International Journal of Robotics and Control Systems
Vol. 5, No. 2, 2025, pp. 661-677
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

Efficient Detection Classifiers for Genetically-Modified Golden Rice Via Machine Learning

Joshua Balistoy Gutierrez [a,1,*], Edwin Romeroso Arboleda [a,2]

[a] Cavite State University, Indang, Cavite, 4122, Philippines
[1] main.joshua.gutierrez@cvu.edu.ph; [2] edwin.r.arboleda@cvsu.edu.ph
*Corresponding Author

ARTICLE INFO

Article history
Received October 28, 2024
Revised December 07, 2024
Accepted February 03, 2025

Keywords
GMO Golden Rice;
Machine Learning;
Image Processing;
MATLAB;
WEKA

ABSTRACT

Rice is a staple food for over half of the global population, especially in the Philippines. However, traditional rice lacks essential micronutrients like vitamin A, contributing to widespread Vitamin A Deficiency (VAD). Golden Rice was developed to combat VAD, and this is biofortified with beta-carotene, a precursor of Vitamin A. However, concerns about cross-contamination, food safety, and ethics have emerged. Current GMO detection methods, such as PCR and ELISA, are not ideal for large-scale or on-site use since these are intended to be performed inside laboratory and requires technical expertise. This study presents a novel machine learning (ML)-based approach for the detection of genetically modified Golden Rice using RGB image data and several classification models as an efficient, rapid, non-destructive method to detect GMO Golden Rice. Two datasets of rice images (340 samples of GMO Golden Rice and 340 samples of Traditional Rice) were processed and split for training and testing (80-20 ratio). This study found that WEKA's Random Tree and MATLAB's Trilayered Neural Network achieved 100% accuracy in detecting GMO Golden Rice, with the fastest computational efficiency in their respective platforms. Additional metrics, such as Precision and Recall, further verified the robustness of these classifiers. This research lays the foundation for developing portable, field-deployable detection tools to empower farmers and regulators while enhancing consumer trust in GMO labeling. Furthermore, the application of ML to GMO rice detection opens new possibilities for biofortified crop monitoring. Future work may explore integrating additional rice features and GMO varieties, validating the results, and expanding this methodology to other GMO rice variants and hybrid varieties to further enhance detection accuracy and scalability.

1. Introduction

More than 3.5 billion people eat rice as part of a staple diet and as a primary source of energy [1]. In the Philippines, rice is not just a dietary staple; it has been part of our country's rich culture, as showcased by our majestic Banaue Rice Terraces. Despite that, rice has gained relevance in the country's history and harsh economy, from being one of the top exporters in the early 1970's, to being the top importer in 2023. Albeit, with approximately 2.5 million hectares of rice fields, 2.4 million Filipino farmers are making a living in the biggest agricultural sector of our country [2]. Be that as it may, traditional rice varieties do not satisfy nutritional needs as they usually lack essential micronutrients, particularly vitamins A, C, and iron. In the 2020 report of the World Health Organization (WHO), the deficiency in vitamin A became one of the public health concerns in the country; this primarily affected children and pregnant women [3]. Vitamin A Deficiency (VAD) is known to cause many health issues, such as blindness, weak immune systems, and an increased mortality rate. 2 million children and maternal deaths were recorded annually worldwide when VAD was first recognized in 1990's, making VAD responsible for the 23-34% of all deaths among children worldwide [4]. In 2013, the mortality rate attributed to VAD reduced to 1.8% in the country [4]. It was estimated that there will be 5,886 VAD-related deaths in the Philippines by 2019 [5]. According to the 2018-2019 Expanded National Nutrition Survey (ENNS) of the Department of Science and Technology's Food and Nutrition Research Institute (DOST-FNRI), 15.5% (roughly one out of six) of Filipino infants and children under 5 years old were reportedly affected by VAD; and across all age/physiologic groups, a VAD prevalence of 22.4% was noted in the poorest wealth quintile which is deemed severe [6]. Moreover, COVID-19 Pandemic hindered the initiatives and programs against VAD which induced mortality rate globally [7].

