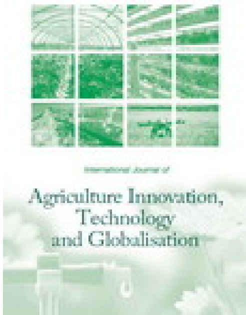


International Journal of Agriculture Innovation, Technology and Globalisation

ISSN online: 2516-1970 - ISSN print: 2516-1962
https://www.inderscience.com/ijaitg

Machine learning-based approach for degree of milling analysis of Indian rice variety

S. Harini, Saritha Chakrasali, G.N. Krishnamurthy

DOI: 10.1504/IJAITG.2023.10058269

Article History:

Received:	24 February 2023
Last revised:	27 February 2023
Accepted:	29 March 2023
Published online:	28 November 2023

Machine learning-based approach for degree of milling analysis of Indian rice variety

S. Harini*

B.M.S. College of Engineering,
Bull Temple Rd, Basavanagudi,
Bengaluru, Karnataka 560019, India
Email: hharinis8@gmail.com
*Corresponding author

Saritha Chakrasali and G.N. Krishnamurthy

BNM Institute of Technology,
12th Main Road, 27th Cross, Banashankari Stage II,
Banashankari, Bengaluru, Karnataka 560070, India
Email: saritha.chakrasali@gmail.com
Email: bnmitprincipal@gmail.com

Abstract: Image processing and machine learning has a wide application in the field of agriculture and food industry. This is because of the non-destructive evaluation process, performance and low cost compared to manual methods. Analysing grain quality manually is laborious and also subjective. It depends on the knowledge and experience of the experts. Rice is one of the staple food grains in major countries of the world. India being one of the top most exporters of rice grains, the quality analysis is very crucial. The food industry and consumers suffer from the lack of a fast, automated solution for identifying the quality of grains. To address this problem, this work proposes a machine learning-based solution for automatic analysis of quality of rice grains using degree of milling (DOM). Various machine learning algorithms are used for the analysis. A noticeable result is obtained for SVM, KNN, decision tree and CNN algorithms with an accuracy of 96%, 90%, 88% and 100% accuracy, respectively.

Keywords: convolution neural network; CNN; SVM; decision tree; KNN; Indian rice; degree of milling; DOM.

Reference to this paper should be made as follows: Harini, S., Chakrasali, S. and Krishnamurthy, G.N. (2023) 'Machine learning-based approach for degree of milling analysis of Indian rice variety', *Int. J. Agriculture Innovation, Technology and Globalisation*, Vol. 3, No. 2, pp.177–192.

Biographical notes: S. Harini is currently working as an Assistant Professor at B.M.S. College of Engineering, Bengaluru and also a research scholar at B.N.M. Institute of Technology in the research area image processing and machine learning. She has over ten years of teaching experience. She has completed her Bachelor's degree in 2010, from Visvesvaraya Technological University and MTech degree in 2012, Jain University, India. She has served as an Assistant Professor in Sapthagiri College of Engineering and B.N.M. Institute of Technology, Bengaluru. She has handled various subjects like machine learning, natural language processing, data structures, operating systems and theoretical foundations of computation. She has published over six papers in journals and conferences.

Saritha Chakrasali is a Professor at B.N.M. Institute of Technology. She has over 21 years of teaching experience. She has completed her Bachelor's degree in 1995 from Bangalore University, MTech degree in 2000 from AAIDU, India and is been awarded PhD in 2014 from Madurai Kamaraj University, India. She has served at BNM Institute of Technology as a Lecturer, Assistant Professor, Associate Professor and Professor. She has handled various courses like cryptography and network security/ information and network security, Unix and shell programming, Unix systems programming, discrete mathematical structures, C# and .Net, Advanced Algorithm, Java and J2EE, design patterns, agile technologies and artificial intelligence. Her area of specialisation is image processing. She has published papers in over 20 journals and conferences.

G.N. Krishnamurthy is the Principal and also a Professor at B.N.M. Institute of Technology. He has over 22 years of teaching experience. He has completed his Bachelor's degree in 1996 from Kuvempu University, MTech in 2000 and is been awarded PhD in 2009 from Visvesvaraya Technological University. He has served as a Faculty in NIIT, Davangere, Assistant Professor in Bapuji Institute of Engineering and Technology, Davangere and Professor at Cambridge Institute of Technology and BNM Institute of Technology, Bengaluru. His area of specialisation is cryptography-block ciphers design and performance analysis. He has handled various courses like agile technologies – principles and practice, wireless and mobile networks, wireless sensor networks, advanced computer architecture, computer systems design and architecture, operating systems and computer system performance analysis. He has published over 22 papers in journals and over six in conferences.

1 Introduction

India's agriculture is includes growing many crops, which includes most staple food staples like rice and wheat. Along with rice and wheat, farmers also grow tea, coffee, pulses, spices and non-food items as cotton, rubber and jute (*Nations Encyclopedia*, 2020). The nutritional value is high in cereal grains; hence, it is good for human population. Rice is one of the leading food crops of the world and is produced in all continents. India has become world's no. 1 exporter of rice since 2011 exporting around 16 million tonnes of rice by growing in around 43 million hectares (m ha) in the country-largest acreage in the world (ICAR-IIRR, 2020). As one of the primary centres of origin of *Oryza sativa* (commonly known as Asian rice), India has rich and diverse genetic wealth of rice.

