Smart Agricultural Technology 12 (2025) 101065

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Kernel to computation: identifying optimal feature set for red rice classification

Suma D [a,1], Narendra V G [a,*], Darshan Holla M [b], Shreyas [b], Raviraja Holla M [b,*]

[a] Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, 576104, Karnataka, India

[b] Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, 576104, Karnataka, India

ARTICLE INFO

Keywords:
Feature extraction
Feature selection
Image processing
Red rice classification

ABSTRACT

While existing research focuses extensively on white rice classification with readily available datasets, automated classification of red rice varieties remains largely unexplored with no publicly available datasets, creating a significant research gap in agricultural image processing applications. This research presents a study on red rice classification, a relatively unexplored area with no prior publicly available datasets or focused investigations on red rice variety identification. This study classifies three distinct red rice varieties—Uma, KCP-1, and Jyothi—primarily cultivated in Karnataka and Kerala, using image processing and machine learning techniques. Six ML models were evaluated with seven unique feature combinations derived from size, shape, and texture characteristics to identify the most discriminative feature set. Feature selection was performed using Recursive Feature Elimination and Backward Feature Elimination to enhance model efficiency. Hyperparameter tuning was applied to optimize classification performance, and k-fold cross-validation with statistical significance testing was used to assess generalization and validate model performance differences. The integration of size, shape, and texture features yielded the highest average accuracy across the models, with K-Nearest Neighbours achieving 98.67 % accuracy and Support Vector Machine reaching 97.34 % accuracy with the size and shape combination. The findings emphasize the importance of optimal feature selection and tuning in improving classification accuracy, contributing to the development of automated classification systems for red rice varieties.

1. Introduction

Red rice varieties are gaining considerable popularity across the globe due to their exceptional nutritional benefits and distinctive culinary profiles. In South India, red rice is a fundamental staple that complements the dietary habits and cultural practices of the region. Among the numerous varieties, Uma, KCP-1, and Jyothi stand out not only for their unique flavours but also for their health benefits which cater to the contemporary focus on healthier eating habits. Particularly, the variety known as Uma, cultivated through organic practices and traditional methods at Kerala Agricultural University, is revered for its high content of iron and calcium. This variety is especially beneficial for individuals with diabetes, offering a lower glycaemic index compared to white rice variants. The traditional methods of parboiling and sun-drying, which are used to process Uma rice, significantly enhance its nutritional value while maintaining its natural qualities. The rice is grown on a single farm, ensuring consistency in its quality and taste, and is parboiled in a large vessel over a wood fire, a method that helps retain most of the bran's nutrients, making it a nutritious choice in the daily diet [1,3].

KCP-1, another prominent variety, is tailored to the coastal agricultural zones of Karnataka and is recognized for its high yield and superior straw quality, making it an economically viable option for farmers. This variety matures in about 130–145 days and can produce yields ranging from 6 to 6.5 tons per hectare, which is considerably higher than many other varieties. It has long, bold grains that are not only ideal for cultivation in lowland areas but are also preferred for their cooking quality. Moreover, the KCP-1 variety has shown resistance to various

* Corresponding authors.
E-mail addresses: suma.d@manipal.edu (S. D), narendra.vg@manipal.edu (N. V G), darshan.m1@learner.manipal.edu (D.H. M), shreyas.mcisma1@learner.manipal.edu (Shreyas), raviraj.holla@manipal.edu (R.H. M).

[1] Contributing authors.

https://doi.org/10.1016/j.atech.2025.101065
Received 8 April 2025; Received in revised form 30 May 2025; Accepted 31 May 2025
Available online 1 June 2025

paddy diseases and pests, including gall midge, further enhancing its appeal to agronomists and farmers looking for reliable crop options. In a similar vein, Jyothi rice, also known as Vadi matta, is cultivated predominantly in the Palakkadan region and is famed for its high fibre content and rich nutritional profile. It is an unpolished, parboiled rice that retains a significant amount of natural nutrients, making it a healthful staple in South Indian cuisine. This variety's adoption is widespread due to its adaptability in various traditional dishes, from daily meals to festive delicacies, underscoring the integration of agricultural diversity with nutritional needs [1,2].

These rice varieties not only represent significant agricultural achievements but also play a crucial role in addressing nutritional security. Their enriched nutrient profiles and adaptability to local agronomic conditions illustrate the potential of traditional and improved varieties in promoting healthful eating practices while supporting sustainable farming methods in South India. As such, the study of these varieties offers valuable insights into the synergy between dietary health and agricultural practices, highlighting the importance of preserving and promoting indigenous crop varieties in the fight against malnutrition and health disorders related to diet.

The proposed method for automated classification of red rice varieties addresses a critical gap in rice classification, where automated systems have primarily focused on white rice, leaving red rice unexplored due to the absence of dedicated datasets and studies. This study introduces a novel framework to classify Uma, KCP-1, and Jyothi varieties, employing a comprehensive methodology that includes sample collection of 300 grains (100 per variety) from the Zonal Agricultural & Horticultural Research Station, Brahmavar, and standardized image acquisition using a 45-megapixel camera under D-65 lighting. Dataset preparation involved pre-processing (grayscale conversion, Otsu's thresholding, morphological operations), followed by feature extraction of size, shape, and texture characteristics. Feature integration evaluated all possible combinations to identify the optimal feature set, achieving up to 98.67 % accuracy with K-Nearest Neighbours. Feature selection, a key novelty, utilized Recursive Feature Elimination (RFE) and Backward Feature Elimination (BFE) to reduce redundant features, enhancing model efficiency. Traditional machine learning models were employed due to the limited dataset size (100 samples per class) and the need for computationally efficient, interpretable models suitable for agricultural applications. Hyperparameter tuning optimized model performance, while k-fold cross-validation ($k = 5$), statistical significance testing, and visualization via confusion matrices and ROC curves ensured robust evaluation and generalization. This approach enhances quality control and authentication in the agricultural supply chain, supporting stakeholders in meeting the growing demand for nutritious red rice in South India.

1.1. Motivation

The impetus for this study stems from the limited research on automated classification of red rice varieties and the lack of publicly available datasets. While machine learning techniques have been applied to classify white rice and other crops, no prior studies focus on specialized red rice varieties like Uma, Jyothi, and KCP-1. This research aims to bridge this gap by developing a robust classification framework that enhances supply chain efficiency, market valuation, and quality assurance. As consumer demand for red rice grows due to its nutritional benefits, accurate classification methods are essential to ensure consistency in quality and meet market expectations. Furthermore, red rice varieties possess unique morphological and textural characteristics that require specialized classification approaches beyond those developed for white rice. The development of this novel framework will not only contribute to academic knowledge but also provide practical tools for stakeholders to authenticate varieties and prevent adulteration in market.

1.2. Key contributions and novelty

- Addresses the gap in prior research by developing an automated classification framework for red rice varieties, which has not been explored before.
- Absence of publicly available red rice datasets necessitated the creation of a custom, high-quality image collection for this research.
- Feature extraction and selection were performed using Recursive Feature Elimination and Backward Feature Elimination to identify the most discriminative feature set.
- Hyper-parameter tuning was conducted to optimize model parameters for achieving the best classification performance.
- Comparison of six machine learning models, including an ensemble learning approach, to determine the most suitable algorithm for red rice variety identification.
- Existing studies on white rice classification rely on deep learning methods, which are computationally expensive, use an excessive number of features (including irrelevant ones), or achieve lower accuracy. This study provides an efficient alternative with optimal feature selection and model tuning for improved performance.

2. Related works

Image processing and machine learning have been effectively utilized in the rice industry for grain classification as demonstrated by Biren Arora et al. [4]. They developed a system that employs Logistic Regression, Decision Tree, Naïve Bayes, KNN, and SVM for the classification of four rice varieties: Surti Kolam, Idli Rice, Long Grain Basmati, and Boiled Rice, aiming to enhance industrial efficiency by automating the sorting process. Nadeesha Nagoda et al. [5] implemented a machine vision-based system for rice quality assessment, using SVM. Their system uses the Watershed Algorithm, Local Binary Pattern (LBP) texture features, and colour features to classify rice into categories like full rice, broken rice, and others, achieving a high segmentation accuracy of 96 % and classification accuracy of 88 %. The study by Kantip Kiratiratanapruk et al. [6] introduces an automated system for classifying paddy rice seeds using machine vision. Their system, which employs InceptionResNetV2 and SVM, manages to classify 14 different rice varieties based on shape, colour, and texture with an accuracy up to 95.15 %. Exploring deep learning's potential in rice grain classification, Kittinun Aukkapinyo et al. [7] employed Mask R-CNN for localization and classification of rice grains. Their model surpasses traditional methods with significant accuracy improvements, demonstrating deep learning's effectiveness in handling visually similar grains. Chandrika Vijaya Krishna et al. [8] proposed an automated rice quality assessment system that utilizes image processing and morphological techniques. Their method, leveraging MATLAB for feature extraction and SVM and Neural Networks for classification, shows superior performance over manual inspection methods.

