International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 3, June 2025, pp. 3202~3211
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i3.pp3202-3211

Classification of morphologically similar Indian rice variety using machine learning algorithms

Harini Shadaksharappa[1], Saritha Chakrasali[2], Krishnamurthy Gorappa Ningappa[2]
[1]Department of Information Science and Engineering, B.M.S. College of Engineering, Bangalore, Karnataka, India
[2]Department of Information Science and Engineering, BNM Institute of Technology, Bangalore, Karnataka, India

Article Info

Article history:

Received Aug 19, 2024
Revised Mar 12, 2025
Accepted Mar 20, 2025

Keywords:

Morphologically similar
Indian rice variety
Convolutional neural network
Support vector machine
K-nearest neighbor
Decision tree
GLCM, RGB and HSV

ABSTRACT

India, among the agriculture-based economy grows wide variety of rice along with other crops. These varieties have different commercial values as they are different in their features. It becomes extremely challenging to classify rice varieties which have similar features but are different in their quality. This study considers four varieties of similar looking rice which conform to be Sona-Masuri. A total of 4180 images are considered to extract 56 features including textural, red, green, blue (RGB) and hue, saturation, value (HSV) color and wavelet decomposition. Machine learning (ML) models like support vector machine (SVM), K-nearest neighbor (KNN), decision tree (DT) and voting classifiers are developed for feature dataset and convolutional neural network (CNN) model for image dataset. The results obtained for every model are obtained using statistical methods and the results are expressed in a table for accuracy, precision, recall and F1-score. A classification accuracy of 72.48% is obtained for SVM using polynomial kernel trick by considering all 56 features. The customized CNN model is designed with three convolution layers has resulted in 97.13% of training accuracy and 87.5% of validation accuracy. Based on the results obtained, it is witnessed that the ML models employed in this study to classify rice types with similar appearances have practical applications.

Corresponding Author:

Harini Shadaksharappa
Department of Information Science and Engineering, B.M.S. College of Engineering
Bangalore - 560019, Karnataka, India
Email: hharinis8@gmail.com

1. INTRODUCTION

India is one of the largest rice exporters 2011 [1]. Being the major cultivators of rice, it is also being a staple food in India and in most of the other countries. India grows nearly 489 varieties of rice including native as well as hybrid [2].

Sona-masuri, a medium-slender rice variety is popular in the southern region of India, is also one of the most valuable commercial varieties of rice, alongside basmati. There are different rice qualities available in the market which are similar to sona-masuri but differ from commercial value. Fraudulent traders quote higher prices for low quality rice. This is a huge loss for both consumers and the food industry.

Identifying the right type of rice variety among the similar type is a challenge. From ages this quality analysis is performed by the human inspection method. This approach is truly based on the knowledge of the expert along with being time consuming and laborious. Hence in recent studies computer vision and machine learning (ML) based automated solution is used for the rice and other food grains classification [3]. The details of the approaches followed by researchers are encompassed in the next section.

The greatest of the research works in analysis of rice quality using image processing and machine learning. These works are limited to:

– Morphologically dissimilar variety of rice
– Limited to long slender rice variety- like Basmati
– Non-Indian variety of rice

The major objective of this research work is to develop a machine vision system capable of accurately classifying visually similar varieties of rice based on their extracted textural, morphological, and color features from images. This study addresses the challenge of distinguishing between rice varieties that look similar but differ in quality and commercial value, which is a significant issue in agricultural marketing and consumer protection.

Successful applications of machine vision in the field of agricultural product inspection have been observed, although a majority of these efforts are still in the research and development stage [4]. The exploration of image processing and ML in the domain of rice quality analysis has been ongoing for decades and has garnered considerable research attention. Harini *et al.* [5] have put forth a review detailing the methods employed in food grain quality analysis. The same authors have also contributed a research on the analysis of rice quality in terms of the other parameters by applying ML approach [6].

Research endeavors have grown rapidly in the past decade. In a study conducted by Koklu *et al.* [7], five distinct varieties of rice cultivated in Turkey were investigated. The study analyzes five rice varieties—Arborio, Basmati, Ipsala, Jasmine, and Karacadag—using morphological, color, and shape features. Remarkable classification accuracies were achieved: 99.87% with artificial neural network (ANN), 99.95% with deep neural network (DNN), and a perfect 100% with convolutional neural network (CNN).

