an image. It's the process of segment the similar attributes into image which is done with the help of the K means classification approach.

Image segmentation is the first step of image analysis. An image is subdivided into its constituent parts or objects. The image is usually subdivided until the objects of interest are isolated. There are so many approaches for segmentation algorithms. In this study, our main goal is to extract the rice from the background. Maximum variance method is used to segment the image into foreground and background regions.

The result of segmentation is usually a binary image. A binary image contains only two types of pixels: the pixels having a gray level value of either 0 or 1.So after segmentation the rice image will be converted into the binary image [0, 1].

- FEATURE EXTRACTION

The next stage is feature extraction where the scale invariant feature transform is used to derive important features from the segmented region. The method retrieves the feature according to the relative position because it does not change from one image to another image.

a. Morphological feature extraction

The following morphological features were extracted from images of individual rice seeds.

i. Area: This refers to the amount of pixels in the region. The algorithm calculated the number of pixels inside, and including the seed boundary.
ii. Major axis: It is the distance between the end points of the longest line that could be drawnthrough the seed. The major axis endpoints were found by computing the pixel distance between every combination of border pixels in the seed boundary.
iii. Minor axis: It is the distance between the end points of the longest line that could be drawn through the seed while maintaining perpendicularity with the major axis.

b. Color feature extraction

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values. Histograms can also be taken from color images - either individual histograms of red, green and blue channels can be taken or a 3-D histogram can be produced, with the three axes representing the red, blue and green channels, and brightness at each point representing the pixel count. The exact output from the operation depends upon the implementation - it may simply be a picture of the required histogram in a suitable image format or it may be a data file of some sort representing the histogram statistics.

- IMAGE CLASSIFICATION

A SVM based classification system has been presented in this work to classify the different types of rice grain images. It provides enhanced accuracy in feature based image identification and classification as compared to other techniques. It is also an effective technique for classifying high dimensional data. Unlike the nearest neighbour classifier, SVM learns the optimal hyper plane that separates training examples from different classes by maximizing the classification margin.

IV. ADVANTAGES

- Good Efficiency.
- Maintenance cost is low.
- Process is fast.
- It is a non-destructive process.

V. SOFTWARE REQUIREMENTS

- Hardware Requirements

The following requirements are only the minimal requirements to run this utility more successfully and efficiently. There should be sufficient memory and software tools for efficient processing.

Processor	Intel Pentium 4 or equivalent
Speed	2 Ghz
RAM	4 GB
Hard Disk	120 GB

- Software requirements

The software requirements explain the software components that we need to develop our project. It is selected such a way that it reduces our work and easy to implement the project anywhere.

Operating System	Windows 7
Tool	Matlab R2017a
Language	Matlab R2017a

VI. RESULTS

The discussed methodology can be made clearer by considering an original image of Rice grain, shown in Figure 1, and then performing all the operations in a sequence.

Fig.1: Original Image

The original RGB image is now converted into gray scale further converting it to get binary image of the rice grains as shown in Figure 2

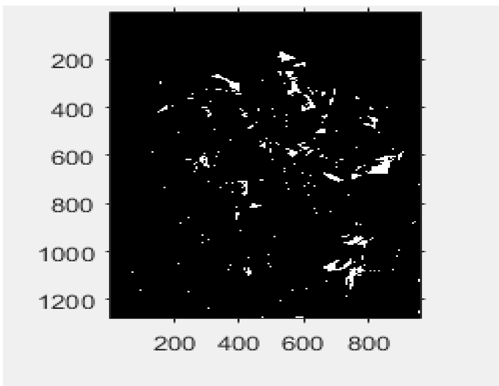


Fig.2: Binary Image

The segmented image is obtained from the binary image to extract geometrical features of the rice grains, as shown in Figure 3.