Continuing this technological advancement, Shafaf Ibrahim et al. [9] developed a multi-class SVM-based classification system for rice grains. They extracted shape and colour features, significantly enhancing the speed and reliability of classification in the agro technology industry with an accuracy of 92.22 %. Rashidah Ruslan et al. [10] presented a system for classifying weedy rice using a comprehensive set of 67 morphological, colour, and texture features. Their method, utilizing multiple machine learning classifiers, achieved the highest accuracy of 97.9 % with Logistic Regression. A cascade network approach was used by Ksh Robert Singh et al. [11] to classify rice grains. They integrated a Backpropagation Neural Network with a fuzzy inference system, improving classification accuracy with the use of morphological features, demonstrating robustness against UCI datasets. Ilkay Cinar et al. [12] analysed a vast dataset of 75,000 rice grain images using a variety of machine learning models. Their study emphasized the critical role of feature selection, with the Multilayer Perceptron model achieving nearly perfect accuracy. In a comparison of deep learning architectures, Farshad Farahnakian et al. [13] assessed the performance of models like ResNet and Mobile Net on a large dataset of rice grains. They noted that EfficientNet not only achieved the highest accuracy but also highlighted the efficiency of deep learning in automating classification tasks.

Further research by Tran Thi Kim Nga et al. [14] optimized feature selection using Binary Particle Swarm Optimization combined with SVM, enhancing the classification accuracy for 17 Vietnamese rice varieties. Tzu-Yi Kuo et al. [15] employed a sparse-representation-based classification system, which proved to be robust and computationally efficient, particularly suitable for the authentication of rice based on microscopic images. Leveraging Near-Infrared Spectroscopy and machine learning, Letícia de Oliveira Carneiro et al. [16] assessed the quality of milled rice, with their models demonstrating a strong correlation between physicochemical properties and rice quality, underscoring the potential of these technologies in non-destructive testing. The comprehensive review by Sheikh Bilal Ahmed et al. [17] traced the evolution of rice grain analysis techniques, highlighting a shift towards deep learning approaches which have shown superior accuracy in classification and disease identification.

Wahyu Srimulyani et al. [18] explored the utility of neural networks in classifying rice varieties using geometric features, with their methods demonstrating remarkable accuracy improvements. The use of a Backpropagation Neural Network enhanced by wavelet decomposition was explored by Ksh. Robert Singh et al. [19], showing superior performance over traditional classifiers due to the robustness of wavelet-based features. Aimi Aznan et al. [20] demonstrated the use of smartphone images for commercial rice grain classification, highlighting the practical applications of ANN with Bayesian Regularization in real-time quality assessment. Reviewing ANN techniques, Anis Sufiya Hamzah et al. [21] emphasized the potential for hybrid methods in enhancing the classification and quality assessment of white rice grains. Sharmin Sultana et al. [22] discussed the impact of quantitative trait loci on rice grain quality, advocating for molecular breeding techniques to improve rice quality parameters.

Md Taimur Ahad et al. [23] evaluated CNN architectures for rice disease classification, highlighting the effectiveness of ensemble models in real-time disease detection. The study by Nga Tran-Thi-Kim et al. [24] underscored the advantages of CNNs over traditional ML techniques for rice variety classification, using advanced image processing and data augmentation strategies. Bharati Patel and Aakanksha Sharaff [25] introduced a deep learning-based approach for rice variety classification and yield prediction, demonstrating the applicability of semantic segmentation in this field. Chinmay Kurade et al. [26] developed a cost-effective, automated rice quality assessment system using a Raspberry-Pi-based module, which performed robustly across various rice varieties. Deepika C. et al. [27] applied Linear Discriminant Analysis in conjunction with digital imaging to classify rice grain quality, achieving high accuracy in identifying high-yielding aromatic hybrids. Teresa Mary Philip and H. B. Anita [28] presented a novel classification system based on Fourier Transform and morphological features, achieving high classification accuracy across multiple South Indian rice varieties.

While prior studies [4–6] have advanced rice classification using image processing and machine learning, they are not related to red rice varieties like Uma, KCP-1, and Jyothi, which are nutritionally significant and in growing demand in South India. These varieties possess distinct morphological characteristics, such as colour intensity, grain shape, and texture, which require specialized classification techniques. Many deep learning approaches [7,13,24] rely on large datasets and computationally intensive models, making them less suitable for resource-constrained environments where high-end hardware and large annotated datasets are not readily available. Additionally, training deep models on small datasets often leads to overfitting and poor generalization. Studies [9,10,12] use extensive feature sets that may contain redundant or irrelevant attributes, increasing computational load and complicating model interpretability.

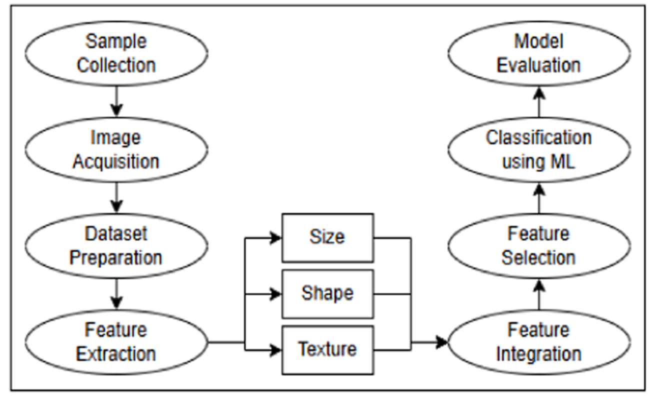

Fig. 1. A Block Diagram Depicting the Overall Methodology.

In contrast, the proposed method focuses on red rice classification using optimized traditional machine learning models on a small dataset (100 samples per class), achieving 98.67 % accuracy through rigorous feature selection techniques such as Recursive Feature Elimination and Backward Feature Elimination. This selective approach ensures that only the most informative morphological features are retained, resulting in faster training and testing times. Moreover, the model design supports easier deployment on field-level systems or mobile-based agricultural tools. Unlike previous studies [8,28], which lack comprehensive cross-validation, this study employs k-fold cross-validation ($k = 5$) and statistical significance testing to ensure robust generalization. By focusing on red rice, optimizing feature sets, and prioritizing computational efficiency, the proposed framework offers a practical and effective solution for automated red rice classification that is both scalable and applicable in real-world agricultural settings.

3. Proposed method

The methodology adopted in this study, illustrated in Fig. 1, begins with sample collection and image acquisition, followed by dataset preparation for optimal pre-processing. Features extracted are of 3 types - size, shape, and texture. These features are integrated and refined through feature selection before being input into various machine learning models for classification. The process concludes with model evaluation using k-fold cross-validation and statistical significance testing, followed by data visualization, ensuring precise classification and deep insights into rice grain varieties.

3.1. Sample collection

The rice samples for this study were collected from the Zonal Agricultural & Horticultural Research Station, Brahmavar. The samples consisted of 300 rice grains, with 100 grains from each of the three distinct rice varieties: Uma, Jyothi, and KCP-1. Each variety was selected based on its prevalence and significance in regional agriculture, providing a diverse basis for the comparative analysis of grain characteristics.

3.2. Image acquisition

The acquisition of rice grain images was conducted under standardized conditions to ensure consistency across all samples. A 45-megapixel camera was utilized, and the images were captured under D-65 lighting, which provides a consistent illumination. A black velvet fabric background was used to avoid shadows and reflections. The non-reflective quality of velvet is crucial as it minimizes light scattering, ensuring uniform lighting across all images. This setup, as shown in Fig. 2, helps in capturing high-quality images that are critical for detailed image analysis.

