Classification and Grading of Wheat Granules using SVM and Naive Bayes Classifier

Ms. Rupali S. Zambre
Research Scholar,
G.H.R.I.E.M., Jalgaon
zambrerupali@gmail.com

Prof. Sonal P. Patil
Assistant Professor,
G.H.R.I.E.M., Jalgaon
sonalpatil3@gmail.com

Prof. Ganesh N. Dhanokar
Assistant Professor,
G.H.R.I.E.M., Jalgaon
ganesh.dhanokar@raisoni.net

Abstract- India is the second leading producer of wheat in the world. Specifying the quality of wheat manually is very time consuming and requires an expert judgment. With the help of image processing techniques, a system can be made to avoid the human inspection. Classification of wheat grains is carried out according to their grades to determine the quality. Images are acquired for wheat grains using digital camera. Conversions to gray scale, Smoothing, Thresholding, Canny edge detection are the checks that are performed on the acquired image using image processing technique. Classification and Grading of wheat grain is carried out by extracting morphological, color and texture features. These features are given to SVM and Naive Bayes Classifier for classification. To evaluate the classification accuracy, from the total of 1300 data sets 50% were used for training and the remaining 50% was used for testing. The classification system was supervised corresponding to the predefined classes of grades. Results showed that overall accuracy of SVM and Naive Bayes classifier is 94.45%, 92.60% respectively. So, the classification performance of SVM is better than Naive Bayes Classifier.

Keyword: *Smoothing; Thresholding; Canny Edge detection; Support Vector Machine; Naive Bayes Classifier.*

*****

I. INTRODUCTION

Wheat is the very important and widely grown food crop in the world. India is the second largest producer of wheat in the world because of its valuable nutrients. It is the main food resource of protein, energy and dietary fiber in human nutrition and an important industrial raw material as well.

Grain grading system gives assurance that a particular lot of grain meets the required set standards of customers. The quality of wheat has different effect on the yield of wheat; therefore it is important to have the proper inspection of wheat quality. In many countries, grading of grain depends on some properties; test weight, moisture contents, broken foreign material or the percentage fragments example broken corn foreign materials, damaged kernels. These grains consist of several impurities like stones, weed seeds, damaged seeds, more broken granules etc [1].

In the present system, grain type and quality are fastly assessed by visual inspection. Determining the quality of wheat manually is time consuming and requires an expert judgment. The decision-making capabilities of a grain supervisor can be critically affected by his/her physical condition, working conditions such as climate, improper lighting etc and loyalty for traders. Also, the knowledge and experience of inspectors are required to accurately perform this evaluation process. Thus, these tasks require automation and develop imaging systems that can be supportive to grade wheat grains. Different algorithms can be used to classify wheat according to its quality. In this paper, the proposed work contains the Support Vector Machine and Naive Bayes Classifier for classifying the wheat grains according to their grades.

II. RELATED WORK

Alireza Pazoki et al. [2] proposed a system that classifies Six Rain fed wheat cultivars. Color, morphological features and shape factors were extracted. UTA method was used for feature selection. The extracted features were feed to ANN (MLP) for classification. Alireza Khoshroo et al. [3] has done classification of four Irani wheat cultivars. Morphological features were extracted from images. Feature selection (Discriminate Analysis) was conducted to reduce the redundancy in the morphological feature set. The Artificial Neural Network (MLP) was used for classification. Harpreet Kaur, et al., [4] proposed an algorithm to grade the rice kernels. Maximum Variance method was used for segmentation purpose. Ten geometric features are extracted and given to multi-Class SVM is used to classify the rice kernel by examining the Shape, Chalkiness and Percentage of Broken kernels. Nandini Sidnal et al.[5], identified the unknown grain types, impurities and its quality. The morphological and color features are presented to the Probabilistic neural network for identification of grain types, impurities. Sanjivani Shantaiya et al. [6] given an approach to perform texture and morphological based retrieval on a quantity of food grain images. Color, textural, morphological features are extracted and given to Discriminate Analysis (DA) and then classified using ANN. Meesha Punn et al. [7] classified wheat grains using machine algorithms(Neural Network (LM), Support Vector Machine(OVR).). Thresholding is used for image segmentation. The accuracy for classification using NN was 94.5% and 86.8% by using SVM. Xian-Zhong Han et al. [8] researched on grading of wheat seeds. Otsu method for segmentation was used. Morphological features were extracted. Analytical Hierarchy Process (AHP) model calculates maximum feature value, establishes judgment matrix and feature vector matrix, It obtains the main weight