The development of golden rice was the response of the International Rice Research Institute (IRRI) to combat VAD. This kind of rice is biofortified with beta-carotene, a precursor to vitamin A, specially engineered to satisfy 30-50% of the estimated average requirement (EAR) for vitamin A in children [8], [9]. Since 2001, initiatives have been made locally to address VAD issue, particularly when Philippine Rice Research Institute has acquired license for the development of its own Golden Rice. However, it was only after twenty years when Philippines finally authorized the commercial planting of golden rice, specifically the GR2E variant, the first in the world—hailing the action as an essential milestone in the fight against malnutrition [10]. Anent to this, the "Malusog Rice" Program was initiated by the government to address micronutrient deficiencies while enhancing food security through local agriculture. "Malusog Rice", also known as 'Healthy Rice' is a genetically modified golden rice additionally enriched with iron and zinc. In 2022, our country made its first large-scale harvest of golden rice, and by 2023, the growing interest in biofortified crops delighted 100,000 farmers to cultivate golden rice [11].

Despite the uplifting developments, the adoption and distribution of Golden Rice and other GMO's have since ignited public skepticism and scrutiny. Environmental and public safety groups oppose Golden Rice initiatives simply because it is genetically modified, fearing cross-contamination [12]. Albeit, a report revealed that Golden Rice cannot just cross-contaminate other rice varieties, and the chance of 'gene flow' across rice variants is low because rice is itself self-pollinating [13]. Also, the safety for consumption of Golden Rice has been confirmed by the regulatory authorities across the world, not just in the Philippines [14]. Another issue that concerns the consumers is fraudulent practice of rice mixing in the market as these compromises not the biological aspect of the rice but rather the quality as crop and staple food [15]. With this, strict labeling regulations, careful monitoring measures and anti-fraudulent techniques [16] are implemented to secure genetically-modified crops, addressing food safety, authenticity, and ethical implications. It is paramount to establish methods for identifying and labeling GMO products to eliminate apprehension among consumers, is often attributed to a lack of accessibility to reliable information [11]. Equally important, rice mixing between GMO and non-GMO crops could negatively affect not only the environment but also the general market, particularly those that cater to non-GMO products [17]. These issues highlight the need to develop a reliable detection tool that can differentiate between GMO and non-GMO crops, ensuring regulatory compliance, environmental protection, and addressing consumer concerns while safeguarding market integrity.

Polymerase chain reaction (PCR) and enzyme-linked immunosorbent assay (ELISA) are the standard identification methods for detecting genetically modified crops, and these are ordinarily performed inside laboratory premises along with specialized equipment and done by trained personnel [18]-[20]. These methods, which directly target the chemical properties of rice, are unsuitable for practical application, especially for large-scale and on-site use. Stressing the practical need for a more accessible, rapid, cost-effective, and non-destructive detection method that will not require extensive expertise and laboratory procedures. Consequently, addressing this technological gap will not only simplify the process of regulatory compliance but also empower local farmers and consumers with information about the rice they are purchasing. Nowadays, machine learning methods have emerged as a promising and convenient approach for the classification of crops, such as rice. Taking advantage of its effectiveness, the detection between traditional rice and GMO Golden Rice is a safe bet, especially when equipped for analyzing its RGB histogram characteristics. Studies have demonstrated the reliability of machine learning models in classifying traditional rice varieties, such as Arborio, Basmati, Ipsala, Jasmine, and Karacadag [21]. Platforms such as MATLAB Classification Learner, WEKA, and Python-based libraries are commonly used in these types of projects and offer the best models for the intended detection applications.

In this context, RGB histogram analysis paired with ML emerges as a promising, practical approach for GMO detection. Unlike traditional methods, RGB histogram analysis focuses on analyzing color distribution and intensity, which could reveal differences between biofortified and non-GMO rice. Machine learning models, already proven effective in rice classification through studies utilizing platforms such as MATLAB, WEKA, and Python-based libraries, offer an adaptable and efficient means to process these visual data, supporting large-scale, field-level detection of GMO rice. Thus, this study aims to explore the potential of RGB histogram analysis and machine learning techniques as a reliable, accessible alternative for GMO detection, providing a feasible solution for rural farmers and consumers in the Philippines. In response to this gap, the main goal of this research is to train and evaluate available machine learning models for identifying the most efficient classifier for the detection of GMO Golden Rice, focusing on overall accuracy and computational efficiency. Specifically, the objectives are as follows: (1) to train all available machine learning classifiers using RGB histogram data of GMO and non-GMO rice samples and assess their accuracy in classification task; (2) to identify the models with the highest accuracy for GMO Golden Rice detection, assessing their robustness and generalization; (3) to compare the training and testing efficiency of the implemented models, aiming to identify the fastest model suitable for practical application; and (4) to evaluate the performance of the models on both MATLAB and WEKA, analyzing differences in accuracy and training efficiency across platforms.