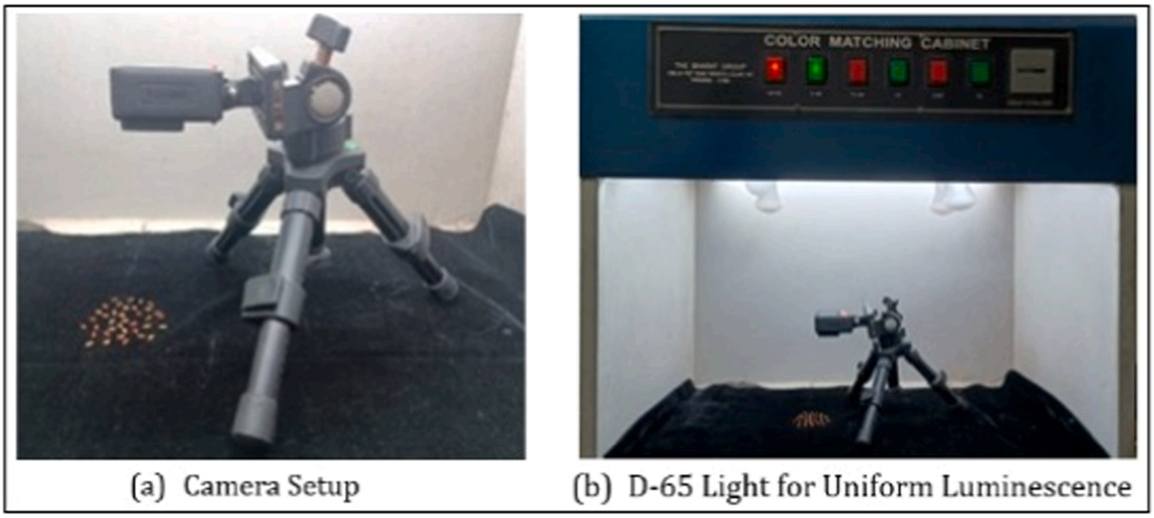

Fig. 2. Image Acquisition Setup.

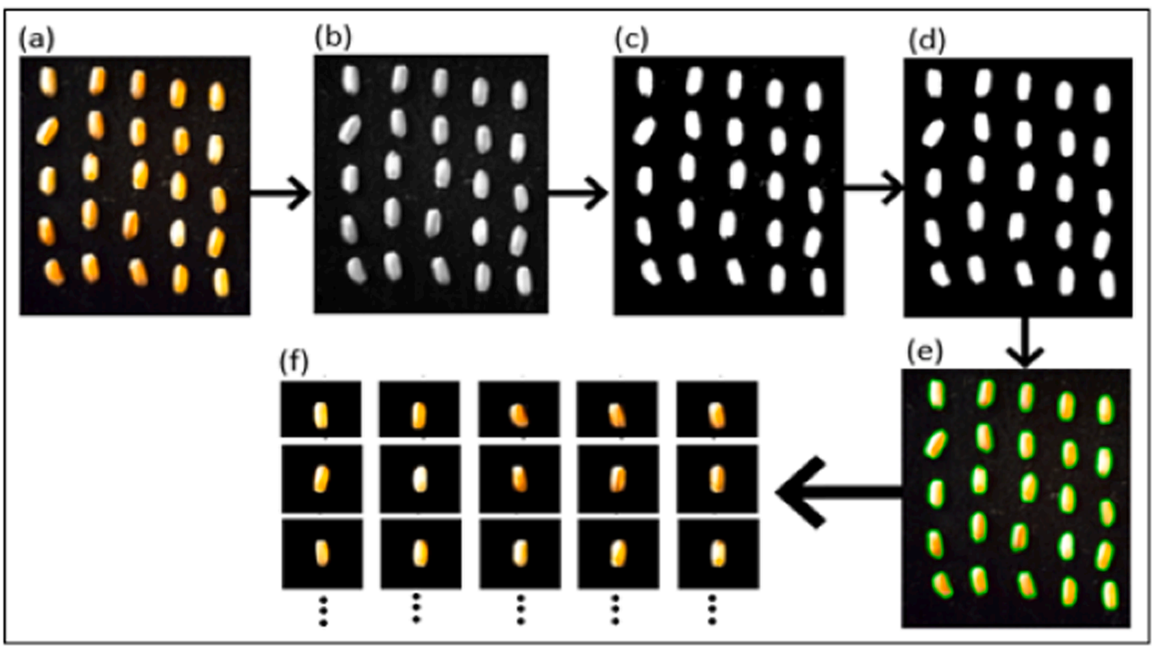

Fig. 3. Steps Involved in Dataset Preparation. (a) Original Image (b) Grayscale Image (c) Thresholded Image (d) Opened Image (e) Identified Contours (f) Cropped Rice grains.

3.3. Dataset preparation

The dataset preparation workflow began with the original image shown in Fig. 3(a), structured into distinct stages to refine the 300 samples—100 grains each of Uma, Jyothi, and KCP-1—for comprehensive feature extraction. The dataset was then split into training and testing subsets, with 75 % allocated for training and 25 % for testing.

3.3.1. Pre-processing

Each image was initially converted from RGB to grayscale to simplify the analysis by reducing computational complexity while retaining necessary textural details. This conversion is mathematically represented by (1), where the grayscale value is calculated as a weighted sum of the colour intensities.

$$\mathrm{Gray} = 0.299\mathrm{R} + 0.587\mathrm{G} + 0.114\mathrm{B} \tag{1}$$

Here, R, G, and B stand for the intensities of the red, green, and blue channels, respectively. This conversion process is demonstrated in Fig. 3(b). Subsequently, Otsu's thresholding, described by (2), was applied to binarize the images. This method calculates an optimal threshold that minimizes the intra-class variance of the black and white pixels, effectively separating the grains from the background, as demonstrated in Fig. 3(c).

$$\sigma_w^2(\mathrm{t}) = \omega_0(\mathrm{t})\sigma_0^2(\mathrm{t}) + \omega_1(\mathrm{t})\sigma_1^2(\mathrm{t}) \tag{2}$$

Here, $\omega_0(\mathrm{t})$and $\omega_1(\mathrm{t})$ are the probabilities of the two classes separated by the threshold t and $\sigma_0^2(\mathrm{t})$ and $\sigma_1^2(\mathrm{t})$ are the variances of these classes.

3.3.2. Noise removal

To ensure that the grains are distinctly separated and free from artefacts, morphological operations such as opening were employed. This crucial step, guided by (3), cleans up the image by removing small-scale noise and separating grains that are in close contact. These modifica-

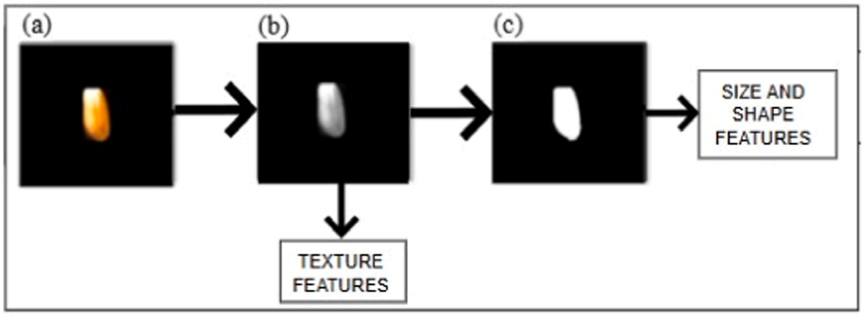

Fig. 4. Steps Involved in Dataset Preparation. (a) Original Image (b) Grayscale Image (c) Segmented Image (d) Labelled Image.

tions are essential for the subsequent accurate detection of contours. The transformation from the segmented image in Fig. 3(c) to the cleaned image in Fig. 3(d) vividly illustrates the effectiveness of this noise removal process.

$$Opening(Img) = Dilation(Erosion(Img, S), S) \tag{3}$$

Here, *Img* refers to the segmented image, a binary representation where the morphological operations are applied to clean and separate image structures. *S* denotes the structuring element, a 4 × 4 matrix of ones, which shapes how the erosion and dilation processes modify the image *Img*.

3.3.3. Contour detection

Contours of individual grains were detected from the cleaned images. Defining the boundaries of each grain is essential for extracting both size and shape features, highlighted in Fig. 3(e).

3.3.4. Cropping and saving grains

Post contour detection, bounding boxes for each contour were computed using (4), which facilitated the cropping of individual grains. Each grain was then isolated and saved as a distinct image. To ensure each cropped grain is centrally positioned within a standardized frame, horizontal and vertical centring calculations were applied as defined in (5) and (6). This process supports independent feature extraction and enhances the precision of the data collected, as depicted in Fig. 3(f).

$$(x, y, w, h) = boundingRect(C_i) \tag{4}$$

Eq. (4) calculates the coordinates x and y of the top-left corner, along with the width w and height h of the rectangle that minimally encloses the i^{th} contour C_i.

$$X_{offset} = \frac{(256 - w)}{2} \tag{5}$$

$$Y_{offset} = \frac{(256 - h)}{2} \tag{6}$$

Eq. (5) and (6) determine the offsets for centring the cropped grain images within a new 256×256 pixel canvas, ensuring uniformity and consistency in image analysis.

3.4. Feature extraction

The feature extraction process in this study captures a comprehensive set of characteristics from the rice grains, focusing on size, shape, and texture features. The process initiates with the original image, converted into grayscale, then thresholded for segmentation as illustrated in Fig. 4(a), Fig. 4(b), and Fig. 4(c) respectively. Texture features were extracted from grayscale images as they depend on intensity variations which are preserved in grayscale images, while size and shape are structure-based features that depend on the outline of the grain and hence extracted from segmented images.