 ISSN: 2321-8169
5322 - 5327

coefficients from unlike solutions, and compares the weight coefficients to obtain the solution which is best.

Related work shows that a lot of work has been done on classification and grading of wheat and classification and grading of other grains like rice, barley etc. Many systems have extracted only morphological or color or textural features. Different methods are used for segmentation and edge detection. Some have used combination of two features to increase the accuracy of classification. Different classifiers like ANN, SVM, and KNN is used for the purpose of classification.

III. METHODOLOGY

Classification and Grading of wheat grains is a system which loads a image of wheat grains, preprocesses the image, edges of each grain is detected, extracts proper image features i.e. morphological, color and textural features, classify the wheat grains based on the extracted image features and the known features are stored in the image model library, and classifies the image according to the degree of similarity between the loaded image and the image models. The Dataflow diagram of Classification and Grading system of wheat grains is given in Fig. 1.

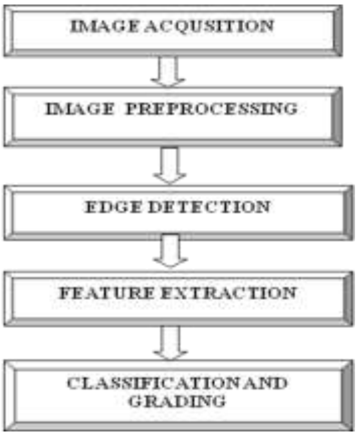

Fig. 1: Dataflow Diagram for Proposed Work

A. Image Acquisition

The Initial step in using a machine vision system is to obtain a digital image. Proper light plays important role in obtaining a good image. This can cause alteration of object features in the image. Determining ideal illumination source is not easy and it depends on the nature [9]. Here, the samples of Wheat grains are collected from local farmers. Wheat grain images are captured using a digital camera. A fixed distance between the camera and the grain samples is maintained. A uniform background which is black in color. The grains are spread on a black sheet randomly. Although grains are placed randomly, make sure that they are not in contact with other. The images were captured and stored in hard disk in JPEG (Joint Photographers Expert Groups) format. The basic building block for image capturing is shown in Fig. 2.

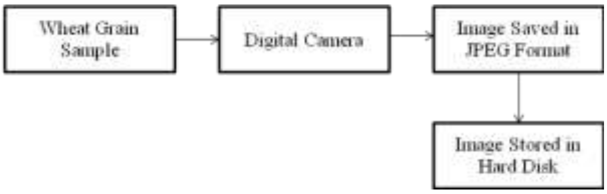

Fig. 2: Basic Building Block for Image Capturing

B. Image Preprocessing

The raw data is subjected to several preliminary processing steps to make it functional in the descriptive stages of classification and grading. In order to get wheat grain features accurately, wheat grain images are preprocessed through different preprocessing methods.

First the image will be converted into gray scale image. Generally, the image will pre-processed using a smoothing filter that include operations which can enhance images and remove noise from an image [4]. The noise present into the input image is removed by using median filter.

Conversion of a gray-scale image into a binary image is called as thresholding or binarization. The output binary image has values of 1 (white) for all pixels in the input image and 0 (black) for all other pixels. Thresholding separate the wheat kernel from the background by converting the image into binary image. Thresholding is also used for segmentation which subdivides an image into different parts or objects i.e. the objects of interest are isolated from their background. Otsu method is used for binarization. Due to some defects on the surface of wheat grains or over drying of the wheat grain, there exist holes. These holes were filled with neighboring pixels.