Grain quality always refers to the quality of grain. The quality required depends on the use or purpose of the grain. The overall quality of grain is affected by various factors. It includes, weather condition, soil type, growing practices, time and type of harvesting, postharvest handling, storage management and transportation practices (Siddagangappa and Kulkarni, 2014). Measurement of grain quality on the commodity crops such as rice, wheat, barley, corn, maize is a wide research area. In the units which handles grain, the quality refers to morphological characteristics like shape, size, moisture content and kernel hardness. The quality also includes the percentage of visual attributes such as the presence of damaged, discoloured kernels, infested kernels and foreign materials. Acceptable grain quality ensure that the grain is free from adulterants and components that cause health hazards (Xu et al., 2010).

In spite of being a tedious, measuring the individual kernel's is very important for the qualitative analysis. Manual method of analysing the grains is tedious and time consuming. This method is complicated and has chances of having error since it is subjective. It totally depends on the knowledge of the expert who is analysing the sample. Hence, to solve such kind of problems, the machine-based automatic solutions are discovered (Prajapati and Patel, 2013). This gives a more accurate and qualitative results compared to manual method of grain quality inspection. Most of the recent works have used image processing to process the grain quality which reduces the job of the grain quality assessment (Hobson et al., 2007).

The food industry and research institutes are aiming to avoid distribution of sub-standard rice by finding a solution using computer vision and artificial intelligence techniques. Common physical properties of rice are size, shape, colour, uniformity, and general appearance. In general, the degree of milling (DOM) and percentage of broken kernels (PBK) are used as a quality assessing factor of milled rice (Zareiforoush et al., 2015). As limited amount of research is carried out in Indian rice quality analysis on automating the classification of food grains, this work focuses on automating the quality analysis of rice.

The DOM is expressed as weight percentage of bran removed from brown rice (Wadsworth et al., 1991). To determine the head rice yield, insect infestation, sensory quality and start gelatinisation DOM is very good parameter (Sun and Siebenmorgen, 1993; McGaughey, 1970; Champagne et al., 1990; Piggrott et al., 1991). The grains with different levels of DOM also helps in finding the nutritional value and commercial value of the sample. Vellupillai and Pandey (1987) reported that 65%–73% of the bran was removed in the first 20s of milling. As milling time increases, head rice yield decreases and DOM increases (Andrews et al., 1992).

After harvesting and drying, the paddy is subjected to the primary milling operation which includes de-husking as well as the removal of bran layers (polishing). Brown rice is milled to produce more visually appealing white rice. This rice which is obtained after milling is called raw rice. Another process through which rice is obtained after milling is called 'parboiling rice'. Nearly 60% of the total rice produced in India is subjected to parboiling.

In this work, analysis of quality of rice grains is performed using DOM as it is one of the major factors in food industry and for consumers. In both the cases, the analysis is carried out manually which could lead to human errors. This shows the need for machine learning-based automated solution for analysis of Indian rice varieties using DOM.

2 Related work

Many researchers have worked in providing the automated solution for analysis of rice grain quality. S. Mittal et.al, have proposed a Non-destructive image processing based system for assessment of rice quality and defects for classification according to inferred commercial value. This work is found out a cost-effective and automatic image processing based solution.

This is used further for classifying rice samples into different types based on their commercial value. This is a non-destructively prototype and works in real-time. Accuracy achieved is 93%. The advantage is that it consistently classifies the rice sample for different degrees of milling leading to better accuracy (Mittal et al., 2019).

Kiratiratanapruk and Sinthupinyo (011) classify defects of corn seed in more than ten categories and extracted colour, texture feature. They used support vector machine (SVM) as type classifier. Accuracy of the system is 96.5% for normal seed type and 86.5% in the case of defect seed (group) types.

Zareiforoush et al. (2015) developed a decision-support system for qualitative grading of milled rice using fuzzy inference system (FIS) coupled with image processing technique. The author has used DOM and PBK as the quality indices. These were used for grading rice into different classes. They have taken experts help for authenticating the different grades for the rice samples which are considered. Then, images of the same samples were captured using a machine vision system. The information obtained from the sample image processing was transferred to FIS. They have used CCD camera to capture the images along with a uniform lighting source and a computer. The uniform lighting is provided with the help of LED strip lights. To get the clear images, the standard background is maintained. They have used black background which would help in further pre-processing stages (Zareiforoush et al., 2015).

Perdon et al. (2001) have proposed Classification of DOM into broad categories, i.e., from well milled to reasonably well milled to lightly milled and lastly under milled.