In this study, distinct feature categories are used to evaluate the inherent characteristics of rice types. Features F_1 to F_{10}, listed in Table 1, are designated as texture features, critical for assessing surface details that differentiate rice varieties. Features F_{11} to F_{16} are focused on size attributes, essential for determining the physical contours of the grains. Features F_{17} to F_{25} delve into shape properties, providing intricate insights into the structural nuances of rice grains.

3.5. Feature integration

Feature integration is a crucial aspect of the methodology, involving the consolidation of distinct feature sets—size, shape, and texture—into a unified dataset. This integration facilitates the evaluation of all meaningful combinations of these feature sets, totalling seven different scenarios, which include every possible single and combined feature grouping. This approach forms the basis of a comparative study aimed at identifying which combinations most effectively enhance the classification accuracy of rice grain varieties. By systematically analysing these combinations, the most informative features can be determined, thereby optimizing the ML models for precise and efficient grain type differentiation.

3.6. Feature selection

Feature selection enhances machine learning model performance by identifying and retaining the most relevant features. Recursive Feature Elimination (RFE) was used for SVM, RF, and LR because these models provide feature importance scores or coefficients, enabling features to be ranked and the least important removed iteratively. For KNN and GNB, which lack built-in feature importance, Backward Feature Elimination (BFE) was applied by removing features one at a time while monitoring performance. These methods helped reduce model complexity, lower computational cost, and improve classification accuracy by removing redundant or irrelevant features. The working of RFE and BFE is illustrated in Fig. 5 and Fig. 6 respectively. In these figures, *S* represents the current set of features, f_k (in Fig. 5) *and* f_i (in Fig. 6) refers to individual features with their importance ranking, and *N* is the target number of features to retain after the elimination process.

3.7. Classification using ML models

In this study, a variety of ML models including Support Vector Machines (SVM), Random Forest (RF), Gaussian Naive Bayes (GNB), Logistic Regression (LR), K Nearest Neighbours (KNN), and Ensemble Learning (EL) techniques were utilized to classify rice grains based on extracted features. Each model was chosen for its unique strengths in handling different aspects of the classification task.

3.7.1. Support vector machine

SVM is recognized for its proficiency in defining an optimal hyperplane that maximizes the margin between classes. This feature is particularly beneficial for datasets where classes are linearly separable,

Table 1
List of features extracted.

Feature Index	Feature Name	Description
F_1	Mean	Statistical measure that describes the brightness and contrast variations within the grain
F_2	Standard Deviation	Statistical measure that describes the dispersion of pixel values
F_3	Variance	Statistical measure that quantifies the spread of pixel values
F_4	Skewness	Describes the asymmetry of the grayscale intensity distribution
F_5	Kurtosis	Describes the tailedness of the grayscale intensity distribution
F_6 [0–9]	Local Binary Pattern (LBP) Histogram	It computed using the Uniform LBP variant with the following parameters: Radius=1, Number of Neighbours = 8; Method = 'uniform'. It captures local texture by encoding the frequency of uniform binary patterns in a pixel's neighbourhood.
F_7	Contrast	Measures the local variations in the grayscale image using the gradient magnitude
F_8	Directionality	Indicates the predominant direction of the gradients
F_9	Edge Density	The proportion of edge pixels after applying an edge detection filter, indicating the roughness or smoothness of the grain boundary
F_{10}	Texture Energy	The sum of squared elements in a smoothed image, representing the textural uniformity
F_{11}	Length	The measurement of the grain along its longest axis
F_{12}	Width	The measurement across the grain perpendicular to the length
F_{13}	Aspect Ratio	The ratio of length to width, indicating the elongation of the grain
F_{14}	Area	The total pixel area occupied by the grain
F_{15}	Perimeter	The total length of the boundary of the grain
F_{16}	Lack of Shape	It captures how irregular the grain shape is compared to a perfect rectangle. It helps in distinguishing rice varieties based on subtle differences in shape structure. It is calculated as: $F_{16} = (F_{11} \times F_{12}) / F_{14}$
F_{17} [1–7]	Hu Moments	Invariant descriptors that capture shape information and are invariant to image transformations such as translation, scale, and rotation
F_{18}	Eccentricity	A measure of how much the shape of the grain deviates from being a perfect circle
F_{19}	Solidity	The ratio of the area of the grain to the area of its convex hull, indicating the compactness of the grain
F_{20}	Extent	The ratio of the pixels in the region to the pixels in the total bounding box
F_{21}	Orientation	The angle at which the grain is oriented with respect to a horizontal axis
F_{22}	Major Axis Length	The length of the longest line that can be drawn through the grain
F_{23}	Minor Axis Length	The length of the shortest line that can be drawn through the grain perpendicular to the major axis
F_{24}	Compactness	A measure that describes how closely the area of the grain approaches that of a circle having the same perimeter as the grain
F_{25}	Convexity	Convexity indicates how curved or bulged the grain is. It's useful for identifying grains with irregular boundaries, which vary between different rice types. It is calculated as: $F_{25} = F_{15}$ / Convex Perimeter; Where 'Convex Perimeter' is the perimeter of the convex hull of the grain — the smallest convex shape that wraps around all outer points, ignoring any indentations.

making SVM highly appropriate for clear and distinct classification tasks. The objective of SVM to maximize the margin is mathematically represented as $\max_{w,b} = \frac{1}{\|w\|}$, where w is the weight vector determining the orientation of the hyperplane, and b is the bias term that allows the hyperplane to shift, optimizing the separation margin between the classes.

3.7.2. Random forest

RF is utilized for its capability to handle overfitting effectively, which is a common challenge in complex models. By employing an ensemble of decision trees, RF not only ensures robust performance but also provides valuable feature importance metrics. These metrics are crucial for identifying the most relevant features in noisy or complex datasets, thus ensuring the robustness of the model.

3.7.3. Gaussian Naïve Bayse

GNB is favored for its simplicity and effectiveness, particularly when the assumption of feature independence between predictors holds true. It is efficient in computational terms and performs admirably in scenarios involving multi-class predictions, making it well-suited for handling large datasets. The Gaussian Naive Bayes model calculates the probability of a label y given predictors $x_1, \ldots, x_n$ based on the assumption of independence among the predictors. This probability model, which facilitates classification, is shown in (7).

$$P(y|x_1, \ldots, x_n) = \frac{P(y) \prod_{i=1}^{n} P(x_i|y)}{P(x_1, \ldots, x_n)} \tag{7}$$

3.7.4. Logistic regression

LR is adept at handling multi-class classification problems using a one-vs-rest scheme. This approach makes it particularly effective in scenarios where there are multiple categories to classify, such as different types of rice grains. Its ability to provide probabilities for different classes enhances its interpretative ability, which is beneficial for detailed class analysis. The logistic function, as modelled by (8), quantifies the probability that the input x belongs to the class $y = 1$.

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n)}} \tag{8}$$

3.7.5. K-Nearest neighbours

KNN is chosen for its simplicity and effectiveness, especially useful when the decision boundary between classes is not clearly defined. This model operates on the principle of proximity, where the classification of a sample is determined based on the majority class among its nearest neighbours. This characteristic makes KNN particularly adept at handling datasets with complex, non-linear class separations, ensuring that each prediction is made with contextual insights derived from the closest data points.

3.7.6. Ensemble learning

Ensemble Learning, specifically through a stacking approach, integrates multiple models to exploit their individual strengths and mitigate their weaknesses. This strategy, depicted in Fig. 7, improves the predictive accuracy and robustness of the classification, making it exceptionally reliable for complex decision-making tasks. In our study, stacking utilized base models such as LR, RF, SVM, and GNB, with GNB also acting as the final estimator to amalgamate the predictions into a cohesive output. GNB was chosen as the final estimator after experimenting with multiple classifiers, as it consistently gave the best validation performance—likely due to its simplicity, efficiency, and ability to handle diverse meta-features generated by the base models.

3.8. Model evaluation

This section outlines the methods used to evaluate the performance

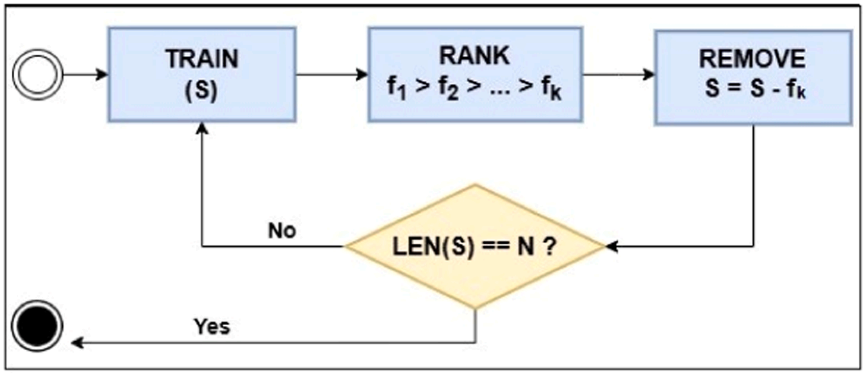

Fig. 5. Recursive Feature Elimination Process.