C. Edge Detection Technique

Edge detection locates areas with strong intensity contrasts. Edge detection is based on recognition of edges by different edge operators. Lack of continuity in grey level, color, texture, etc. is detected by edge operators. For detecting edges of wheat kernels, canny edge detection technique is used.

Canny edge detector is a most selected detector which gives filtered image. The edges in the image are marked only once and false edges are not created because of the noise in the image because canny method has good localization and detection with minimum response. It has the capability to detect weak edges and is significantly strong. Canny edge detector distinguishes the edges by locating the local minima and maxima of the gradient of the intensity function. The advantage of Canny edge detection is that, the detected edges are thick.

D. Feature Extraction

The qualitative and quantitative information is extracted from the objects to be analyzed in the images in feature extraction. These extracted attributes are called features. The various features that are extracted are color features, morphological features and texture features because their structural forms like shape and size and also their visual color differences identify wheat grains of different grades by human vision in the traditional system. Hence, the classification system proposed is based on morphological, color and textural features.

Morphology is one of the most essential features that can be useful in recognizing various objects. The word morphology denotes visual and shape characteristics of an object. Here, four morphological features are extracted. They are area, major axis length, minor axis length and perimeter of the wheat grain.

Color is an important feature for image representation. Color Features are extracted using color moment, color histogram, and Color autocorrelogram. The mean value of R, the mean value of G, the mean value of B, the standard deviation values of R,G,B, color autocorrelogram, hsvcolor histogram are calculated in an image. The color histogram represents the color content of an image. Color auto correlograms when used as feature vectors, they make color histograms and color co-occurrence matrices exceptionally good [11].

Texture gives us information on structural arrangement of surfaces and intensity variation of objects on the image. They are suitable for visualizing pattern and surface properties of images [12]. Mean Squared Energy, Mean Amplitude features are calculated using Gabor wavelet and Wavelet Moments i.e. Mean coefficient, Standard coefficient are calculated by using Wavelet Transform.

E. Classification

The grading of wheat is determined by classifying the wheat into different grades according to their quality. Support Vector Machine and Naive Bayes Classifiers are used for classification of wheat into grade classes.

i) Support Vector Machine

Support Vector Machine is supervised learning classifier which is used for classification. Fundamentally SVM is a binary linear classifier in kernel leading feature space. It is expressed systematically as a weighted combination of kernel functions on training examples. The inner product of two vectors in linear or nonlinear feature space is represented by the kernel function. In a high dimensional space, a support vector machine creates a hyper plane or set of hyper planes, which can used for the purpose of classification or other tasks [13]. Though, SVM's have several good properties than other classifiers, which includes in general faster training, configuration is easy and satisfying theoretical properties. However, a main drawback of SVMs is that, theoretically, they were initially developed to resolve binary classification problems. Multi-class classifiers are normally constructed by combining several binary classifiers.

ii) Naive Bayes Classifier

Naive Bayes classifier is a classifier which is based on probability distribution. It classifies an object into the class to which it is probably to fit based on the observed features. It results from applying Bayes Theorem with independent assumptions between the features.

Simply, a Naive Bayes classifier considers that the value of a particular feature is not associated to the presence or absence of any other feature, given the class variable. Naive Bayes does quite well when the training data does not include all possibilities so it can be very good with low amounts of data. The Bayesian classification approach is described as follows [10].

Assume that there are N classes C_1, C_2 . . . C_N, and an unfamiliar pattern x in a d-dimensional feature space x= [x_1, x_2, . . ., x_d] . Compute the probability of belongingness of the pattern x to each class C_i **,** i = 1, 2, . . . , N**.** The pattern is classified to the class C_k if probability of its belongingness to C_k is a maximum. While classifying a pattern based on Bayesian classification, we distinguish two kinds of probabilities. They are priori probability and posteriori probability [32]. The priori probability denotes the probability that the pattern should fit in to a class, say C_k**,** based on the prior belief or evidence or knowledge. In cases where there exists no prior knowledge about the class membership of the pattern, usually a uniform distribution is used to model it. The posteriori probability $P(C_i/x)$, on the other hand, indicates the final probability of belongingness of the pattern x to a class C_i .The posteriori probability is computed based on the feature vector of the pattern, class conditional probability density functions $P(x/\ C_i)$ for each class C_i and priori probability $P(C_i)$ of each class C_i.
Bayesian classification states that the posteriori probability of a pattern belonging to a class C_k is given by, C_i