Singh *et al.* [8] have introduced a model aimed at classifying four varieties of rice grain through the integration of image processing and ML algorithms. The methodology extracts color, texture, and wavelet features, using support vector machine (SVM), K-nearest neighbor (KNN), naïve Bayes (NB), and backpropagation neural network (BPNN) for classification analysis. BPNN achieves the highest accuracy of 96% across all feature sets, though the study lacks details on rice varieties and morphological similarities.

A work proposed by Asif *et al.* [9] classifies six different variety of rice and also aims in quality analysis of rice. The study used PCA and Canny edge detection for classification and extracted morphological features for rice quality analysis, achieving 92.3% and 89.5% accuracy, respectively. However, it does not discuss the morphological similarity of the rice varieties considered.

In the study conducted by Lin *et al.* [10], a deep convolutional neural network (DCNN) was used to classify rice kernel images, outperforming traditional methods like PHOG-KNN, PHOG-SVM, GIST-KNN, and GIST-SVM. The DCNN achieved a peak training accuracy of 99.4% with a batch size of 15 and 20 epochs. Cinar *et al.* [11] have emerged with a computerized vision-based system was developed to differentiate between two proprietary rice species using 3,810 rice grain images and seven morphological features. Lin *et al.* employed various ML models, achieving classification accuracies of LR - 93.02%, MLP - 92.86%, SVM - 92.83%, DT - 92.49%, RF - 92.39%, NB - 91.71%, and KNN - 88.58%.

Identifying and classifying visually similar food grains is challenging, but automated recognition systems help mitigate misrecognition. Gujjar *et al.* [12] proposed a method using texture, morphology, and color-based retrieval on Basmati rice images, employing image warping, discriminant analysis, and a back-propagation neural network to classify six rice varieties. They developed a digital image analysis algorithm to identify six different varieties of Basmati rice seeds commonly cultivated in India. Discriminate analysis was conducted using nine features each for color, morphological, and texture aspects. Additionally, a back-propagation neural network-based classifier was implemented for the identification of unknown grain types [13].

Kolkure *et al.* [14] have introduced an automated evaluation method was introduced for rice quality assessment using color and geometric features. A neural network model, trained on appearance features, effectively identified unknown grains and demonstrated promising results in quality evaluation. Srimulyani *et al.* [15] analyzed nine Indonesian rice varieties using images captured with a Canon D600. Their study found that combining six geometric features with color, texture, and shape improved performance, achieving 100% accuracy with the back-propagation algorithm and a 5.2% increase with learning vector quantization (LVQ).

The primary goal of this study is to create a machine vision system capable of classifying visually similar varieties of rice grain based on their extracted textural, morphological, and color features from images. The research focuses on developing an algorithm designed to categorize images of four similar-looking rice varieties using these extracted features. Several ML models are evaluated for their effectiveness in achieving this classification task. The methodology and the obtained results are further elaborated in the subsequent sections of this paper.

2. METHOD

Since there is a need for automated solutions, extensive research has been conducted over the past decade on the quality analysis of food grains using image processing and machine learning. Image processing involves several steps like Pre-processing, segmentation, transformation, feature extraction and feature analysis [3]. The process begins with image collection using a standard setup to ensure uniformity and minimize noise, followed by preprocessing to remove irrelevant information. Image segmentation, performed in the spatial or frequency domain, identifies meaningful regions and objects. Segmentation or mathematical transformations help extract essential features, which are further refined through filtering. These features, including geometric, histogram, and color properties, are then analyzed for classification. This study focuses on a particular type of rice grains. The initial phase involved in gathering samples of the rice and capturing pictures to establish dataset. The entire procedure employed for this investigation is outlined in Figure 1.

Four different types of rice shown in Figure 2, that are available in the market which have morphological similarities were gathered for quality analysis and assessed at M/s. Shri Bhagyalakshmi Agro. Food Pvt. Ltd. These varieties, comparable to Sona-Masuri and categorized as medium-sized rice, are designated by the names: i) B.R.T Daiwan – referred as Rawh (Figure 2(a)), ii) Grine world - referred as Rawl (Figure 2(b)), iii) Royal Sona - referred as Steamh (Figure 2(c)), and iv) Daiwan Sona - referred as Steaml (Figure 2(d))

The considered varieties of rice samples have two raw and two steam rice categories. The major challenge in this classification model is the similar features among these varieties. The underling challenge of this work is to make proper rice variety classification. This influences the commercial market of the grain. Figure 2 shows the sample images.