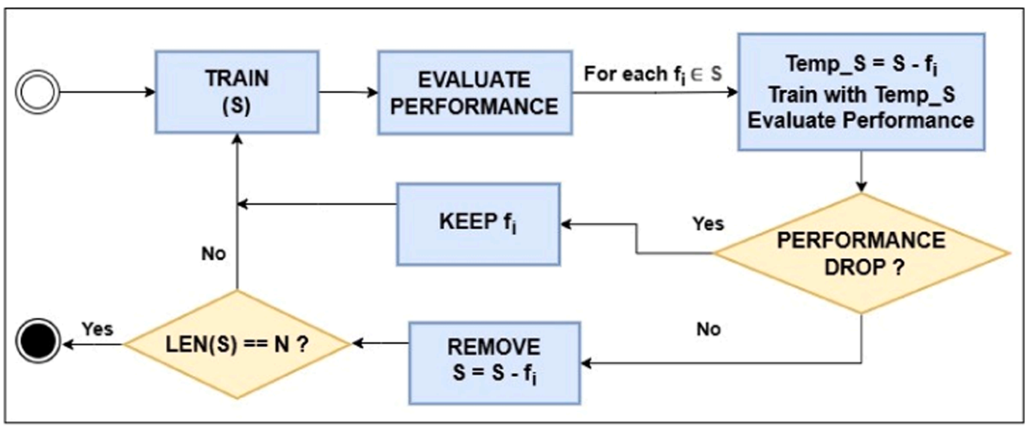

Fig. 6. Backward Feature Elimination Process.

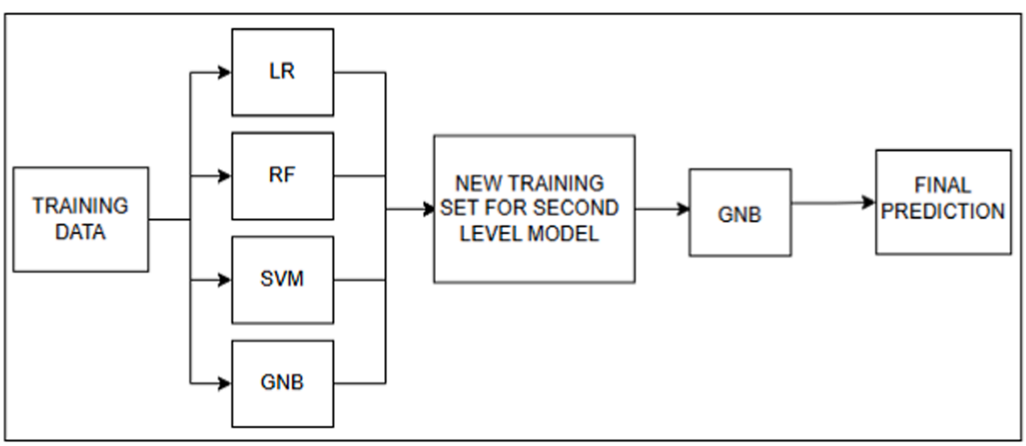

Fig. 7. Workflow of Stacking Ensemble Model.

of the model. Multiple evaluation metrics were employed, including Accuracy, Precision, Recall, F1-score, Confusion Matrices, Receiver Operating Characteristic (ROC) curves, and Area Under the Curve (AUC) to comprehensively assess classification effectiveness.

3.8.1. Performance metrics

The primary metrics for evaluating the models include accuracy, precision, recall, and the F1-score, as illustrated in (9), (10), (11) and (12). Accuracy measures the overall correctness of the model across all classes, while precision and recall provide insights into the model's performance concerning individual classes, which is crucial for applications where the balance between false positives and false negatives is critical. The F1-score is particularly useful as it is a harmonic mean of precision and recall, providing a single metric to assess the balance between the two.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

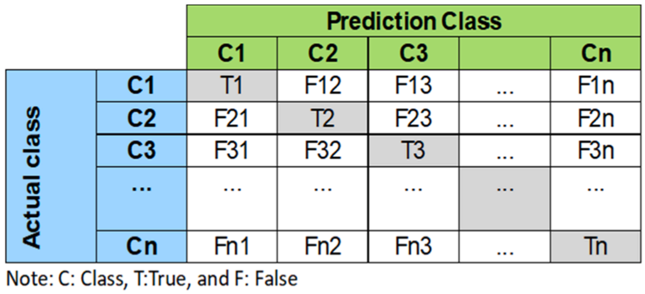

Fig. 8. General Confusion Matrix for Multiclass Classification.

Table 2
Accuracy Comparison Across Individual Feature Sets.

	Size	Shape	Texture
SVM	90.67	92	80
RF	90.67	92	85.34
GNB	94.67	86.67	82.67
LR	93.34	92	90.67
KNN	93.34	94.67	94.67
EL	93.34	92	90.67

Table 3
Accuracy comparison across combined feature sets.

	SIZE & SHAPE	SIZE & TEXTURE	SHAPE & TEXTURE	SIZE, SHAPE & TEXTURE
SVM	97.34	90.67	96	96
RF	96	93.34	93.34	93.34
GNB	96	90.67	96	96
LR	92	96	98.67	97.34
KNN	96	98.67	95	98.67
EL	93.34	96	97.34	97.34

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

3.8.2. Confusion matrix

A confusion matrix, shown in Fig. 8, is an essential tool for understanding the performance of a classification model. It categorizes the predictions into four types: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). This matrix helps in visualizing the accuracy of predictions across different classes by indicating not only the correct predictions (TP and TN) but also the errors in the form of FP, where the model incorrectly predicts the positive class, and FN, where it fails to detect the positive class. This detailed analysis helps identify patterns in misclassification, enabling researchers to refine their classification strategies.

Table 4
K-fold cross-validation results for classification models.

Model	Mean CV Accuracy (%)	Standard Deviation
SVM	95.56	3.44
RF	92.50	1.02
GNB	92.60	3.10
LR	96.44	1.09
KNN	96.80	3.00
EL	95.56	1.41

3.8.3. ROC and AUC metrics

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are used to evaluate the performance of the models at various threshold settings. The ROC curve plots the true positive rate against the false positive rate, providing insights into the trade-offs between benefiting from correct predictions and suffering from false alarms. The AUC provides a single value that summarizes the performance of the model across all thresholds, with higher values indicating better discrimination capabilities between the classes.

4. Experimental findings and results

Comprehensive analysis was conducted on all seven feature combinations using six ML models, evaluating their performance through accuracy, precision, recall, F1-score, confusion matrices, and feature selection patterns. For brevity and clarity, while Table 2 and Table 3 present the accuracy comparisons across all feature combinations, the detailed performance metrics are presented for the Size + Shape + Texture combination, which attained the highest average accuracy of 96.44 %, as depicted in Fig. 9, across various models.

4.1. K-fold cross-alidation: evaluating model generalization

To validate the generalization capability of the models and assess potential overfitting, k-fold cross-validation was implemented, a robust technique that partitions the dataset into k equal subsets or folds. For each iteration, k-1 folds are used for training while the remaining fold serves as a validation set, with this process repeating until each fold has been used exactly once for validation. This approach provides a comprehensive evaluation of model performance across different data subsets, offering more reliable metrics than a single train-test split. This

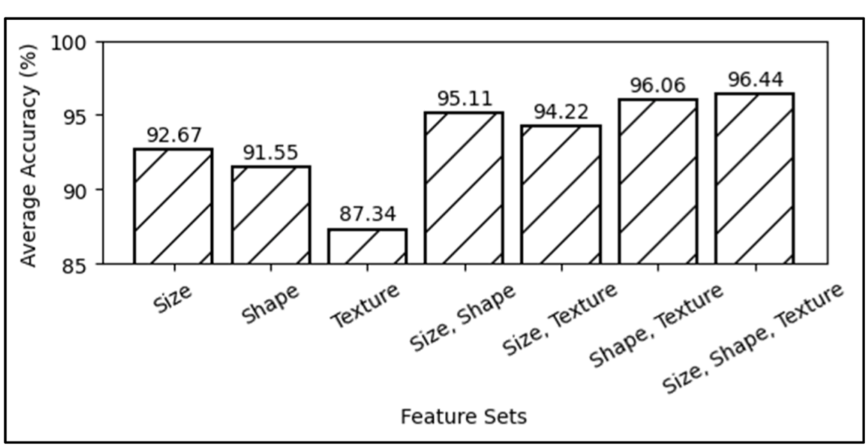

Fig. 9. Average Accuracy across Models for Various Feature Sets.