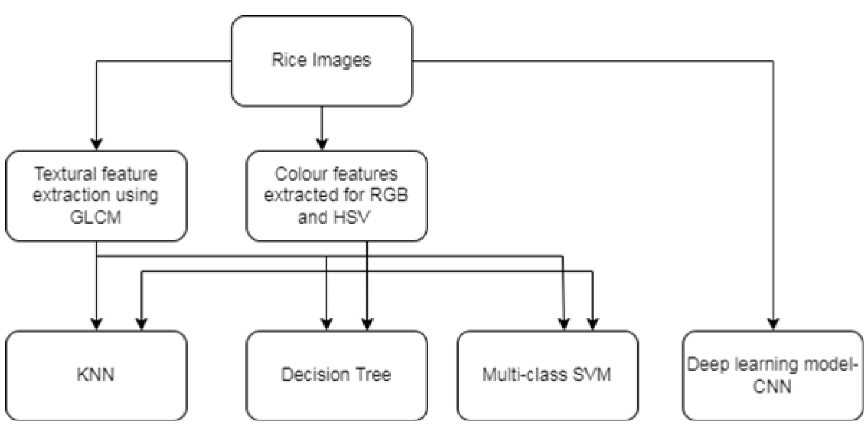
Archer and Siebenmorgen (1995) and Chen and Siebenmorgen (1997) have used milling metres/SDM (Satake DOM) which is an optical instrument to identify the DOM of the rice samples. This uses both reflectance and transmittance measurement to quantify DOM level on scale between 0–199. Level zero on SDM depicts brown rice and 199 represents fully milled white rice.

Though the above work is non-destructive, it is totally dependents on the instrument or equipments. Hence, a computer vision based machine learning solution would be a cost effective and less time consuming compared to the previous approaches.

3 Methodology

The rice quality is analysed using DOM by applying various machine learning algorithms. The architecture and algorithms used on various features are of the rice samples are depicted in Figure 1.

Figure 1 The different machine learning algorithms used to classify DOM



Textural and colour features extracted are used for first level of analysis using KNN, decision tree and SVM. The samples are further used on convolution neural network (CNN) model to check the improvement in results.

3.1 Sample collection

The rice samples are collected from a nearby commercial rice mill which processes tons of paddy per hour. The paddy undergoes many stages before polishing like, cleaning, de-husking, steaming, etc. This brown unpolished rice is further polished at three different stages. The third level polished rice will have no bran and will have very less nutritional value in it and has high commercial value because of the whiteness.

The laboratory analyses are carried out at Sri Bhagyalakshmi Agro Foods Pvt. Ltd. The evaluated rice variety, RNR 15048 (also called as 'Telangana Sona') commonly grown and consumed rice variety in southern part of India. This varierty is considered as long rice grain as per the standards given by CODEX STANDARD FOR RICE, Codex Standard 198-1995. This variety is cold tolerant rice with good cooking quality (Mandal, 2018). This rice variety is low in glycemic index (GI) and is good for diabetic patients when compared to Sona Masoori rice (https://kvk.icar.gov.in/API/Content/PPupload/k0331_1.pdf). The whiteness, transparency and milling degree of the sample rice grain are measured by using milling meter (SATAKE Corporation, Milling Metre MM1D) and the dimensions of the grains are measured by vernier calipher.

The readings taken by milling meter are by using rice samples randomly. The whiteness, transparency and milling degree increases as the milling degree/ level increase. The Milling meter shows highest values for whiteness, transparency and milling degree for degree 3 milled rice. These values support our work in identifying the rice quality based on DOM. Based on the values measured using verinier calipher, the expert at Sri Bhagyalakshmi Agro Foods Pvt. Ltd. have identified that the rice samples belongs to RNR 15048 variety.

4 Image acquisition

A standard experiment setup is developed and is fabricated for this purpose. The digital images of the rice kernel are acquired using the set up. As this work concentrates in finding a wieldy solution, a camera with 8 mega pixel resolutions and LED flash is used. Each image has a dimension of 2,448 × 3,264 pixels. An average of 30 to 40 rice kernels of each class (different levels of milling) is spread over the tray with black background. The kernels are separated manually from each other. This sample on a tray with black background is placed under the camera. This process is repeated over 25 times for each class of rice kernels. A uniform distance of 30 cm in maintained between the sample and the camera. The sample images captured by using the camera are transferred through a USB chord and are then stored for further pre-processing. The acquired images are then pre-processed and textural features are extracted.

The sample images captured are shown in Figure 2.

Figure 2(a) shows the images of rice kernel samples which have undergone stage one level of milling. Each sample has nearly 8 to 30 rice kernels. Similarly, Figures 2(b) and 2(c) shows the images of rice kernel samples which have undergone stage two and stage three level of milling, respectively.

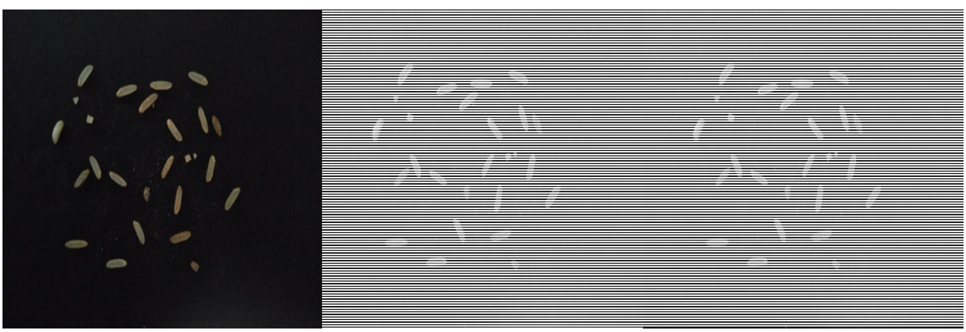
Figure 2 Samples of rice kernels with (a) degree 1 (stage 1) milling, (b) degree 2 (stage 2) milling and (c) degree 3 (stage 3) milling respectively

(a) (b) (c)

4.1 Image pre-processing module

The acquired images are converted to grey scale. The converted grey images are filtered using Gaussian blur to reduce the noise. The sample images after pre-processing are shown in Figure 3.