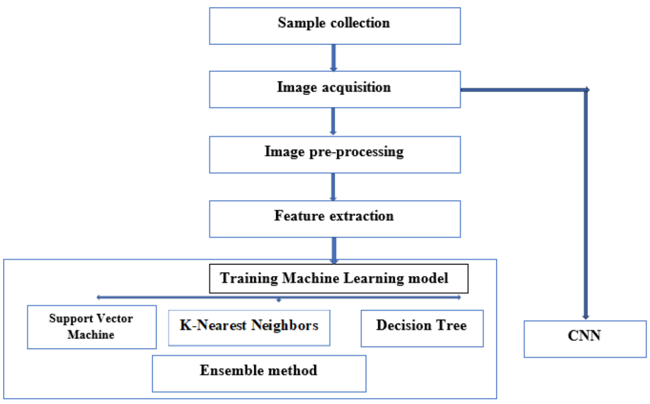

Figure 1. Method followed for classification of similar looking Indian rice grains

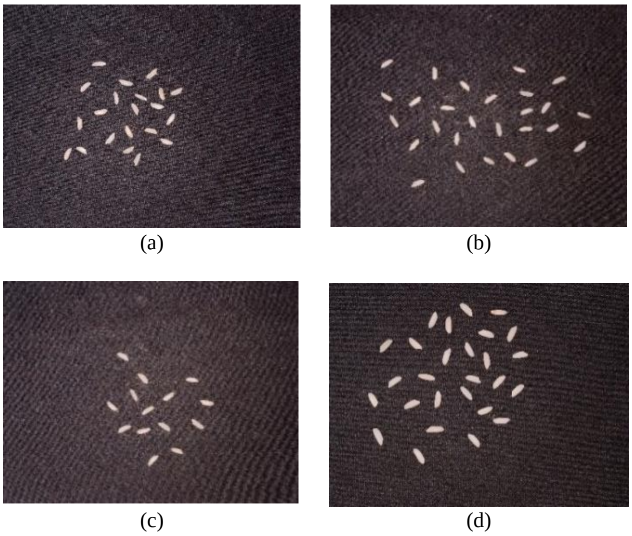

Figure 2. Sample images of (a) Rawh), (b) Rawl, (c) Steamh, and (d) Steaml

2.1. Dataset creation

A total of 4180 images (which include 1045 images of each variety) depicting rice kernels from four visually analogous rice varieties were acquired through a standard image acquisition model. This includes Each image that was captured using a mobile camera featuring 50MP resolution, 2.5X zoom, and flashlight in a standardized setup. The acquired images are channeled for initial steps which includes pre-processed and features extraction. A total of 56 features are extracted including textural, color and wavelet features are considered. The features are extracted using grey level co-occurrence matrix (GLCM), red, green, blue (RGB) and hue, saturation, value (HSV) color spaces. The features are shown in Table 1. The details of GLCM features are given in Table 2.

Table 1. Features extracted from the rice grain image

Feature Type	Features	Total number of features
GLCM	Width, Contrast, Dissimilarity Height, Aspect Ratio, Energy, Correlation and Homogeneity.	8
RGB and HSV Color	Skewness, Standard deviation, Kurtosis, Mean and Wavelet decomposition	48 (24 features each of RGB and HSV)

Table 2. Details of GLCM features [16]

Sl. No.	Feature	Explanation
1.	Contrast	Contrast is the distinction measured between top and bottom value of pixels adjacent in a given image. It has low spatial frequencies if its contrast value is low and is the concentration term around the major diagonal.
2.	Dissimilarity	Dissimilarity measures the local variations linearly in a given image.
3.	Homogeneity	It looks at bigger values when there is a lesser difference between the pair of grey tone elements. Homogeneity is sensitive to nearly diagonal elements. When the pixels in images are identical, the measure of homogeneity will be highest. This is the opposite of contrast.
4.	Energy	Energy is nothing but the square root of an angular second moment. Energy will bare higher values, when the window is orderly arranged.
5.	Correlation	Correlation gives linear dependencies among the gray tone of the image.