This study demonstrates the potential of machine learning classifiers in the efficient detection of genetically modified Golden Rice, a breakthrough approach that could significantly enhance food security and address global nutritional deficiencies, particularly in rice-dependent regions like the Philippines. By leveraging advanced technology, the research explores how these classifiers can offer a reliable, scalable solution for monitoring the presence of genetically modified crops, ensuring compliance with regulatory standards, and fostering sustainable agricultural practices.

2. Literature Review

2.1. Machine Learning Approach

Machine Learning (ML) has been a prominent approach for rice classification, due to its flexibility to assess diverse features particularly the color, morphology, texture, and spectral features. Feature Extraction Method is used by traditional ML to extract grain characteristics and become basis for training models. However, Feature Ranking Algorithms, such as ANOVA, MRMR, and ReliefF, are used complementarily to give priority to relevant features for better interpretability and optimization of the classifier. ANOVA-based ML model that effectively captured compositional differences in rice, reported 93% accuracy [23]. Similarly, an accuracy of 87% was achieved using ML algorithms that prioritized its shape and texture with the help of Python's scikit-learn [20]. On an expanded focus, Multi-feature ML models trained also with Python's scikit-learn, additionally utilized color on top of shape and texture, gained 91.5% accuracy [24]. With MATLAB, Backpropagation Neural Network earned 88% accuracy [25]; while using grain surface texture with MATLAB's Vision Toolbox coincidentally gained 88% as well [26]. These studies collectively prove ML's capability to recognize traits and markers especially beneficial for recognition of biofortified crops. The classification of GMO Golden Rice is proven feasible when ML-based feature extraction methods combine with tools like MATLAB and WEKA.

2.2. Deep Learning Approach

When it comes to utilizing images with complex data, Deep Learning (DL) algorithms are adopted for rice classification due to their adept capability to detect fine visual distinctions. adaptive neuro-fuzzy inference system (ANFIS) yielded 82.95% accuracy as its best [27]. Convolutional Neural Networks (CNN), equipped with MATLAB, achieved 91% accuracy in identifying fine visual differences in rice grains [28]. RiceNet, a CNN-based model developed in TensorFlow, garnered 94% accuracy when dealt with capturing morphological markers [29]. On a comparative study of DL methods for rice variety detection, that CNN models DenseNet21, ResNet50, NasNetLarge, and InceptionV3 recorded as high as 99.04%, 98.84%, 98.12%, 98.10% of accuracy respectively [30]. On a study that uses three-dimensional imaging to assess physical features of rice reached 95% accuracy [31]; while the use of hyperspectral imaging with 3D-2DCNN model in Keras gave 98.93% accuracy [32]. These findings show that DL, through CNN models, can perform effectively with complex images, and still provide an almost-perfect accuracy.

2.3. Hyperspectral Imaging Approach

Hyperspectral and Spectroscopic Imaging are also proven effective in analyzing optical and molecular level compositions of rice samples. Near Infrared Spectroscopy backed by SIMCA garnered 90.5% accuracy in MATLAB [33]. Hyperspectral imaging was also utilized with models combining ML and DL to achieve an accuracy of 94% in TensorFlow [34]. Moreover, FT-IR Spectroscopy achieved 90% accuracy when paired with sparse chemometric analysis [35]. However, hyperspectral imaging is implemented along with Self-Attention Mechanism within CNNs, it earned a significant accuracy of 97.5 [36]. These methods therefore reveal potential effectiveness in detection and analysis of genetically-modified crops.