Table 5
Model performance and selected features for optimal feature set.

Selected Features	Models	Performance Metrics	Class		
			Uma	KCP-1	Jyothi
F_{13}, F_{17} [3], F_{17} [4], F_{21}, F_{22}, F_{23}, F_1, F_2, F_5, F_7, F_{10}	**SVM**	Precision	89	100	100
		Recall	100	96	92
		F1-Score	94	98	96
F_{11}, F_{12}, F_{13}, F_{14}, F_{17} [1], F_{17} [2], F_{17} [3], F_{17} [4], F_{18}, F_{21}, F_{11}, F_{22}, F_{23}, F_{11}, F_1, F_2, F_3, F_4, F_5, F_6 [3], F_6 [5], F_6 [8], F_7, F_9, F_{10}	**RF**	Precision	96	89	95
		Recall	100	96	84
		F1-Score	98	92	89
All	**GNB**	Precision	100	92	96
		Recall	100	96	93
		F1-Score	100	94	94
F_{11}, F_{18}, F_2, F_6 [6], F_{17} [1], F_{17} [2], F_{17} [4], F_{17} [5], F_{17} [7], F_{19}, F_{20}, F_{21}, F_1, F_2, F_6 [0], F_9	**LR**	Precision	100	100	93
		Recall	100	92	100
		F1-Score	100	96	96
F_{11}, F_{13}, F_{14}, F_{16}, F_6, F_7, F_9	**KNN**	Precision	100	100	96
		Recall	100	96	100
		F1-Score	100	98	98
All	**EL**	Precision	100	100	93
		Recall	100	92	100
		F1-Score	100	96	96

study conducted k-fold cross-validation with $k = 5$, allowing each model to be evaluated on five different validation sets. As shown in Table 4, the results demonstrate strong alignment between cross-validation (CV) means and test accuracies across all models, with differences consistently within one standard deviation. This alignment, coupled with relatively low standard deviations (particularly for LR and RF), indicates that the models maintain consistent performance across different data subsets and do not exhibit significant overfitting. The highest performing model, KNN, achieved a test accuracy of 98.67 % with a cross-validation mean of 96.80 %, suggesting excellent generalization capabilities despite its high accuracy.

4.2. Performance analysis of individual feature sets

In the analysis of individual feature sets, as detailed in Table 2, distinct variations in model performance were observed across different machine learning techniques. The accuracy scores for individual features—Size, Shape, and Texture—indicate that each set uniquely contributes to the classification process. The Size feature set consistently showed high accuracy across all models, with an average accuracy of 92.67 %, suggesting that size attributes are highly informative for the models used. Shape features also performed well, achieving an average accuracy of 91.55 %, with the highest accuracy noted for the KNN model. This reflects the effectiveness of shape attributes in differentiating between classes in a spatial context. Conversely, texture features had the lowest impact on model accuracy with an average of 87.34 %, indicating that texture alone may be less discriminative for the models applied compared to size and shape. This section underscores the importance of feature selection in machine learning and hints at the potential benefits of feature combination to enhance model robustness and accuracy.

4.3. Efficacy of combined feature sets

The analysis focuses on the performance enhancement gained through the integration of combined feature sets in machine learning models. The data provided in Table 3 offers a detailed comparative assessment across various classifiers, including SVM, RF, GNB, LR, KNN, and EL, utilizing combinations such as Size + Shape, Size + Texture, Shape + Texture, and Size + Shape + Texture. This comparison highlights the substantial impact that feature integration has on model accuracy. From the results, it is evident that combining features typically results in superior model performance compared to utilizing individual feature sets alone. The combination of Size + Shape + Texture emerges as the most effective, achieving the highest average accuracy of 96.44 %. Notably, the KNN classifier reached 98.67 % accuracy with this feature set, illustrating the significant potential of leveraging comprehensive feature sets in complex classification scenarios. Furthermore, the combination of Shape + Texture also shows promising results (96.06 % average accuracy across models), closely approaching the highest average accuracy and emphasizing the valuable interplay between these two feature types. This suggests that even partial combinations of features can capture critical information, significantly enhancing classifier performance.

4.4. Detailed analysis of optimal feature set

The feature selection process revealed distinct preferences across the various classification models regarding the optimal feature subsets. In Table 5, it is evident that while models like GNB and EL demonstrated superior performance with the complete feature set, others exhibited enhanced efficacy with specifically tailored feature combinations. As Table 5 illustrates, SVM and RF achieved optimal results with condensed feature sets, while KNN demonstrated remarkable classification capability utilizing just a small subset of carefully selected features. This variability in feature preference across models underscores the importance of model-specific feature optimization rather than a one-size-fits-all approach.

The Size + Shape + Texture combination achieved the highest average accuracy of 96.44 % across models (Table 3), likely because it captures a comprehensive set of discriminative characteristics—size for physical dimensions, shape for structural nuances, and texture for surface variations—enabling robust differentiation across Uma, KCP-1, and Jyothi varieties. Conversely, the Size + Shape combination excelled for SVM (97.34 % accuracy, Table 3), possibly due to SVM's ability to effectively model linear and non-linear relationships in a reduced feature space, where size and shape features provide sufficient discriminative power without the additional complexity of texture features, which may introduce noise for this model.

Analysis of the performance metrics presented in Table 5 across the three rice varieties reveals compelling insights into classification effectiveness. As shown in Table 5, for the Uma variety, most classifiers achieved exceptional precision and recall rates, with several models attaining perfect scores (100 %). The KCP-1 variety, as indicated in Table 5, saw perfect precision from SVM, LR, KNN, and EL (100 %), though recall performance showed slight variability, ranging from 92 % to 96 %. For the Jyothi variety, Table 5 demonstrates that precision remained strong across all models (93 %−100 %), while recall exhibited the most variation, with RF achieving 84 % and others reaching up to 100 %. The F1-scores displayed in Table 5, which balance precision and recall considerations, further validate the robustness of the optimized feature sets, with KNN demonstrating the most consistent F1-score performance (98–100 %) across all varieties. The strong performance of KNN with a smaller feature subset (Table 5) may be attributed to its reliance on proximity-based classification, where a compact set of highly discriminative features minimizes distance calculation errors, enhancing accuracy across all varieties. The consistently high-performance metrics shown in Table 5 across multiple evaluation metrics and different rice varieties substantiates the generalizability of the identified optimal feature sets for red rice classification.

4.5. Hyper-parameter tuning

Hyper-parameter tuning was performed to optimize classification performance. For SVM, regularization strength, kernel type, and scale adjustment were optimized, selecting a polynomial kernel with higher

Table 6
Hyper-parameter tuning.

Model	Best Hyper-Parameter	Description of Terms
SVM	$C = 100$, kernel=poly, gamma=scale	Regularization parameter, Kernel type, Influence range (gamma)
RF	n_estimators=200, max_depth=10, min_samples_split=2, min_samples_leaf=1, bootstrap=True	Number of trees, Depth, Split criteria, Leaf criteria, Bootstrap sampling
GNB	var_smoothing=1e-09	Smoothing parameter to handle low variance
LR	$C = 1$, penalty=l1, solver=liblinear	Regularization parameter, Regularization type (penalty), Optimization solver
KNN	n_neighbors=5, weights=uniform, $p = 2$	Nearest neighbours, Weighting scheme, Distance metric ($p = 2$ for Euclidean)

regularization for improved decision boundaries. RF tuning adjusted the number of trees, tree depth, minimum split criteria, and bootstrap sampling, achieving a balanced trade-off between complexity and generalization. GNB refinement focused on smoothing parameter optimization for handling low variance. LR was optimized by adjusting regularization strength, type, and solver, selecting sparse regularization with moderate strength. KNN tuning optimized the number of neighbours, distance metric, and weighting function, ensuring well-balanced classification. The best hyper-parameters for each model and the description of the terms are summarized in Table 6.

4.6. Statistical significance analysis

Statistical significance of model performance differences was evaluated using the Friedman test, which revealed significant overall differences between models ($\chi^2 = 19.971$, $p = 0.001$). Pairwise comparisons using paired *t*-tests demonstrated that KNN, the best-performing model, significantly outperformed SVM ($p = 0.003$), RF ($p = 0.008$), and GNB ($p < 0.001$). However, no statistically significant differences were observed between KNN and the next-best performers, LR ($p = 0.695$) and EL ($p = 0.156$). Effect size analysis using Cohen's d revealed large effects when comparing KNN with RF ($d = 1.919$) and GNB ($d = 1.377$), medium effect with EL ($d = 0.529$), small effect with SVM ($d = 0.384$), and negligible effect with LR ($d = 0.160$). These results confirm that KNN's superior performance is statistically robust and not due to random variation.