The above-mentioned texture features are calculated using the (1) to (5):

$$Contrast = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} P(x,y) * (x-y)^2 \tag{1}$$

$$Contrast = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} P(x,y) * (x-y)^2 \tag{2}$$

$$Dissimilarity = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} P(x,y) * |x-y| \tag{3}$$

$$Energy = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} P(x,y)^2 \tag{4}$$

$$Correlation = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} P(i,j)\frac{(x-\mu)(y-\mu)}{\sigma^2} \tag{5}$$

In the above equations: n - the number of gray levels. P(x,y) – for the given position x and y, this measures the gray-scale's normalized with the sum equal to unity. μ - the GLCM mean. σ^2 - the variance of the intensities of all reference pixels. Along with the above list of textural features, 48 color features are extracted. The list is shown in Table 3.

Table 3. List of color features

Color Space	Feature						
RGB		Std_	Skew_	Kurtosis_	HAAR_RLL	HAAR_RHH	HAAR_	HAAR_
	Mean_RGB_R	RGB_R	RGB_R	RGB_R	HAAR_RLH	HAAR_ GLL	GHL	BLH
	Mean_	Std_	Skew_	Kurtosis_	HAAR_	HAAR_	HAAR_	HAAR_
	RGB_G	RGB_G	RGB_G	RGB_G	RHL	GLH	GHH	BHL
	Mean_	Std_	Skew_	Kurtosis			HAAR_	HAAR_
	RGB_B	RGB_B	RGB_B	_RGB_B			BLL	BHH
HSV		Std_	Skew_	Kurtosis_	HAAR_	HAAR_RHH	HAAR_	HAAR_
	Mean_	HSV_H	HSV_R	HSV_R	RLL	HAAR_ GLL	GHL	BLH
	HSV_H	Std_	Skew_	Kurtosis_	HAAR_RLH	HAAR_	HAAR_	HAAR_
	Mean_ HSV_S	HSV_S	HSV_G	HSV_G	HAAR_	GLH	GHH	BHL
	Mean_	Std_	Skew_	Kurtosis_	RHL		HAAR_	HAAR_
	HSV_V	HSV_V	HSV_B	HSV_B			BLL	BHH

2.2. ML models

Once the features are extracted, the ML models are trained. The entire dataset includes features of 4180 rice kernels. The SVM, KNN and decision tree (DT) are used in this work. This dataset is classified into training and testing set with a ratio of 80:20. The performances of each of these methods are discussed in the result section of the paper.

2.2.1. SVM

SVM is a Supervised Learning algorithm used for classification and even for regression problems. This algorithm tries to find the best hyperplane which separates the two classes. SVM can handle non-linear data using “Kernel trick”. It uses three kernel functions namely: Linear, Polynomial, and Radial basis (RBF).

2.2.2. DT

Decision trees can perform both classification and regression tasks. A decision tree is a predictive model that makes decisions based on the input data, using a flowchart-like structure. The tree consists of branches, a root, leaves, and internal nodes. In this structure, each internal node tests an attribute, every branch represents an attribute value, and each leaf node signifies the final decision or prediction based on the provided dataset.

Attribute selection at internal is carried out either by calculating Information gain (IG) or Gini index. The formula for IG and Gini index is given in (6) and (7):

$$Information\ Gain = Entropy(S) - [(Weighted\ Avg) * Entropy\ (each\ feature) \tag{6}$$

$$Gini\ Index = 1 - \sum jPj2 \tag{7}$$

Entropy which is used in IG calculation is the amount of uncertainty in our dataset or measure of disorder. The calculation of Entropy is carried out using the formula (8):

$$\mathrm{E(S)} = -\mathrm{p}_{(+)}\mathrm{logp}_{(+)} - \mathrm{p}_{(-)}\mathrm{logp}_{(-)} \tag{8}$$

here, E= Entropy, S= Total number of samples, $\mathrm{p}_{(+)}$= probability of positive class,$\mathrm{p}_{(-)}$= probability of negative class

2.2.3. KNN

This algorithm considers the similarity between the available and new data. It then categorizes new data into the class which is in close proximity with available class KNN is a memory-based algorithm as is stores the available data in memory and assigns class to new data based on similarity. KNN calculates distance between data points and the new data using Euclidean distance formula and the new data is categorized to the nearest neighbor. The formula for calculating Euclidean distance is provided in (9). KNN is a non-parametric algorithm because it does not make any assumptions about the underlying data.

$$Euclidean\ distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{9}$$

where (x_1, y_1) and (x_2, y_2) are the co-ordinates of the two data points.