Figure 3 Stages of pre-processing, (a) original image (b) grey image (c) filtered image



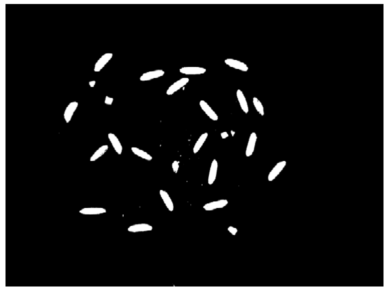
(a) (b) (c)

Figure 3 shows the results of applying pre-processing. Figure 3(a), is the original colour image, Figure 3(b), is the image after converting to grey scale and Figure 3(c), is the image obtained after filtering using Gaussian blurring.

4.2 Image segmentation and feature extraction

Thresholding is the simplest method of segmenting images (Kaur and Kaur, 2010). Thresholding creates a binary image based on setting a threshold value on the pixel intensity of the original image. In Otsu thresholding, the value for the threshold is chosen automatically without the intervention of the human. Here, always a bimodal image is considered. The generated histogram contains two peaks. Hence, a threshold value that lies in the middle of both the histogram peak values plays in any general conditions. Any point (x, y) for which f(x, y) > T (T means threshold), is called an object point, otherwise, the point is called a background point (Gonzalez and Woods, 2005). In this work, Otsu's thresholding is applied on the filtered image to separate rice kernels from the background. The obtained binary image is shown in Figure 4.

Figure 4 Binary image after applying Ostu's thresholding



Each rice kernel in the binary image obtained above is separated by finding contours. The curve joining all the continuous points (along the boundary), having same colour or intensity are simple called as contours. For analysing the shape and for detecting the objects, these contours play a very important role. After finding the contours of each rice kernel in the given sample, the contour area is considered. An ellipse is fitted for each rice kernel over the contour area obtained.

The contour area obtained is then used to draw a rectangle box bounding each rice kernel. This bounding box gives minimal up-right bounding rectangle for the specified point set. The bounding box for few rice kernels are shown in Figure 5.

Figure 5 The bounding box over rice kernels

Both the contours and bounding box are applied on the original image but not on the binary image. This is performed since we need to find the DOM based on the features of the original images.

Text features are used to measure the DOM of the rice kernels (Wan and Long, 2011; Gujjar and Siddappa, 2014). The grey-gradient co-occurrence matrix (GLCM) algorithm is used extract these textural features. Each pixel's distribution of grey-scale intensity and gradient is evaluated in matrix.

The gradient refers to the edge elements of the image and the grey value of each pixel represents the basis of an image.

The intensity value of rice kernels in a grey scale image is used for evaluating DOM along with RGB and HSV colour features. The textural features extracted using GLCM in this work are:

- width
- height
- aspect ratio
- contrast
- dissimilarity
- homogeneity
- energy
- correlation.

Table 1 List of extracted features on RGB and HSV colour channels

Features	*RGB colour space*	*HSV colour space*
Mean	Mean_RGB_R	Mean_HSV_H
	Mean_RGB_G	Mean_HSV_S
	Mean_RGB_B	Mean_HSV_V
Standard deviation	StdDev_RGB_R	StdDev_HSV_H
	StdDev_RGB_G	StdDev_HSV_S
	StdDev_RGB_B	StdDev_HSV_V
Skewness	Skewness_RGB_R	Skewness_HSV_H
	Skewness_RGB_G	Skewness_HSV_S
	Skewness_RGB_B	Skewness_HSV_V
Kurtosis	Kurtosis_RGB_R	Kurtosis_HSV_H
	Kurtosis_RGB_G	Kurtosis_HSV_S
	Kurtosis_RGB_B	Kurtosis_HSV_V
Wavelet decomposition	Haar_RGB_R_LL	Haar_HSV_H_LL
	Haar_RGB_G_LL	Haar_HSV_S_LL
	Haar_RGB_B_LL	Haar_HSV_V_LL
	Haar_RGB_R_LH	Haar_HSV_H_LH
	Haar_RGB_G_LH	Haar_HSV_S_LH
	Haar_RGB_B_LH	Haar_HSV_V_LH
	Haar_RGB_R_HL	Haar_HSV_H_HL
	Haar_RGB_G_HL	Haar_HSV_S_HL
	Haar_RGB_B_HL	Haar_HSV_V_HL
	Haar_RGB_R_HH	Haar_HSV_H_HH
	Haar_RGB_G_HH	Haar_HSV_S_HH
	Haar_RGB_B_HH	Haar_HSV_V_HH

The RGB and HSV colour features extracted in this work are listed in Table 1.

The overall of 56 features are extracted for each rice kernel, i.e., a total of 3,250 rice kernels are examined for this process.