4.7. Analysis of confusion matrices

Analysis of the confusion matrices presented in Fig. 10(a) to Fig. 10(f) reveals significant insights into classification performance across the three rice varieties. The diagonal elements show correct classifications across all models, with SVM correctly identifying all 25 Uma samples and the KNN classifier successfully recognizing 24 Jyothi samples. Examination of off-diagonal elements highlights that KCP-1 and Jyothi varieties experience more frequent cross-classification incidents, as evidenced in Fig. 10(c), Fig. 10(d), and Fig. 10(f), where several models demonstrate confusion between these two varieties. In contrast, the Uma variety consistently shows minimal misclassification across all models, with Fig. 10(a) through Fig. 10(f) demonstrating negligible instances of Uma being confused with other varieties, indicating that Uma possesses more distinctive morphological characteristics that are effectively captured by the selected features, while KCP-1 and Jyothi share greater similarity in their feature representation.

4.8. Evaluation of ROC and AUC metrics

The ROC curves presented in Fig. 11(a) to Fig. 11(f) illustrate the classification models' ability to discriminate between classes across various threshold settings by plotting the true positive rate against the false positive rate. The AUC metric quantifies this discriminative power, with values approaching 1.0 indicating excellent classification performance and values near 0.5 suggesting performance no better than

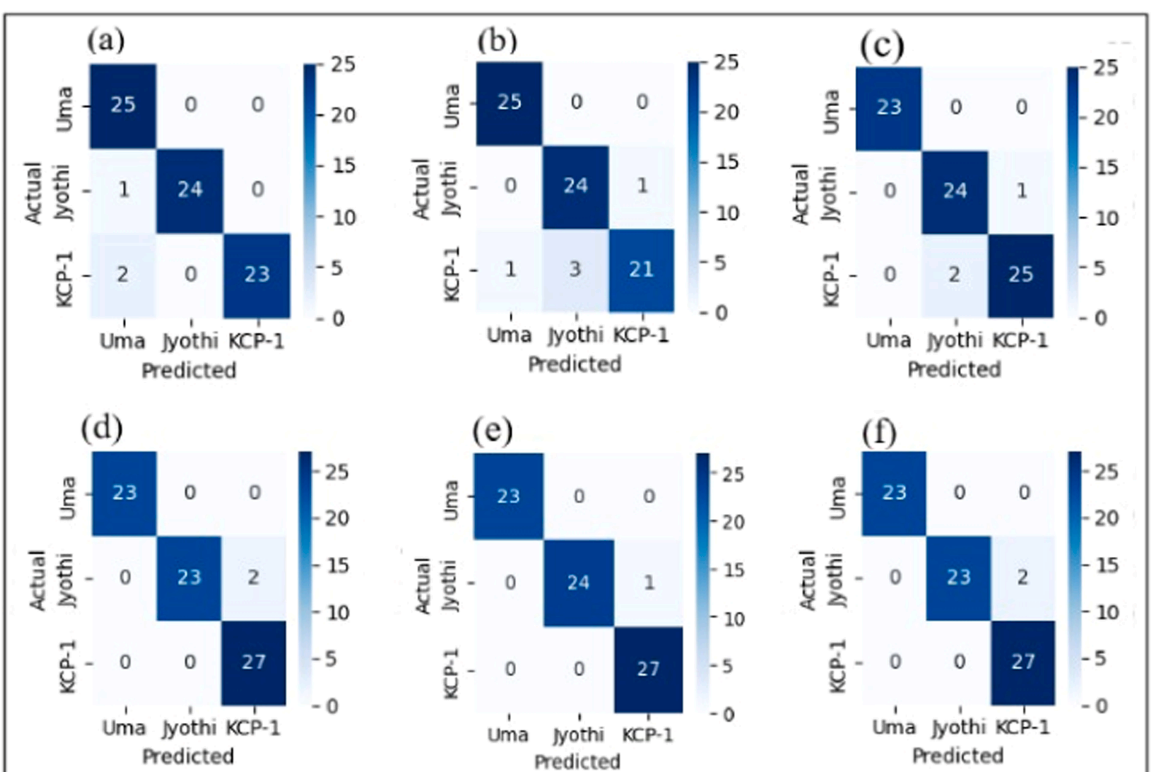

Fig. 10. Confusion Matrices. (a) SVM (b) RF (c) GNB (d) LR (e) KNN (f) EL.

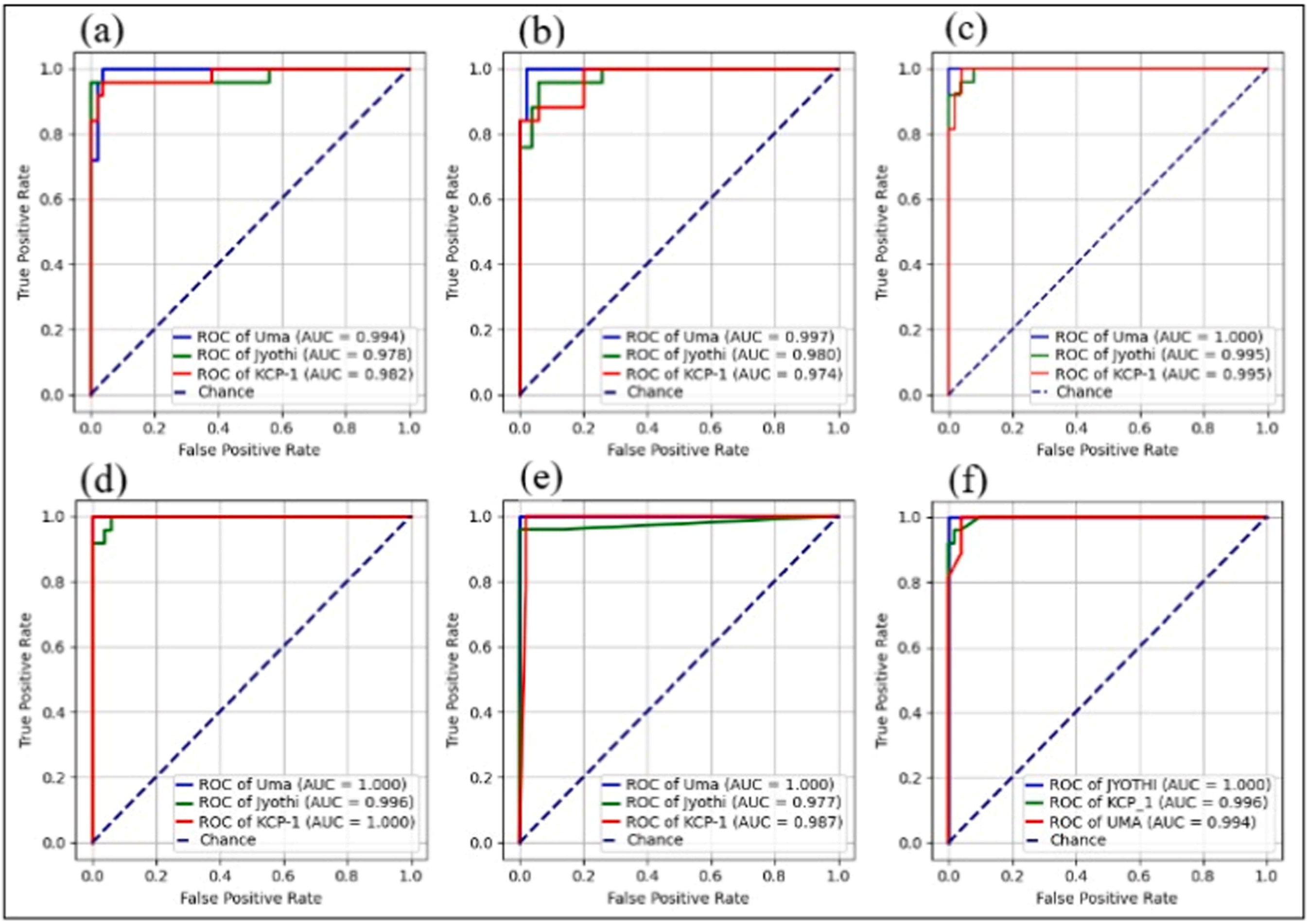

Fig. 11. ROC Curves and AUC Metrics. (a) SVM (b) RF (c) GNB (d) LR (e) KNN (f) EL.

random chance. As demonstrated across all subfigures, the models consistently achieve remarkably high AUC values for all three rice varieties, with many curves reaching or approaching the perfect classification score of 1.0. The consistently steep ascent of the ROC curves toward the top-left corner of each plot indicates that the classifiers maintain high sensitivity while minimizing false positives. This exceptional performance across all models validates the effectiveness of the selected feature combinations in capturing the distinctive characteristics of each rice variety. Near-perfect AUC scores, especially for Uma, confirm the robust classification framework's ability to reliably distinguish Uma, KCP-1, and Jyothi red rice varieties.