2.4. Geographic Detection

In a field-level perspective of performing large-scale rice classification, Remote Sensing and Machine Vision Techniques have been explored and implemented. Sentinel-2 Multispectral Imagery assisted by QGIS earned 93% accuracy in classification of rice varieties [37]. In determination of the geographic origin of rice, apart from its compositional variations, Raman Spectroscopy achieved an amazing 65% accuracy [38]. Raman Spectroscopy could be implemented for GMO Golden Rice with its sensitivity to molecular composition, essential for quantization of beta-carotene levels. Machine vision methods can indeed detect visual traits of a subject at a reliable accuracy—stressing the potential of rice's RGB histogram to differentiate GMO rice from traditional variants.

3. Method

3.1. Framework

This IPO-based conceptual framework visually represents the step-by-step process of identifying the most efficient machine learning model for GMO Golden Rice detection. Fig. 1 begins with sample acquisition, where rice grains were collected and prepared for analysis. Next is image preprocessing, which includes segmentation and masking to isolate individual grains for RGB extraction. This preprocessing step is crucial as models would objectively rely on what they could extract from each sample image, and must be done before training of models in order to rightly obtain a reliable data. The RGB extraction phase generates averaged RGB data based on the color intensity values it obtained from each sample, forming the dataset for ML model building. Afterwards, the dataset preparation stage ensures that data is appropriately formatted and split into training and testing sets. This data feeds into the machine learning classifiers, where models are trained and tested to objectively evaluate performance. Finally, the models undergo ranking and comparison, using criteria particularly the accuracy and computational efficiency to identify the most effective classifiers. This framework highlights the systematic integration of image processing and machine learning, underscoring the importance of each step in achieving reliable classification results [39]. It also demonstrates the feasibility of using visual and computational tools to address a practical agricultural challenge and promote precision agriculture [40].

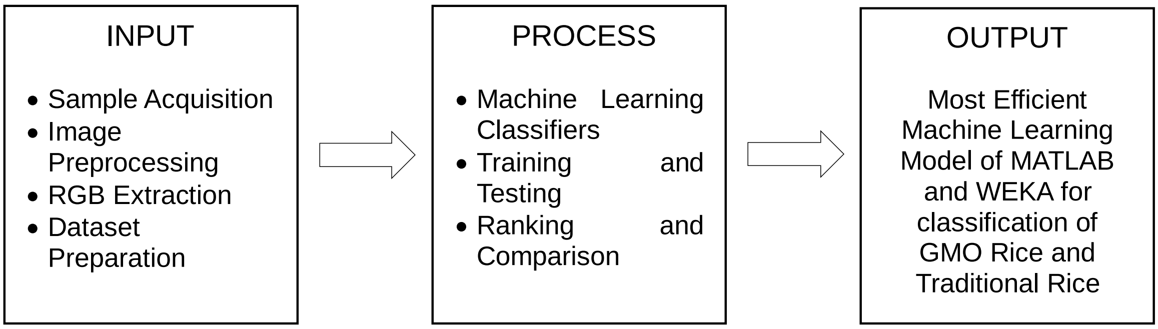


Fig. 1. The conceptual framework of the study

3.2. Materials, Equipment, and Software Used

3.2.1. Software

The author utilized both MATLAB (R2024b trial version) and WEKA (version 3.8.6) for training and testing the models, with MATLAB being used more due to its broader capabilities. MATLAB (short for MATrix LABoratory) is a high-level programming language and interactive environment developed by MathWorks Inc., primarily aimed at engineers and scientists. It supports a wide range of tasks, including data analysis, algorithm development, and the creation of models and applications. The trial version of MATLAB was sufficient for using the Classification Learner App, performing basic image preprocessing, and writing code. In contrast, WEKA (abbreviated from Waikato Environment for Knowledge Analysis), developed by Waikato University, is a comprehensive collection of machine learning algorithms designed for solving data mining problems and building machine learning models.