4.3 Classification of DOM

The analysis and identification of DOM is performed using three major machine learning algorithms and CNN model. The 56 extracted features are used as input for the three supervised Machine learning algorithms namely: SVM, decision tree and K-nearest neighbours (KNN). The acquired images are directly fed as input for CNN model in this work.

Multi-class SVM is the majorly used classification model which uses multiple binary SVM's. The binary classifier here uses either one versus all (OVA) or one versus one (OVO) method to find the classification hyperplane. Either of these methods of binary SVM aims to find an optimal boundary to separate the samples of given two classes (Cortes and Vapnik, 1995). Finding this hyperplane is a challenge because in includes fitting a linear curve. If the given data is linearly separable, then linear classification is best suitable, otherwise the problem becomes more tougher. For such problem, SVM promotes various kernel trick. This is illustrated in Figure 6.

Figure 6 Nonlinear and linear hyperplane using SVM (see online version for colours)

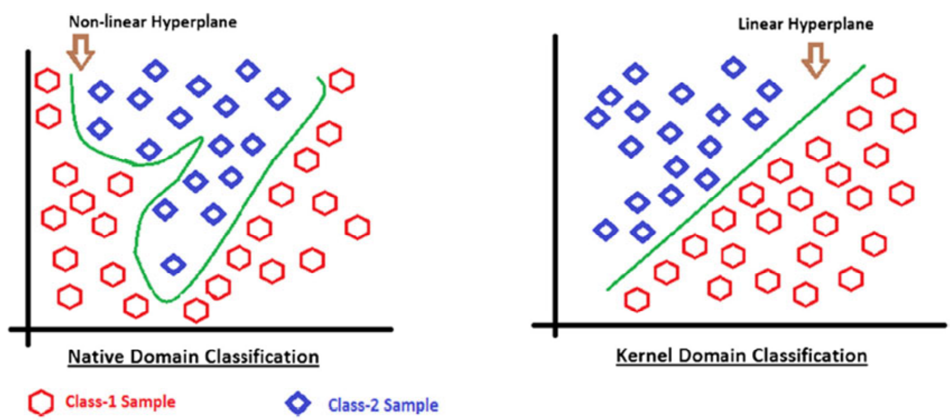


Figure 6 shows the hyperplane constructed for nonlinear and linear dataset using SVM.

The SVM algorithm is generally implemented using a kernel in real time. In linear SVM, the hyperplane is learnt by transforming the problem using some linear algebra. To address linear and nonlinear SVM model, it uses the following different kernel functions:

4.4 Linear kernel SVM

The linear kernel function is applicable when the data are linearly separable. The dot-product is called the kernel. The equation for linear kernel is given below:

$$K(x, xi) = sum(x * xi) \tag{1}$$

4.5 Polynomial kernel SVM

If the given dataset is not linear, polynomial kernel function could be used. The equation for polynomial kernel is given below:

$$k(x, xi) == sum(x * xi)^d \tag{2}$$

where the '*d*' represents the degree of the polynomial. This must be specified by hand to the learning algorithm.

4.6 Radial basis function (RBF) kernel SVM

Nonlinear data can also be transformed using radial kernel, which is widely used. The equation for RBF kernel is given below:

$$K(x, xi) = \exp\left(-gamma * sum\left(x - xi^2\right)\right) \tag{3}$$

where the main parameter is 'gamma'. This should be specified to the learning algorithm. The gamma parameter generally takes values between 0 to 1. Most works have considered the gamma value between 0.1 to 0.5 as a default.

Linear SVM is a parametric model, but the RBF kernel SVM is not a parametric model. The complexity of the RBF kernel grows with the size of the training set. Training such a complex RBF kernel SVM is expensive. This extends to an 'infinite' higher dimensional space. Furthermore, if more number of hyper parameters are to tune, the model selection is more expensive as well. And finally, it is much easier to over fit a complex model (Mittal et al., 2019).

The SVM with RBF kernel function uses the features, creates the non-linear combination and upgrades the dataset to higher-dimension.

4.7 Convolutional neural network

CNN is a deep learning model that is often used in areas such as natural language processing, voice recognition, image processing, and data sets that contain a high number of data (Champagne et al., 1990). The CNN model is widely used for end-to-end classification. The model takes the data as input to the CNN network, extracts features with layers after extracting the features, it learns and classifies the data. CNN is one of the deep learning models which follows five major layers: convolution layer, pooling layer, activation layer, fully connected layer and the classification layer (Lin et al., 2018).

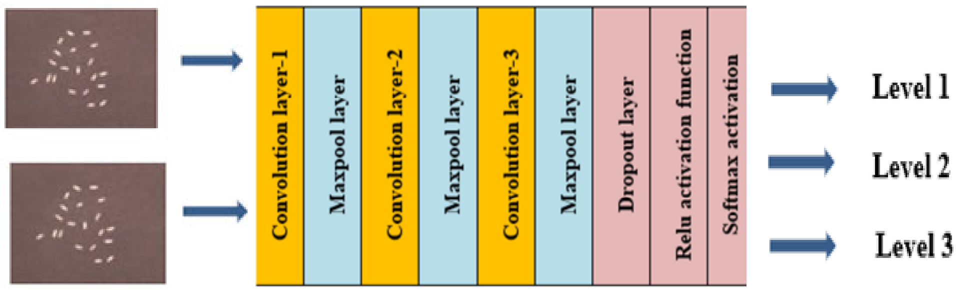
Generally, various filters are used in convolution layers. These filters are applied to extract the features from the each region of the image. This layers allows to vary the number of steps and filters to extract different number of features from the image. Later, the selection of best feature will be done to reduce the challenges of the learning model (Guo et al., 2015).