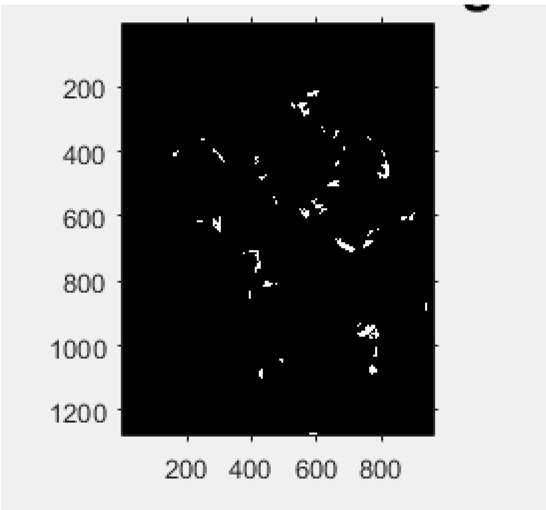


Fig.3: Feature Extraction

To obtain the quality of color features, we have plotted histogram of rice grains as shown in Figure 4.

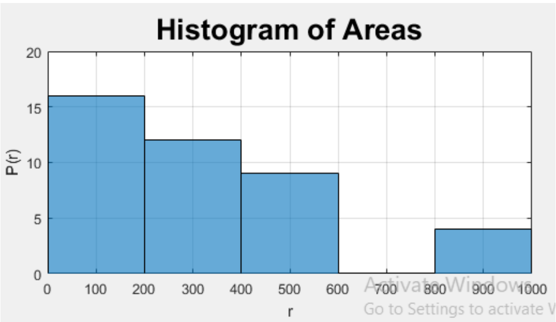


Fig.4: Histogram plotting of the image

The calculated smallest area of individual rice grain is displayed in the dialogue box, shown in Figure 5.

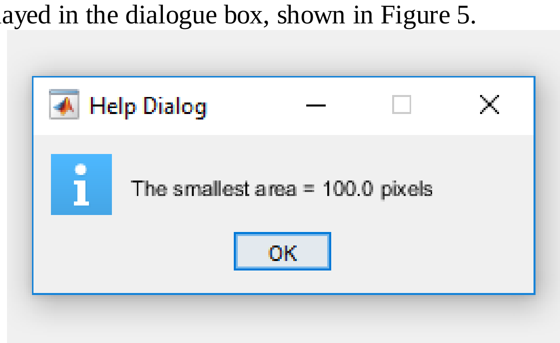


Fig.5: Calculation of smallest area of individual rice grains

The result of classification is displayed in dialogue box, shown in Figure 6.

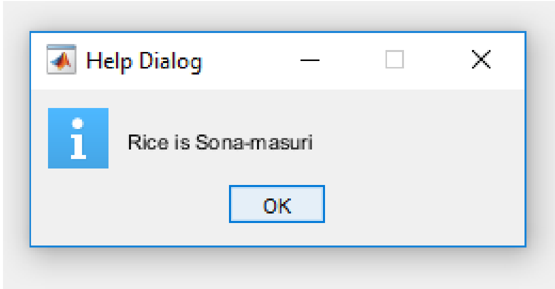


Fig.6: Classification of rice grains

The grading of classified rice grain is displayed in the dialogue box, shown in Figure 7.

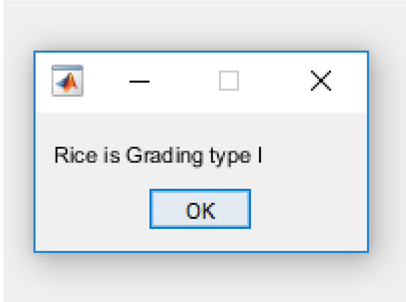


Fig.7: Grading of the classified rice

The outlook of the output window is shown in Figure 8, Figure 9 and Figure 10.