3.2.2. Hardware

Acer Extensa 15 215-51 laptop equipped with a Core i5 10th Gen. Processor was the computer used by the author for the time being, the processor was enough to perform the key processes as well as the technical paper writing process. In obtaining samples, images were captured using an Infinix Zero 40 5G smartphone equipped with camera that has highly reliable resolution of 108 Megapixels. As for the samples, the GMO rice variant was specifically the golden rice variant of Penaranda Rice, PSBRc82 as the only accessible GMO variant the author could gather. Jasmine Rice was selected to represent Traditional Rice due to its morphological resemblance with Penaranda Rice, only differing in odor and taste; as well as its prevalence in the local market as one of the consumer-friendly and popular traditional rice names.

3.3. Preprocessing, Data Acquisition, and Dataset Preparation

340 samples, each from GMO rice and traditional rice, were placed on white sheets of paper, as shown in Fig. 2. The samples were preprocessed using MATLAB's Image Segmenter, where each grain was masked for later analysis. The segmented images were then processed through a custom code that can identify and label each grain, extracting and averaging their RGB values. The code included a function to randomly select 80% of the processed data for the Training Dataset, while the remaining 20% was automatically apportioned to the Testing Dataset. This procedure ensured adherence to the conventional 80-20 distribution ratio. From the combined total of 680 rice samples, the preprocessed GMO and traditional rice grains were divided into separate training and testing datasets. Ultimately, 544 rice grain samples were used for training, while 136 samples were allocated for testing. Additionally, the code includes a function to save the datasets in both '.csv' (for MATLAB) and '.arff' (for WEKA) file formats. Note that in order to acquire.

Fig. 2 displays the raw images of both rice types before preprocessing through segmentation. Segmentation ensures each grain is isolated and prepared for accurate RGB value extraction, which minimizes errors caused by overlapping grains or background interference. By focusing on individual grains, the study ensures uniformity in data preparation, improving the reliability of machine learning training. This preprocessing step stresses the importance of isolating visual characteristics to capture unique color features, making it crucial for classification accuracy.

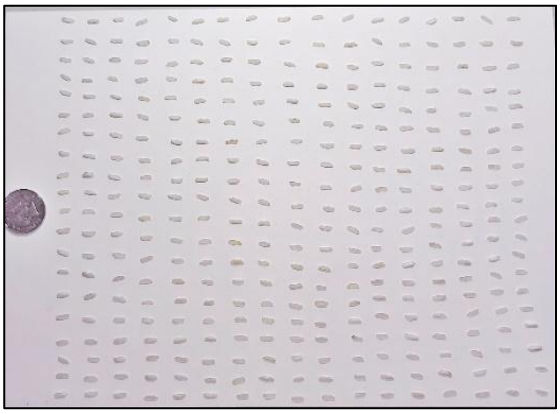
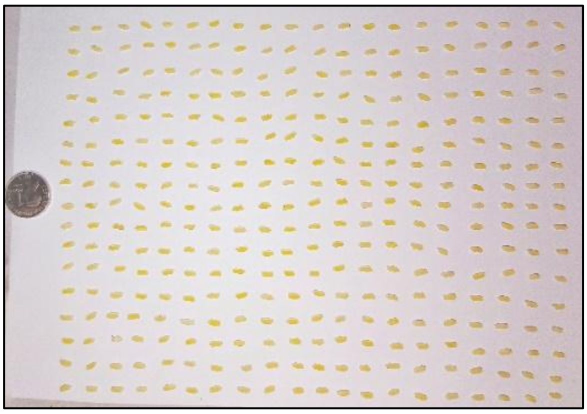



Fig. 2. The traditional rice (left) and the gmo golden rice (right)

Fig. 3 visually contrasts Jasmine rice, characterized by its lighter, almost white appearance, with GMO Golden Rice, which has a yellowish hue as manifested by its high carotenoid content [22], [8]. This difference in coloration stems from the genetic modification in Golden Rice to combat Vitamin A deficiency. The yellow-golden color is a unique identifier, which the study leverages through RGB histogram analysis. This visual disparity provides a foundational basis for differentiating between the two rice types, emphasizing the significance of visual data in automated GMO detection.