The pooling layer is used to reduce the complexity by reducing the number of data coming from the convolution layer. This layer should to adjusted to get optimal results in the performance of the classification (Scherer et al., 2010).

An activation layer is added to draw the data to certain range. This is followed by a fully connected layer, this reduces the features and so the level of the neural network. The

Softmax activation function is used to handle this multi-class problem (Sainath et al., 2013).

Figure 7 Customised CNN model used for classification of DOM (see online version for colours)



In this work, the customised CNN network is built and used for training. The model includes three convolutional layers followed by max-pooling layer. The images are resized to 224 × 224 and is fed as input to the first layer. The dropout layer is added after the third layer which is flattened and is followed by Softmax activation function which has three neurons to parse three different DOM. Figure 7 shows the CNN model used in this work.

5 Results

The machine learning algorithms like multi-class SVM, decision tree, K-nearest neighbours and CNN are used in this work and the results obtained are discussed below. The dataset of total 3,250 rice kernel images are considered. Among these 3,250 images, 80% of them that is 2,600 images are considered for training and 20%, i.e., 650 are used for testing. Initially, the experiments are carried out by considering only textural features extracted by GLCM. The work is carried out with the textural features dataset with normalised or scaled data.

The results were obtained using the above machine learning models are discussed in Table 2. The dataset are used in both un-normalised and normalised form for the study.

The experiment has given noticeable results for multi-class SVM with RBF kernel function. The results show that an approximate of 92% accuracy is obtained for both selected and complete set of GLCM features. The overall accuracy of 92.03% is achievement for non-overlapping normalised test datasets with all eight features of rice kernels.

The same environment is trailed for the extended feature set of 56 (GLCM, RGB and HSV) on the normalised dataset. The results obtained with multi-class SVM with linear, polynomial, RBF kernel function, KNN and decision tree for the normalised dataset is given in Table 3.

Table 3 results shows that multi-class SVM with RBF and polynomial kernel functions gives highest accuracy of 97% among the other algorithms.

Table 2 Performance of multi-class SVM, decision tree and KNN on datasets with only GLCM features

Algorithms	*Accuracy considering all 8 un-normalised features*	*Accuracy considering all 8 normalised features*	*Accuracy considering only contrast, dissimilarity, energy, homogeneity and correlation un-normalised features*	*Accuracy considering only contrast, dissimilarity, energy, homogeneity and correlation normalised features*
SVM with linear kernel function	72.88%	87.06%	73.38%	87.81%
SVM with polynomial kernel function	82.33%	85.07%	87.31%	80.34%
SVM with RBF kernel function	61.94%	92.03%	75.87%	92.03%
Decision tree	80.09%	79.85%	80.09%	80.34%
KNN	55.22%	85.07%	64.17%	90.04%

Table 3 Accuracy table for multi-class SVM with linear, RBF and polynomial kernel functions, decision tree and KNN

Algorithms	*Accuracy with 5 features (contrast, dissimilarity, energy, homogeneity and correlation)*	*Accuracy with 8 features*	*Accuracy with 56 features*
SVM with linear kernel function	87.81%	87.06%	92%
SVM with polynomial kernel function	80.34%	85.07%	96.7%
SVM with RBF kernel function (c = 9 and gamma = 0.5)	92.03%	92.03%	96.2%
Decision tree	80.34%	79.85%	90%
KNN	90.04%	85.07%	88%

Figure 8 is the confusion matrix obtained for SVM with RBF kernel function. The performance matrix including accuracy, precision, recall and F1-score are calculated for SVM with polynomial function using equations (4), (5), (6) and (7) and the results are listed in Table 4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{7}$$

where *TP* – true positive, *FP* – false positive, *TN* – true negative and *FN* – false negative.

Figure 8 Confusion matrix of SVM with RBF kernel function for DOM classification (see online version for colours)

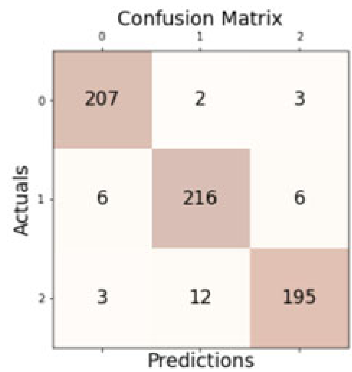
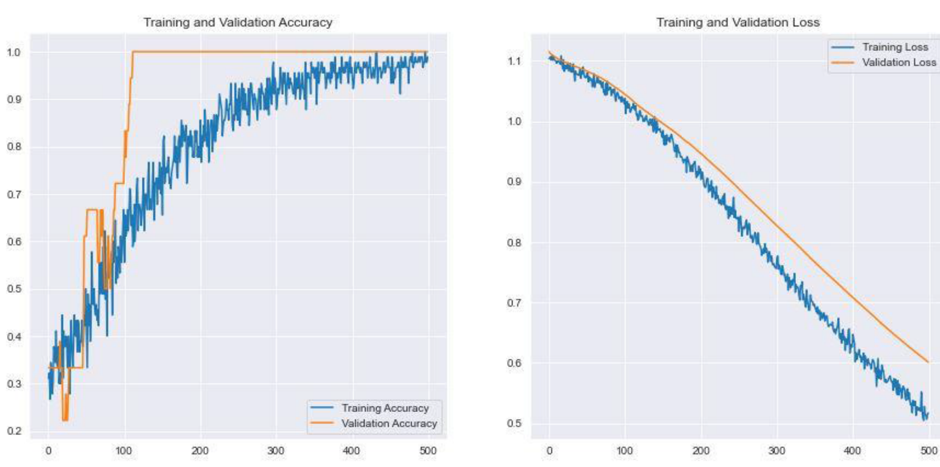