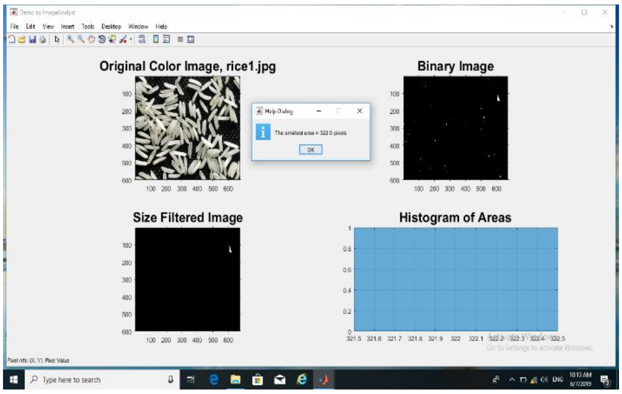


Fig.8: Area Calculation

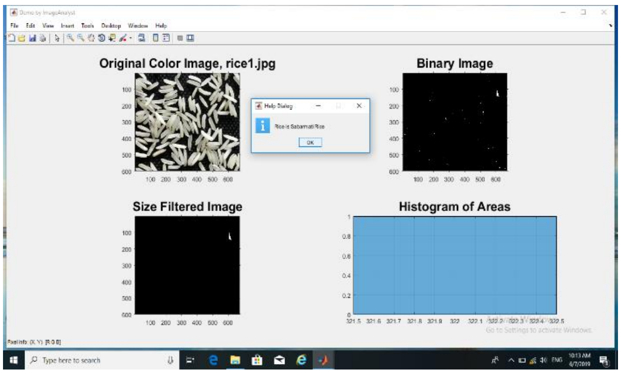


Fig.9: Classification

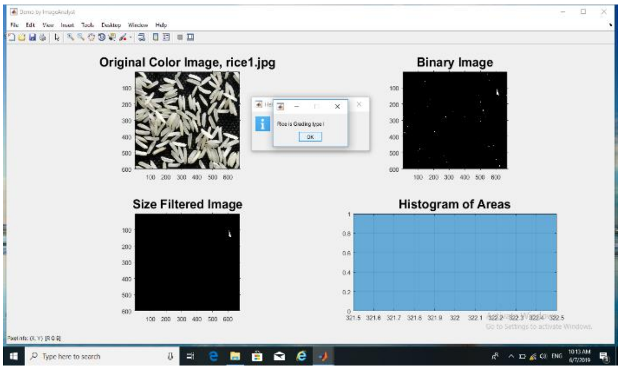


Fig.10: Grading

For this work here, we have considered 35 different varieties of rice grain images. We have used SVM classifier for classifying the grains and grade them as grade 1, grade 2 and grade 3.

Rice Varieties	Number of Samples	Classification Results	Accuracy (%)
Sabarmati	10	8	80%
Basmati	10	6	60%
SonaMasuri	10	6	60%
Brown rice	10	9	90%
White rice	10	6	60%
Boiled rice	10	8	80%
BT rice	10	7	70%
Bullet rice	10	9	90%
Society rice	10	6	60%
Steamed rice	10	8	80%

VII. CONCLUSION

This work is focused on providing a better approach for identification of varieties of rice grains and rice quality based on color and geometrical features using image processing conceptsand SVM classifier. Firstly the image is preprocessed and segmented, then color and geometrical features have been extracted from grain images. These extracted features are used to classify in SVM classifier for further matching process. Different types of grains are considered for identification. Three grading quality of rice is considered and each of the rice grains have been graded into three grades. Even though the problem being worked upon is not completely new, the earlier approaches employed very large number of morphological features, textural and colorfeatures which made the algorithm to be limited to get the output for less number of varieties. Proper setup can be used to reduce these limitations. Here we have considered non-touching grain samples in the approach and combined both the quality analysis and grading of rice to provide a good setup and be able to consider more number of varieties.

VIII. FUTURE SCOPE

The work presented here recognizes only three grading quality of rice. It can be extended to be able to detect quality of rice and also the rate for each quality for customer usage. Other extensions on food grains like identification of fissures, broken rice, mixed rice and plastic rice can also be done. A major developmental work would be to devise an application for daily purposes which is greatly helpful to the consumers.

IX. REFERENCES

[1]. P. Lin, X. Li, Y. Chen, and Y. He, “A deep convolutional neural network architecture for boosting image discrimination accuracy of rice species,” *Food and bioprocess technology*, vol. 11, no. 4, pp. 765–773, 2018.

[2]. D. M. Hobson, R. M. Carter, and Y. Yan, “Characterisation and identification of rice grains through digital image analysis,” in 2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007. IEEE, 2007, pp. 1–5.

[3]. B. Lurstwut and C. Pornpanomchai, “Image analysis based on color, shape and texture for rice seed (oryza sativa l.) germination evaluation,” Agriculture and Natural Resources, vol. 51, no. 5, pp. 383–389, 2017.