2.2.4. Ensemble method (EM)

Ensemble methods refer to a machine learning method that combines multiple base models to create an optimal predictive model. These ensembles are particularly useful in classification tasks, as they can achieve high accuracies by leveraging the diverse strengths of individual base models. Even if the accuracy of each base classifier is relatively low, the ensemble can excel if different base models misclassify distinct training examples. In this work, an ensemble model consisting of SVM, DT and KNN is used. The results obtained are discussed in the result section.

2.3. CNN

CNN is a model that makes use of deep learning for uses including natural language, speech recognition, picture processing, and datasets where large quantities of data exist [17]. End-to-end classification can broadly apply using the CNN model. Data as the input for the CNN model enters to have layers afterwards the extraction of the feature followed by its classification through which learning happens to categorize data. CNN is a deep learning model which has five key layers: one or more convolution layer, pooling, activation, fully connected and the classification layer [18] [10].

Convolution layers use various filters to extract features from different regions of an image, adjusting steps and filter counts to enhance feature extraction [19]. The pooling layer reduces data complexity from the convolution layer, requiring proper tuning for optimal classification. An activation layer captures non-linear relationships between input and output data, followed by a fully connected layer [20]. In multi-class classification, the Softmax activation function is commonly used for final predictions [21].

In this study, a customized CNN was developed and used for training. This neural network model consists of three convolutional layers, each one has a max-pooling layer following them. Images are scaled to 224×224 pixels before being processed by the first layer. After the third layer, a Dropout layer is applied, and the output is flattened and passed through a Softmax activation function with four neurons to classify four different types of rice. The CNN model used in this study is shown in Figure 3.

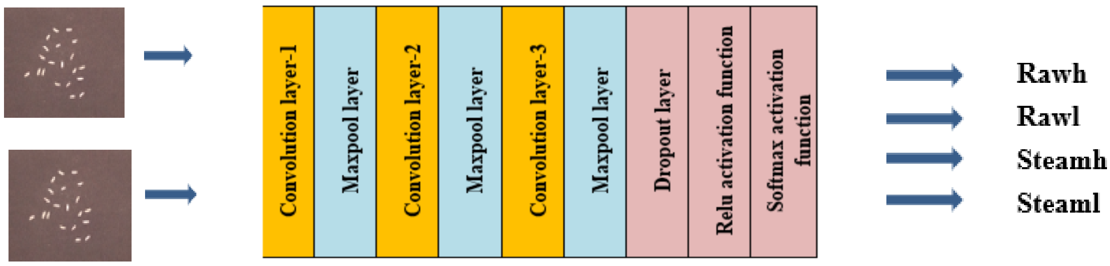

Figure 3. A customized CNN architecture

3. RESULTS AND DISCUSSION

This section presents research findings while also providing a thorough discussion. The analysis of performance measures based on the extracted features is discussed in this section [22]. SVM, KNN, DT, and EM (Voting classifier) are individually trained on each of these feature sets. A comparative analysis is then performed to identify the optimal feature set and model for classifying morphologically similar rice varieties. For the SVM with an RBF kernel, two key hyperparameters, 'C' and 'gamma,' are crucial in the training process. The model is tested with various gamma values to determine the most suitable one. Figure 4 presents the results obtained for each model across the different feature sets and gamma values. The outcomes derived from this investigation are organized and presented in Table 4 which includes the results of KNN, DT and EM along with SVM. According to the findings presented in Table 4, notable testing accuracy is attained with color, wavelet, and textural features. The best performance, reaching 72.48%, is observed when utilizing Polynomial kernel function of SVM with all 56 features.

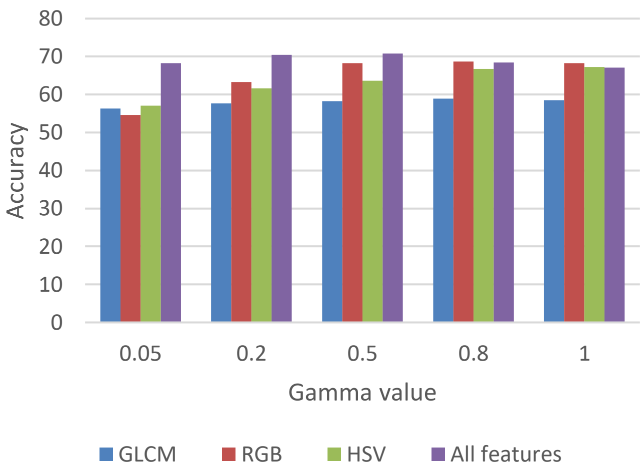

Figure 4. Graph showing the accuracy of SVM RBF kernel function for different gamma value