Fig. 3. The jasmine rice (left) and the gmo golden rice (right)

3.4. Model Training, Testing, and Evaluation

To conduct a comprehensive comparison of the performance of all ML classifiers, only the available models in MATLAB and WEKA were used for training, testing, and thorough evaluation. 5-fold cross-validation was also used across platforms to preserve consistency and fairness. The Cross-Validation method aims to protect data from overfitting by partitioning the training data through folds. For testing, 136 images from the test files of both Traditional Rice and GMO Golden Rice were used on both platforms. The models were then categorically evaluated and ranked according to the following essential criteria in sequence: 1.) Validation Accuracy, 2.) Testing Accuracy, and 3.) Training Speed.

4. Results and Discussion

4.1. RGB Histogram of Rice Samples

The RGB histogram in Fig. 4 reveals a distinct pattern between GMO Golden Rice and traditional rice. The GMO rice exhibits prominent peaks in the red and green channels, with a noticeable intensity in the blue channel at lower levels, consistent with its golden coloration. Conversely, traditional rice shows a more balanced RGB distribution but significantly lower blue intensity, reflecting its lighter appearance. These differences underscore the feasibility of using RGB histogram data as a feature for machine learning classifiers. The results demonstrate that color properties differentiate these rice types, objectively validating the study's approach.

Table 1 provides the measured ranges of color characteristics from both rice types, with traditional rice having higher red channel values but lower blue channel values than GMO Golden Rice. This numerical representation provides evidence for the visual distinctions observed in the histogram. The broader intensity range in GMO Golden Rice suggests more significant variability in its coloration, certainly due to the genetic modifications. These quantitative findings support the study's premise that RGB data can reliably distinguish GMO Golden Rice from traditional varieties, providing a robust feature set for machine learning models.

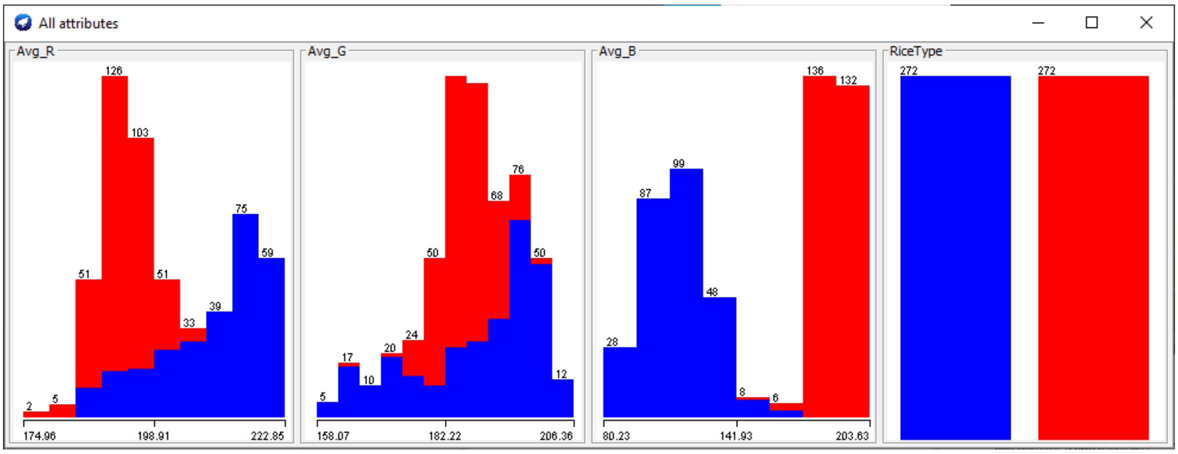


Fig. 4. RGB histogram data

Table 1. RGB histogram ranges

Channel	GMO Golden Rice	Traditional Rice
Red	175-205	183-223
Green	164-200	156-207
Blue	153-204	83-169

4.2. Comparison of the Accuracies of MATLAB CLA's Classifiers

Fig. 5 illustrates that nearly all classifiers in MATLAB's Classification Learner App achieve high validation accuracy, with several models reaching 100%. This consistency demonstrates MATLAB's robustness for the classification task and its ability to generalize well during training. The results indicate that the RGB histogram features are highly informative, allowing even basic classifiers to perform effectively. The result reinforces that MATLAB provides an efficient platform for training machine learning models for visual-based GMO detection.