5. Conclusion and future scope

This study has successfully developed a novel approach for red rice classification, addressing a significant research gap where no prior studies or datasets exist for red rice variety classification. Through rigorous hyperparameter tuning across six diverse machine learning models, including SVM, RF, GNB, LR, KNN, and EL, combined with strategic feature selection using Recursive and Backward Feature Elimination techniques, the research achieved exceptional classification performance with KNN reaching 98.67 % accuracy. The combination of size, shape, and texture features was found to be the most effective feature set, achieving an average accuracy of 96.44 % across all models, though the shape and texture combination performed remarkably close with an average accuracy of 96.06 %. Comprehensive evaluation based on multiple indicators shows that the study has achieved excellent classification performance. Implementation of k-fold cross-validation with $k = 5$ confirmed robust generalization capabilities across all models without significant overfitting, validating the effectiveness of the optimized feature combinations. The comprehensive evaluation methodology and impressive performance metrics establish a solid foundation for automated red rice classification systems that can enhance agricultural quality control processes in South Indian rice production.

For **future research**, several enhancements are proposed to further strengthen and broaden the scope of this work. First, expanding the dataset size is crucial to improve the model's robustness and its ability to generalize to more diverse and unseen samples. Second, incorporating additional red rice varieties beyond Uma, Jyothi, and KCP-1 could help in developing a more comprehensive and scalable classification system. Lastly, with larger datasets, deep learning techniques such as Convolutional Neural Networks (CNNs) can be explored, which may uncover more complex patterns and further boost classification accuracy beyond traditional machine learning methods.

Ethical statement

This study does not involve any experiments on humans or animals, and no ethical approval was required. Therefore, an ethical statement is not applicable.

Consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Funding

Not applicable.

Code availability (software application or custom code)

Not Applicable.

CRediT authorship contribution statement

Suma D: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Narendra V G:** Writing – review & editing, Supervision. **Darshan Holla M:** Writing – review & editing, Software. **Shreyas:** Writing – review & editing, Software. **Raviraja Holla M:** Writing – original draft, Software, Methodology.

Declaration of competing interest

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.

Acknowledgement

The authors are thankful to Dr. Sridevi Jakkaral, Assistant Professor with Specialization in Rice Breeding for Quality and Salinity, Department of Plant Science Comprising Genetics and Plant Breeding, Plant Biotechnology, Crop Physiology, and Biochemistry, Zonal Agricultural & Horticultural Research Station, Brahmavar, Udupi District, Karnataka, India, for generously providing labelled rice grain varieties that were essential for our research work.

Data availability

Data will be made available on request.

References

[1] AICRIP: KCP-1. http://aicrip-intranet.in/Documents/AicripSite/Centers/14%20Brahmavar.pdf. Accessed 4 Feb 2025.
[2] A.S. Husain, Impact of the Rice variety, Uma (Mo16) On Farmers, Department of Agricultural Extension, College of Horticulture, Vellanikkara, 2015. M.S. thesis.
[3] Bio basics: uma Kerala red rice - parboiled. https://biobasics.org/products/uma-kerala-red-rice-parboiled?variant=37905433559199. Accessed 4 Feb 2025.
[4] B. Arora, N. Bhagat, L. Saritha, S. Arcot, Rice grain classification using image processing & machine learning techniques, in: Proceedings of International Conference on Inventive Computation Technologies (ICICT), 2020, pp. 205–208.
[5] N. Nagoda, L. Ranathunga, Rice sample segmentation and classification using image processing and support vector machine, in: Proceedings of IEEE 13th International Conference on Industrial and Information Systems (ICIIS), 2018, pp. 179–184.
[6] K. Kiratiratanapruk, et al., Development of paddy rice seed classification process using machine learning techniques for automatic grading machine, J. Sensors 2020 (2020) 1–15.
[7] K. Aukkapinyo, S. Sawangwong, P. Pooyoi, W. Kusakunniran, Localization and classification of rice-grain images using region proposals-based convolutional neural network, Int. J. Auto. Comp. 17 (2) (2020) 233–246.
[8] C.V. Krishna, et al., Quality analysis of rice grains using morphological techniques, in: Proceedings of International Conference on Communication, Computing and Internet of Things (IC3IoT), 2022, pp. 1–5.
[9] S. Ibrahim, N.A. Zulkifli, N. Sabri, A.A. Shari, M.R.M. Noordin, Rice grain classification using multi-class support vector machine (SVM), IAES Int. J. Artif. Intell. 8 (3) (2019) 215–222.
[10] R. Ruslan, S. Khairunniza-Bejo, M. Jahari, M.F. Ibrahim, Weedy rice classification using image processing and a machine learning approach, Agriculture 12 (5) (2022) 645–658.
[11] K.R. Singh, S. Chaudhury, A cascade network for the classification of rice grain based on single rice kernel, Complex Intell. Syst. 6 (2020) 321–334.
[12] I. Cinar, M. Koklu, Identification of rice varieties using machine learning algorithms, J. Agri. Sci. 28 (2) (2022) 9–17.
[13] F. Farahnakian, J. Sheikh, F. Farahnakian, J. Heikkonen, A comparative study of state-of-the-art deep learning architectures for rice grain classification, J. Agri. Food Res. 15 (2024) 100890.
[14] T.T.K. Nga, T.V. Pham, D.M. Tam, I. Koo, V.Y. Mariano, T. Do-Hong, Combining binary particle swarm optimization with support vector machine for enhancing rice varieties classification accuracy, IEEE Access 9 (2021) 66062–66078.
[15] T.Y. Kuo, C.L. Chung, S.Y. Chen, H.A. Lin, Y.F. Kuo, Identifying rice grains using image analysis and sparse-representation-based classification, Comp. Elect. Agri 127 (2016) 716–725.
[16] L.O. Carneiro, et al., Characterizing and predicting the quality of milled rice grains using machine learning models, AgriEngineering 5 (3) (2023) 1196–1215.
[17] S.B. Ahmed, S.F. Ali, A.Z. Khan, On the frontiers of rice grain analysis, classification and quality grading: a review, IEEE Access 9 (2021) 160779–160796.
[18] W. Srimulyani, A. Musdholifah, Identification of rice variety using geometric features and neural network, Indonesian J.Comp. Cybern. Syst. 13 (3) (2019) 301–312.
[19] K.R. Singh, S. Chaudhury, Efficient technique for rice grain classification using back-propagation neural network and wavelet decomposition, IET Comp. Vision 10 (8) (2016) 780–787.
[20] A. Aznan, C. Gonzalez Viejo, A. Pang, S. Fuentes, Computer vision and machine learning analysis of commercial rice grains: a potential digital approach for consumer perception studies, Sensors 21 (19) (2021) 6354–6370.
[21] A.S. Hamzah, A. Mohamed, Classification of white rice grain quality using ANN: a review, IAES Int. J. Arti. Int. 9 (4) (2020) 600–608.
[22] S. Sultana, M. Faruque, M.R. Islam, Rice grain quality parameters and determination tools: a review on the current developments and future prospects, Int. J. Food Properties 25 (1) (2022) 1063–1078.
[23] M.T. Ahad, Y. Li, B. Song, T. Bhuiyan, Comparison of CNN-based deep learning architectures for rice diseases classification, Artif. Intell. Agri. 9 (2023) 22–35.
[24] N. Tran-Thi-Kim, T. Pham-Viet, I. Koo, V. Mariano, T. Do-Hong, Enhancing the classification accuracy of rice varieties by using convolutional neural networks, Int. J. Elec. Electro. Eng. Telecom. 12 (2) (2023) 150–160.
[25] B. Patel, A. Sharaff, Rice variety classification & yield prediction using semantic segmentation of agro-morphological characteristics, Multimed. Tools Appl. 82 (29) (2023) 45567–45584.
[26] C. Kurade, et al., An automated image processing module for quality evaluation of milled rice, Foods 12 (6) (2023) 1273–1285.
[27] C. Deepika, R. Gnanamalar, K. Thangaraj, N. Revathy, A. Karthikeyan, Linear discriminant analysis of grain quality traits in rice (oryza sativa l.) using the digital imaging technique, J. Cereal Sci. 109 (2023) 103609.
[28] T.M. Philip, H.B. Anita, Rice grain classification using fourier transform and morphological features, Indian J. Sci.Tech. 10 (14) (2017) 1–6.