Table 4. The testing accuracies achieved by SVM (with linear, polynomial and RBF functions), KNN, and DT models on various feature sets are as follows

Feature Set	Linear	Polynomial	RBF	KNN	DT	EM
GLCM	56.33%	58.85%	58.9% (gamma=0.8)	56.81%	57.65%	59.68%
RGB (Color and Wavelet)	54.06%	72.24%	68.66% (gamma=0.8)	42.94%	48.32%	55.26%
HSV (Color and Wavelet)	56.93%	69.13%	67.22% (gamma=0.8)	52.15%	51.43%	58.85%
ALL	68.18%	72.48%	68.99%	58.49%	65.43%	65.78%

3.1. Training CNN model

Machine learning models trained over the extracted features provide less accuracy because of the incomplete feature set. Therefore, in this research CNN model is utilized which extracts high quantity feature set from the set of images. The 280 images assimilated are split into train and test set into 80:20 ratios. The images are resized to the size 224×224 prior to giving it as input to the CNN model. A CNN model with three convolution layers is developed for classification of similar looking rice varieties. The initial experiments involved using 32 filters of size 3×3 in all three layers which are later tuned and the results are recorded for varying epoch which is tabulated in Table 5.

Table 5. Performance of the CNN model the tuned number of filters in three layers against varying epochs

Filter size	Accuracy	Epochs				
		100	200	300	400	500
All three layers with 32 filters, each (3x3).	Training	45%	72.9%	89%	91.95%	94.25%
	Validation	50%	67.5%	80%	82.5%	87.5%
32 filters - Layer one and two	Training	50.7%	73.08%	81%	80.7%	90.7%
64 filters each (3×3)- layer three	Validation	50%	70%	73%	73%	73%
Layer one with 32 filters and two, three	Training	69.5%	85.6%	89.08%	94.8%	97.13%
with 64 filters each (3×3).	Validation	57.5%	80%	80%	80%	87.5%

The accuracy during training and validation phase is depicted in Figure 5 along with the loss curve, for the CNN model having 32 filters in layer 1 and 64 filters (each of size 3×3) in layer 2 and layer 3. It can be seen from the graph that loss decreases with the increment in the number of epochs and, simultaneously, the accuracy also increases with the increase in epochs. Table 6 depicts the total accuracy, recall, and F1-score in distinguishing morphologically analogous rice varieties when a CNN with 32 filters in layer 1 and 64 filters (3×3) in layer 2 and layer 3 is applied.

Figure 5. The Training and validation loss and accuracy curves of the CNN model

The result achieved in this work is compared with other significant work in the same field. The work carried out by Koklu *et al.* [7] has considered five distinct variety of grains grown in Turkey. By applying NN, DNN and CNN they have achieved an accuracy of 99.87%,99.95% and 100% respectively. Asif *et al.*

[9] have researched on classifying 5 distinct types of rice grown in Pakistan and have applied Principal component analysis (PCA) and Canny edge detection. They have got an accuracy of 89.5%. Singh *et al.* [8], Lin *et al.* [10] and Ahmed *et al.* [23] have worked on distinct variety of rice grains from their local region. They have applied types of neural network models and have obtained 99.4%, 96% and 88.07% of accuracy. Liu *et al.* [24] and Pranshu Saxena [25] have experimented on classification of five varieties of distinct rice types. They have achieved over 99.85% of accuracy by applying SVM and 93.02% by applying Random Forest Classifier.

Table 6. CNN model's precision, recall, and F1-score measures are provided with 32 filters at Layer 1 and 64 filters at both Layer 2 and Layer 3 at epoch=500

Type	Precision	Recall	F1-score
B.R.T Daiwan (Rawh)	86%	60%	71%
Grine world (Rawl)	100%	100%	100%
Royal Sona (Steamh)	69%	90%	78%
Daiwan Sona (Steaml)	100%	100%	100%
Average	89%	88%	87%

4. CONCLUSION

This novel research work to classify the morphologically similar Indian rice variety (medium slender) is quite a challenging task. The objective of this study is to assist consumers and the food industry in accurately identifying the appropriate type of rice. Samples of four medium-sized rice types, similar to Sona-Masoori rice, were purchased from the commercial market and analyzed at Shri Bhagyalakshmi Agro Food Pvt. Ltd. The results were compared with the earlier work carried out by varies researchers.