$$P\left(\frac{C_k}{x}\right) = \frac{P\left(\frac{x}{C_k}\right)P(C_k)}{\sum_{i=1}^{N}\left(P\left(\frac{x}{Ci}\right)P(Ci)\right)}$$

The denominator $\sum_i^N P\left({}^{x}/_{Ci}\right)P(Ci)$ in the above equation is the scaling term which provide the normalized value of the posteriori probability that the pattern x belongs to class C_i. Hence, x belongs to class *Cp* when $P(C_p|x) = \max\{P(C_1|x), P(C_2|x), \ldots, P(C_N|x)\}$.

IV. RESULTS

Matlab is the tool used for implementation. Image Processing is used for determining quality of wheat grains. In-order to train the classifiers, a set of training wheat grains was required and the grades of the wheat grains were predefined. Total 1300 wheat grains were taken for training and testing. These wheat grains were predefined into grades.

There are two basic phases of classification. They are training and testing phases. In the training phase, data is repeatedly presented to the classifier, while weights are updated to obtain a desired response. In testing phase, the trained system is applied to data that it has never seen to check the performance of the classification. Hence, we need to design the classifier by partitioning the total data set into training and testing data set. From the total of 1300 data sets, 650 were used for training and 650 were used for testing. In this study, the input features to the classifiers are the morphological, color and textural features and there were five output classes because i.e. Grade1, Grade2, Grade3, Grade4, Grade5.

A) Result of Support Vector Machine

The summary result of correctly classifying wheat grain into specific grade by SVM is shown in Table 1.

Table 1: Summary Result by SVM

	By SVM		
Class (Grade)	No of Images Tested	No of Images Classified Correctly	No of Images Classified Incorrect
Grade 1	130	125	5
Grade 2	130	119	11
Grade 3	130	120	10
Grade 4	130	126	4
Grade 5	130	124	6

The performance result of classifying wheat grain into grade by SVM is showed in Table 3. The classification accuracy of wheat for Grade1, Grade2, Grade3, Grade4, and Grade5 is 96.15%, 91.53%, 92.30%, 96.92%, and 95.38% respectively is shown in Fig. 3. The overall accuracy is 94.45% is shown in Fig. 4.

B) Result of Naive Bayes Classifier

Bayesian classifier is a probabistic classifier i.e. it is based on probability distribution. It classifies an object into the class to which it is most likely to belong based on the observed features. Hence, there were five predefined grade classes. The summary result of correctly classifying wheat grain into specific grade by Naive Bayes Classifier is shown in Table 2.

Table 2: Summary result by Naive Bayes Classifier

	By Naive Bayes Classifier		
Class (Grade)	No of Images Tested	No of Images Classified Correctly	No of Images Classified Incorrect
Grade 1	130	121	9
Grade 2	130	119	11
Grade 3	130	118	12
Grade 4	130	122	8
Grade 5	130	123	7

The result of classifying wheat grain into grade by Naive Bayes Classifier is showed in Table 3. The classification accuracy of wheat for Grade1, Grade2, Grade3, Grade4, Grade5 is 92.30%, 91.53%, 90.76%, 93.84%, 94.61% respectively is shown in Fig. 3. The overall accuracy by Naive Bayes Classifier is 92.60%. is shown in Fig. 4.