[4]. R. Kiruthika, S. Muruganand, and A. Periasamy, “Matching of different rice grains using digital image processing,” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 7, pp. 2937–2941, 2013.

[5]. S. Kaur and D. Singh, “Geometric feature extraction of selected rice grains using image processing techniques,” International Journal of Computer Applications, vol. 124, no. 8, 2015.

[6]. H. Zareiforoush, S. Minaei, M. R. Alizadeh, and A. Banakar, “Qualitative classification of milled rice grains using computer vision and metaheuristic techniques,” Journal of food science and technology, vol. 53, no. 1, pp. 118–131, 2016.

[7]. S. Shantaiya and U. Ansari, “Identification of food grains and its quality using pattern classification,” in International Conference [ICCT-2010], December, 2010, pp. 3–5.

[8]. J. D. Guzman, E. K. Peralta et al., “Classification of philippine rice grains using machine vision and artificial neural networks,” in World conference on Agricultural information and IT, vol. 6, 2008, pp. 41–8.

[9]. O. C. Agustin and B.-J. Oh, “Automatic milled rice quality analysis,” in 2008 Second International Conference on Future Generation Communication and Networking, vol. 2. IEEE, 2008, pp. 112–115.

[10]. Q. Yao, J. Chen, Z. Guan, C. Sun, and Z. Zhu, “Inspection of rice appearance quality using machine vision,” in 2009 WRI Global Congress on Intelligent Systems, vol. 4. IEEE, 2009, pp. 274–279.

[11]. L. Guangrong, “Detection of chalk degree of rice based on image processing,” in 2011 International Conference on Intelligence Science and Information Engineering. IEEE, 2011, pp. 515–518.

[12]. J. S. Aulakh and V. Banga, “Grading of rice grains by image processing,” International journal of engineering research & technology (IJERT), vol. 1, no. 4, 2012.

[13]. B. Verma, “Image processing techniques for grading & classification of rice,” in 2010 International Conference on Computer and Communication Technology (ICCCT). IEEE, 2010, pp. 220–223.

[14]. J. Liu, Z. Tang, P. Xu, W. Liu, J. Zhang, and J. Zhu, "Quality-related monitoring and grading of granulated products by weibull-distribution modeling of visual images with semi-supervised learning," Sensors, vol. 16, no. 7, p. 998, 2016.
[15]. K. Kawamura, H. Asai, S. Kobayashi, S. Souvannasing, P. Sinavong, and T. Inthavong, "The relationship between the physical quality of rice and the market price: A case study in savannakhet, laos, using a bayesian approach," Sustainability, vol. 10, no. 11, p. 4151, 2018.
[16]. K. SRIKAEO and U. PANYA, "Efficiencies of chemical techniques for rice grain freshness analysis," Rice Science, vol. 20, no. 4, pp. 292–297, 2013.
[17]. H. Mebatsion, J. Paliwal, and D. Jayas, "Automatic classification of non-touching cereal grains in digital images using limited morphological and color features," Computers and electronics in agriculture, vol. 90, pp. 99–105, 2013.
[18]. L. Pabamalie and H. Premaratne, "A grain quality classification system," in 2010 International Conference on Information Society. IEEE, 2010, pp. 56–61.
[19]. H. Kaur and B. Singh, "Classification and grading rice using multi-class svm," International Journal of Scientific and Research Publications, vol. 3, no. 4, pp. 1–5, 2013.
[20]. D. Savakar, "Recognition and classification of similar looking food grain images using artificial neural networks," Journal of Applied Computer Science and Mathematics, vol. 6, no. 13, pp. 61–65, 2012.
[21]. B. Anami, D. Savakar, A. Makandar, and P. Unki, "A neural network model for classification of bulk grain samples based on color and texture," in Proceedings of the International Conference on Cognition and Recognition, 2005, pp. 359–368.
[22]. M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "An overview of ensem-´ ble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes," Pattern Recognition, vol. 44, no. 8, pp. 1761–1776, 2011.
[23]. W. S. Noble, "What is a support vector machine?" Nature biotechnology, vol. 24, no. 12, p. 1565, 2006.
[24]. T. Brosnan and D.-W. Sun, "Improving quality inspection of food products by computer vision—-a review," Journal of food engineering, vol. 61, no. 1, pp. 3–16, 2004.