By looking at earlier work on rice quality analysis, it is evident that not much work is done on a morphologically similar variety of rice. This new work has attained highest training accuracy of 97.13% and testing accuracy of 87.5% on morphologically similar medium slender Indian rice classification. The study's approach involved image acquisition, pre-processing, and feature extraction, contributing to the outcomes gathered using various machine learning models. Various features, including textural, color, and wavelet decomposition in both RGB and HSV color spaces, were utilized as input for SVM, KNN, DT and EM (voting classifier) algorithms. Furthermore, the images acquired are rescaled to a pixel size of 224x224 and is given as input into the CNN model. With the entire 56-feature set, accuracy of 72.48% is obtained. The CNN model with the tuned parameters has achieved the best training accuracy of 97% and validation accuracy of 88%. The outcome indicates that textural, color, and wavelet features are essential for achieving high accuracy. While SVM performs best among ML models, CNN outperforms them by automatically extracting more features for better classification.

ACKNOWLEDGMENTS

We gratefully acknowledge Sri Bhagyalakshmi Agro Foods Pvt. Ltd., Bangalore, India, for the use of laboratory equipment to carry out the work. We also appreciate BNM Institute of Technology for giving permission to access the VGST-funded laboratory facilities, through which we were able to perform our experiments.

FUNDING INFORMATION

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Harini Shadaksharappa	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	
Saritha Chakrasali	✓	✓		✓	✓	✓			✓		✓	✓		
Krishnamurthy Gorappa Ningappa	✓	✓		✓	✓	✓			✓		✓	✓		

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest (financial, personal, or professional) in connection with manuscripts. Non-financial competing interests include a declaration of political, personal, religious, ideological, academic, and intellectual competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. If there are no conflicts of interest, please include the following author's statement: Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, HS. The data, which contain information that could compromise the privacy of research participants are not publicly available due to certain restrictions.

REFERENCES

[1] "ICAR-Indian institute of rice research," *ICAR-Indian Institute of Rice Research*. https://www.icar-iirr.org/index.php/en/ (accessed Mar. 14, 2023).

[2] "India and IRRI," *International Rice Research Institute*. https://www.irri.org/where-we-work/countries/india (accessed Mar. 14, 2025).

[3] S. E. Umbaugh, *Computer vision and image processing: A practical approach using CVIP tools*. Prentice Hall, 1997.

[4] G. Zhang, D. S. Jayas, and N. D. G. White, "Separation of touching grain kernels in an image by ellipse fitting algorithm," *Biosystems Engineering*, vol. 92, no. 2, pp. 135–142, Oct. 2005, doi: 10.1016/j.biosystemseng.2005.06.010.

[5] S. Harini, S. Chakrasali, and G. N. Krishnamurthy, "A survey on food grain identification techniques," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 7, no. 6, pp. 6627–6633, 2018.

[6] S. Harini, S. Chakrasali, and G. N. Krishnamurthy, "Analysis of Indian rice quality using multi-class support vector machine," 2023, pp. 791–802. doi: 10.1007/978-981-19-7346-8_68.

[7] M. Koklu, I. Cinar, and Y. S. Taspinar, "Classification of rice varieties with deep learning methods," *Computers and Electronics in Agriculture*, vol. 187, Aug. 2021, doi: 10.1016/j.compag.2021.106285.

[8] K. R. Singh and S. Chaudhury, "Efficient technique for rice grain classification using back-propagation neural network and wavelet decomposition," *IET Computer Vision*, vol. 10, no. 8, pp. 780–787, Jul. 2016, doi: 10.1049/iet-cvi.2015.0486.

[9] M. J. Asif, T. Shahbaz, S. Tahir Hussain Rizvi, and S. Iqbal, "Rice grain identification and quality analysis using image processing based on principal component analysis," in *2018 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, Oct. 2018, pp. 1–6. doi: 10.1109/RAEE.2018.8706891.

[10] 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, Apr. 2018, doi: 10.1007/s11947-017-2050-9.