C) Comparative Result for SVM and Naive Bayes Classifier

The comparative result for SVM and Naive Bayes Classifier for classifying grain correctly in Grade classes is shown in Table 3 and Fig. 3. The overall accuracy result for SVM and Naive Bayes Classifier is shown in Fig. 4. The accuracy for Grade 1, Grade 3, Grade 4, and Grade 5 is greater for SVM than Naive Bayes classifier and for Grade 2, the SVM and Naive Bayes Classifier classifies with same accuracy i.e. 91.53%. The overall accuracy for SVM is better than Naive Bayes Classification.

Table 3: Grade wise and Overall Accuracy Result for SVM and Naive Bayes Classifier

Classifier	Grade1	Grade 2	Grade 3	Grade 4	Grade 5	Over-all Accuracy
SVM	96.15 %	91.53 %	92.30 %	96.92 %	95.38 %	94.45 %
Naive Bayes Classifier	92.30 %	91.53 %	90.76 %	93.84 %	94.61 %	92.60 %

ISSN: 2321-8169
5322 - 5327

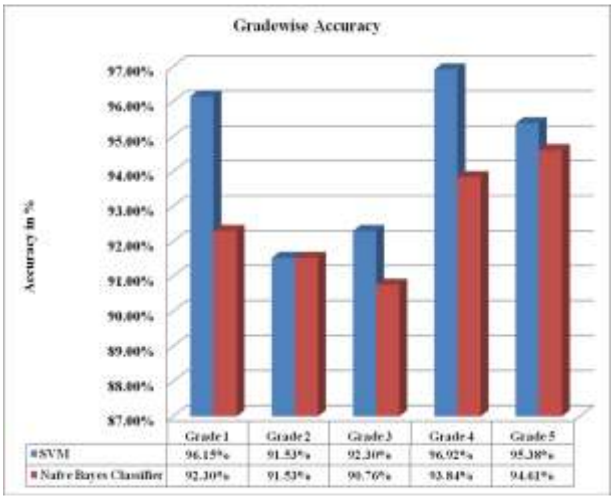

Fig. 3: Grade-wise Accuracy for SVM and Naive Bayes Classifier

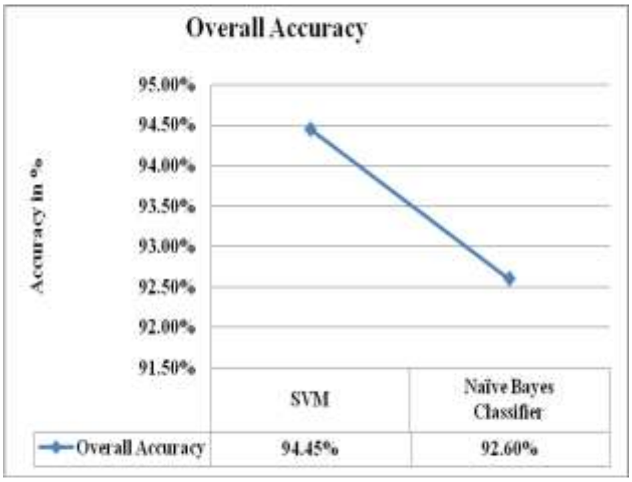

Fig. 4: Overall Accuracy

V. DISCUSSION

The experiments are conducted by using same feature sets as input to both SVM and Naive Bayes Classifier. Then, the experiment results are compared the performance of the SVM and Naive Bayes classifier. No standard database was available so, created the dataset containing total 1300 images of wheat grain. The 50% dataset i.e. 650 images used for training and 50% dataset i.e. 650 images used for testing. The accuracy of classification varies differently for different classifiers The morphological, color and textural features are extracted and same feature set is given to Support Vector Machine and Naive Bayes Classifier for classification. The performance result i.e. accuracy of each grade for both classifiers i.e. for Support Vector Machine and Naive Bayes Classifier is shown in Table 3 and in Fig. 3. The SVM perform better than Naive Bayes Classifier for Grade 1, Grade 3, Grade 4, and Grade 5 while the classification result for Grade 2 is same for both SVM and Naive Bayes Classifier. The overall accuracy of SVM and Naive Bayes is 94.45% and 92.60% respectively. The Overall accuracy for classifying wheat grains in to grade by SVM is better than Naive Bayes Classifier. So, the performance of SVM is better than Naive Bayes Classifier.