Table 4 The table of various performance matrix obtained from multi-class SVM with RBF kernel functions

DOM	*Accuracy*	*Precision*	*Recall*	*F1-score*
0	97.8%	95.83%	97.64%	96.72%
1	96%	93.9%	94.7%	94.29%
2	96.30%	95.58%	92.85%	94.20%
Average	*96.7%*	*95.10%*	*95.97%*	*95.07%*

Figure 9 Testing and validation accuracy/loss, (a) training and validation accuracy against no. of epochs (b) training and validation loss against no. of epochs (see online version for colours)



(a) (b)

From Table 4, it is observed that the SVM with polynomial kernel function and RBF kernel functions yield higher results compared to other algorithms.

The model is optimised using Adam optimiser and the model is training with varying number of iteration. The results of 100% accuracy are obtained using our customised CNN model shown in Figure 9.

The testing and validation accuracy and loss obtained from our CNN model in classifying the DOM is shown in the graph Figure 9. The graph in Figure 9(a) shows the increase of training and validation accuracy as the number of epochs increases. The validation accuracy has remained stable for epochs greater than 120 till 500. The graph in Figure 9(b) shows the decrease in training and validation loss as the number of epochs increases.

We can observe from the above figure that with the increase of iterations (epochs), the accuracy increases and the loss reduce.

The overall precision, recall, accuracy and F1-score for the classification of DOM using our customised CNN model is shown in Table 5.

The results in Table 5 shows that 100% accuracy, precision, recall and F1-score is obtained for the classification of DOM using our customised CNN model.

Table 5 The table of various performance matrix obtained from CNN model

DOM	*Accuracy*	*Precision*	*Recall*	*F1-score*
0	100%	100%	100%	100%
1	100%	100%	100%	100%
2	100%	100%	100%	100%
Average	*100%*	*100%*	*100%*	*100%*

6 Conclusions

An attempt has been made to analyse quality of rice based on DOM using SVM, decision tree, KNN and CNN. The work is carried out based on the images acquired using the developed standard setup. Textural features are extracted using GLCM from the images acquired and colour features are extracted from RGB and HSV colour space. Different supervised learning algorithms like multi-class SVM, decision tree and KNN are applied on the 56 extracted features. The customised CNN model is directed applied on the images collected. The work shows promising result of 100% for CNN model and good result for multi-class SVM. The following observations are made based on the results achieved using CNN and multi-class SVM:

- An accuracy of 100% is obtained for customised CNN model as it extracts all possible features from the image provided as the input. The model consists of three convolution layers each followed by a maxpooling layer and a dropout layer at the end. The Softmax activation function is used to parse the output to three different DOM.
- An accuracy of 97% is obtained using multi-class SVM for the normalised dataset with all 56 features both for RBF kernel function (with gamma = 1.0 and C = 9) and polynomial kernel function.

References

Andrews, S.B., Siebenimorgen, T.J. and Mauromoustako, M. (1992) 'Evaluation of the McGill No. 2 rice miller', *Cereal Chemistry*, Vol. 69, No. 1, pp.35–43.

Archer, T.R. and Siebenmorgen, T.J. (1995) 'Milling quality as affected by brown rice temperature', *Cer. Chem*., Vol. 72, No. 3, pp.304–307.

Champagne, E.T., Marshall, W.E. and Goynes, W.R. (1990) 'Effect of degree of milling and lipid removal on starch gelatinization in the brown rice kernel', *Cereal Chemistry*, Vol. 67, No. 6, pp.570–574.

Chen, H. and Siebenmorgen, T.J. (1997) 'Effect of rice kernel thickness on degree of milling and associated optical measurements', *Cer. Chem*., Vol. 74, No. 6, pp.821–825.

Cortes, C. and Vapnik, V. (1995) 'Support-vector networks', *Machine Learning*, Vol. 20, No. 3, pp.273–297.

Gonzalez, R.C. and Woods, R.E. (2005) *Digital Image Processing*, 2nd ed., 779pp, Prentice-Hall.

Gujjar, H.S. and Siddappa, M. (2014) 'Recognition and classification of different types of food grains and detection of foreign bodies using neural networks', *International Journal of Computer Applications (0975-8887), International Conference on Information and Communication Technologies (ICICT – 2014)*.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. and Lew, M.S. (2015) 'Deep learning for visual understanding: a review', *Neurocomputing*, Vol. 187, pp.27–48, https://doi.org/10.1016/j.neucom.2015.09.116.