[11] M. Koklu and I. Cinar, "Classification of rice varieties using artificial intelligence methods," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 7, no. 3, pp. 188–194, Sep. 2019, doi: 10.18201/ijisae.2019355381.

[12] P. K. Mallick and B. P. Mohanty, "Identification and classification of similar looking food grains," *International Journal Of Recent Advances in Engineering & Technology*, vol. 08, no. 04, pp. 21–25, Apr. 2020, doi: 10.46564/ijraet.2020.v08i04.006.

[13] H. S. Gujjar and M. Siddappa, "A method for identification of basmati rice grain of india and its quality using pattern classification," *International Journal of Engineering Research and Applications (IJERA)*, vol. 3, no. 1, pp. 268–273, 2013.

[14] V. S. Kolkure and B. N. Shaikh, "Identification and quality testing of rice grains using image processing and neural network," *International Journal of Recent Trends in Engineering & Research (IJRTER)*, vol. 3, no. 1, pp. 130–135, 2017.

[15] W. Srimulyani and A. Musdholifah, "Identification of rice variety using geometric features and neural network," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 3, p. 301, Jul. 2019, doi: 10.22146/ijccs.48203.

[16] N. Iqbal, R. Mumtaz, U. Shafi, and S. M. H. Zaidi, "Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms," *PeerJ Computer Science*, vol. 7, May 2021, doi: 10.7717/peerj-cs.536.

[17] E. T. Champagne, W. E. Marshall, and W. R. Goynes, "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, 1990.

[18] N. Handa, Y. Kaushik, N. Sharma, M. Dixit, and M. Garg, "Image classification using convolutional neural networks," 2021, pp. 510–517. doi: 10.1007/978-981-16-3660-8_48.

[19] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.

[20] D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *Artificial Neural Networks – ICANN 2010 (ICANN 2010)*, 2010, pp. 92–101. doi: 10.1007/978-3-642-15825-4_10.

[21] T. N. Sainath, A. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for LVCSR," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 8614–8618. doi: 10.1109/ICASSP.2013.6639347.

[22] R. Choudhary, J. Paliwal, and D. S. Jayas, "Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images," *Biosystems Engineering*, vol. 99, no. 3, pp. 330–337, Mar. 2008, doi:

10.1016/j.biosystemseng.2007.11.013.

[23] T. Ahmed, C. R. Rahman, and M. F. M. Abid, "Rice grain disease identification using dual phase convolutional neural network based system aimed at small dataset.," *arXiv:2004.09870*, Jul. 2021, doi: 10.31220/agriRxiv.2021.00062.

[24] T. Liu *et al.*, "A shadow-based method to calculate the percentage of filled rice grains," *Biosystems Engineering*, vol. 150, pp. 79–88, Oct. 2016, doi: 10.1016/j.biosystemseng.2016.07.011.

[25] P. Saxena, K. Priya, S. Goel, P. K. Aggarwal, A. Sinha, and P. Ain, "Rice varieties classification using machine learning algorithms," *Journal of Pharmaceutical Negative Results*, vol. 13, no. Special issue 07, 2022, doi: 10.47750/pnr.2022.13.S07.479.

BIOGRAPHIES OF AUTHORS

Harini Shadaksharappa working as an assistant professor, in information science and engineering, BMSCE. She obtained her M.Tech, in computer science and engineering, from School of Engineering, Jain University, Bangalore in the year 2012, and registered for Ph.D under VTU in December 2017. Her research areas include image processing, high performance computing and machine learning. She can be contacted at email: hharinis8@gmail.com

Dr. Saritha Chakrasali holds a Ph.D. in computer science from Madurai Kamaraj University, focusing on medical image security. She is the head of training & placement and a professor in the information science & engineering department at BNM Institute of Technology. She has received a ₹20 lakh VGST grant under the K-FIST L1 scheme for a project on GPGPU computing and is a life member of ISTE and CSI. She can be contacted at email: saritha.chakrasali@gmail.com

Dr. Krishnamurthy Gorappa Ningappa holds a B.E. from Kuvempu University and an M.Tech. & Ph.D. from Visvesvaraya Technological University. He has served as registrar (evaluation) at VTU and as principal and is currently the deputy director and professor of CSE at BNM Institute of Technology. With over 40 publications, he has also represented VTU in international delegations, visiting universities in Europe and beyond for academic collaborations. He can be contacted at email: bnmitprincipal@gmail.com.