VI. CONCLUSION

Quality of grain is an important requirement for today's market, to protect the consumers from poor quality products. The system identifies the quality of wheat grains by distributing them into grade classes. There is no standard dataset of wheat is available so created own dataset for training and testing. The classification and grading is based on the appearance features such as morphological and color features and also based on texture features. Same morphological, color and texture feature set is used for both the classifier. The SVM and Naive Bayes classifiers are used for classifying the wheat grain into grade classes. The accuracy of classification varies differently for different classifiers. The result shows that the overall accuracy of SVM is better than Naive Bayes Classifier i.e. 94.45%, 92.60% respectively. So, the performance of SVM is better than Naive Bayes Classifier. The present work can be extended for other food grains also the accuracy rate can be increased by extracting other features. Feature selection can be applied for selecting a subset of features ideally necessary to perform the classification task from a larger set of candidate features for better accuracy.

REFERENCES

[1] Neelamegam. P, Abirami. S, Vishnu Priya. K, Rubalya Valantina. S, "Analysis of rice granules using Image Processing and Neural Network", *IEEE Conference on Information and Communication Technologies (ICT 2013)* ,2013.

[2] Alireza Pazoki and Zohreh Pazoki,"Classification system for rain fed wheat grain cultivars using artificial neural network", *African Journal of Biotechnology,* Vol. 10(41), pp. 8031-8038, August 2011.

[3] Alireza khoshroo, Arman Arefi, Asad Masoumiasl and Gholam-hossein Jowkar, "Classification of Wheat Cultivars Using Image Processing and Artificial Neural Networks", *Agricultural Communications,* pp. 17-22, 2014.

[4] Harpreet Kaur,Baljit Singh , "Classification and grading Wheat Using Multi-class SVM", *International Journal of Science and research publication*, vol. 3, 2013.

[5] Nandini Sidnal, "Grading And Quality Testing Of Food Grains Using Neural Network", *IJRET: International Journal of Research in Engineering and Technolog* , Vol. 02 Issue: 11 , November 2013.

[6] Sanjivani Shantaiya, Mrs. Uzma Ansari, "Identification Of Food Grains And Its Quality Using Pattern Classification," *International Journal of Computer & Communication Technology*, Vol. 2, 2010.

[7] Meesha Punn, Nidhi Bhalla, "Classification of Wheat Grains Using Machine Algorithms", *International Journal of Science and Research (IJSR),* Vol. 2 Issue 8, August 2013.

[8] Xian-Zhong Han, Ke-Jian Wang, Yingchun Yuan, Chen Chen, and Liang Liang, "Research on Grading Detection of the Wheat Seeds", *Scientific World Journal*, 2014.

[9] Megha R. Siddagangappa, A. H. Kulkarni, "Classification and Quality Analysis of Food Grains",

IOSR Journal of Computer Engineering (IOSR-JCE), Vol. 16, Issue 4, Ver. III, July - August 2014.

[10] Tinku Acharya, Ajoy K. Ray, "Image Processing Principles and Applications, Jhon Wiley", 2005.

[11] Robson Barcellos, Rogério Oliani Saranz, Luciana Tarlá Lorenzi, Adilson Gonzaga, "Content Based Image Retrieval Using Color Autocorrelograms in HSV Color Space".

[12] Ebrahim Ebrahimi , Kaveh Mollazade , Sirwan Babaei , " Toward an automatic wheat purity measuring device: A machine vision-based neural networks-assisted imperialist competitive algorithm approach", *Elsevier Ltd., Measurements55 (2014)*, pp.196–205, 2014.

[13] Sanghamitra Mohanty, Himadri Nandini Das Bebartta, "Performance Comparison of SVM and *K*-NN for Oriya Character Recognition", *International Journal of Advanced Computer Science and Applications (IJACSA) Special Issue on Image Processing and Analysis*, 2012.