Hobson, D.M., Carter, R.M. and Yan, Y. (2007) 'Characterisation and identification of rice grains through digital image analysis', *IEEE Instrumentation and Measurement Technology Conference – IMTC*, No. 1-4244-0589.

ICAR-IIRR (2020) *ICAR-Indian Institute of Rice Research* [online] https://www.icar-iirr.org/index.php (accessed December 2022).

Kaur, D. and Kaur, Y. (2010) 'Various image segmentation techniques: a review', *International Journal of Computer Science and Mobile Computing*, May 2014, Vol. 3, No. 5, pp.809–814, ISSN 2320-088X.

Kiratiratanapruk, K. and Sinthupinyo, W. (2011) 'Color and texture for corn seed classification by machine vision', *International Symposium on Intelligent Signal Processing & Communication Systems (ISPACS)*, pp.1–5.

Lin, P., Li, X., Chen, Y. and He, Y. (2018) 'A deep convolutional neural network architecture for boosting image discrimination accuracy of rice species', *Food Bioprocess Technol*., Vol. 11, No. 4, pp.765–773, https://doi.org/10.1007/s11947-017-2050-9.

Mandal, D. (2018) 'Adaptive neuro-fuzzy inference system based grading of basmati rice grains using image processing technique', article in *Applied System Innovation*, DPI, Received: 10 April 2018; Accepted: 15 June 2018; Published: 20 June 2018.

McGaughey, W.H. (1970) 'Effect of degree of milling and variety on insect development in milled rice', *Journal of Economic Entomology*, 1 August, Vol. 63, No. 4, pp.1375–1376.

Mittal, S., Dutta, M.K. and Issac, A. (2019) 'Non-destructive image processing based system for assessment of rice quality and defects for classification according to inferred commercial value', *Measurement*, Vol. 148, No. 106969, 0263-2241/_, 2019 Elsevier Ltd. All rights reserved.

Nations Encyclopedia (2020) 'Own a farm, live a little', 8 January [online] https://www.nationsencyclopedia.com/economies/Asia-and-the-Pacific/IndiaAGRICULTURE.html (accessed January 2022).

Perdon, A.A., Siebenmorgen, T.J., Mauromoustakos, A., Griffin, V.K. and Johnson, E.R. (2001) 'Degree of milling effects on rice pasting properties', *Cer. Chem*., Vol. 78, No. 2, pp.205–209.

Piggrott, J.R., Morrison, W.R. and Clyne, J. (1991) 'Changes in lipids and in sensory attributes on storage of rice milled to different degrees', *International Journal Food Science Technology*, Vol. 26, No. 6, pp.615–628.

Prajapati, B.B. and Patel, S. (2013) 'Algorithmic approach to quality analysis of Indian basmati rice using digital image processing', *International Journal of Emerging Technology and Advanced Engineering*, Vol. 3, No. 2, pp.503–504, ISSN: 2250-2459.

Sainath, T.N., Mohamed, A-r., Kingsbury, B. and Ramabhadran, B. (2013) 'Deep convolutional neural networks for LVCSR', *IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, pp.8614–8618, https://doi.org/10.1109/ICASSP.2013.6639347.

Scherer, D., Müller, A. and Behnke, S. (2010) 'Evaluation of pooling operations in convolutional architectures for object recognition', *International Conference on Artificial Neural Networks*, Springer, pp.92–101, DOI: 10.1007/978-3-642-15825-4_10.

Siddagangappa, M.R. and Kulkarni, A.H. (2014) 'Classification and quality analysis of food grains', *IOSR Journal of Computer Engineering (IOSR-JCE)*, Vol. 16, No. 4, Ver. III, e-ISSN: 2278-0661, p-ISSN: 2278-8727.

Sun, H. and Siebenmorgen, T.J. (1993) 'Milling characteristics of various rough rice kernel thickness fractions', *Cereal Chemistry*, Vol. 70, No. 6, pp.727–733.

Vellupillai, L. and Pandey, J.P. (1987) *Color and Bran Removal in Rice Processing*, Manuscript 87-07-1342, Louisiana Agricultural Experimental Station, Baton Rouge, LA.

Wadsworth, J.I., Sequeira, D.J., Vellupillai, L. and Verma, L.R. (1991) *Rice Degree of Milling Measured by NIR*, ASAE Paper 91-6030 ASAE, St. Joseph, MI.

Wan, P. and Long, C. (2011) 'An inspection method of rice milling degree based on machine vision and gray-gradient co-occurrence matrix', in *Computer and Computing Technologies in Agriculture IV*, Springer, pp.195–202.

Xu, X., Fan, C., Lu, S-J. and Hou, L-L. (2010) 'The analysis of wheat appearance quality based on digital image processing', *Environmental Science and Information Application Technology (ESIAT), The 2nd Conference on Environmental Science and Information Application Technology*, pp.231–234.

Zareiforoush, H., Minaei, S., Alizadeh, M.R. and Banakar, A. (2015) 'A hybrid intelligenet approach based on computer vision and fuzzy logic for quality measurement of milled rice', *Measurement*, Vol. 66, pp.26–34, Elsevier.

Websites

https://kvk.icar.gov.in/API/Content/PPupload/k0331_1.pdf (accessed 25 April 2021).