Similar to the validation accuracy results, Fig. 6 shows that many classifiers achieve 100% accuracy during testing, confirming their ability to generalize well to unseen data. This highlights the reliability of the extracted RGB histogram features for differentiating between GMO Golden Rice and traditional rice. The ability to replicate high accuracy in both validation and testing phases ensures the practical applicability of these models, making them suitable for real-world deployment in GMO detection tasks.

4.3. Comparison of the Training Speed of MATLAB CLA's Classifiers

Fig. 7 compares the prediction speed (measured in observations per second) of various machine learning classifiers in MATLAB's Classification Learner App (CLA). Each bar represents a classifier, with higher bars indicating faster prediction speeds. Prediction speed is a critical metric for real-time applications, and this chart reveals significant variability among the classifiers. Certain models exhibit exceptionally high speeds, making them ideal for rapid field-level detection tasks. In this case, Quadratic SVM stood out for its ability to handle large datasets quickly, whereas others show moderate performance. This variation highlights the trade-offs between model complexity and efficiency, emphasizing the importance of selecting models tailored to specific deployment scenarios.

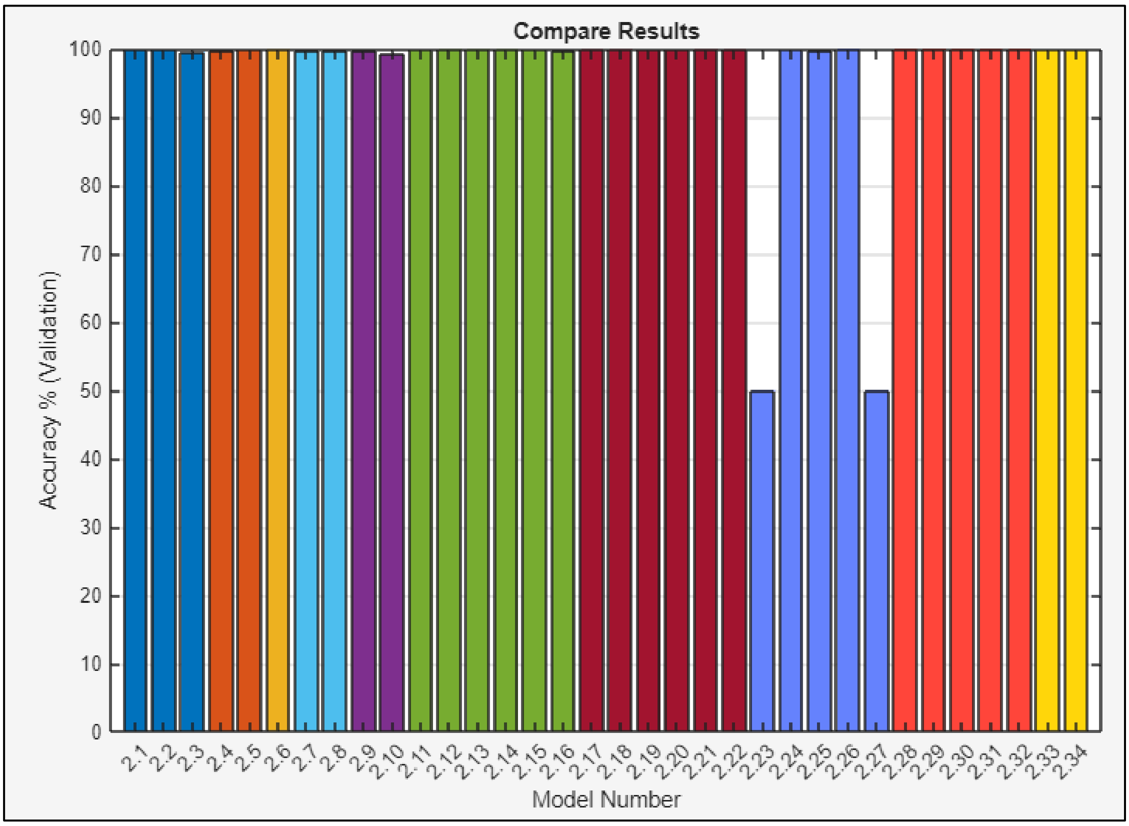


Fig. 5. Validation accuracy of MATLAB CLA's machine learning classifiers

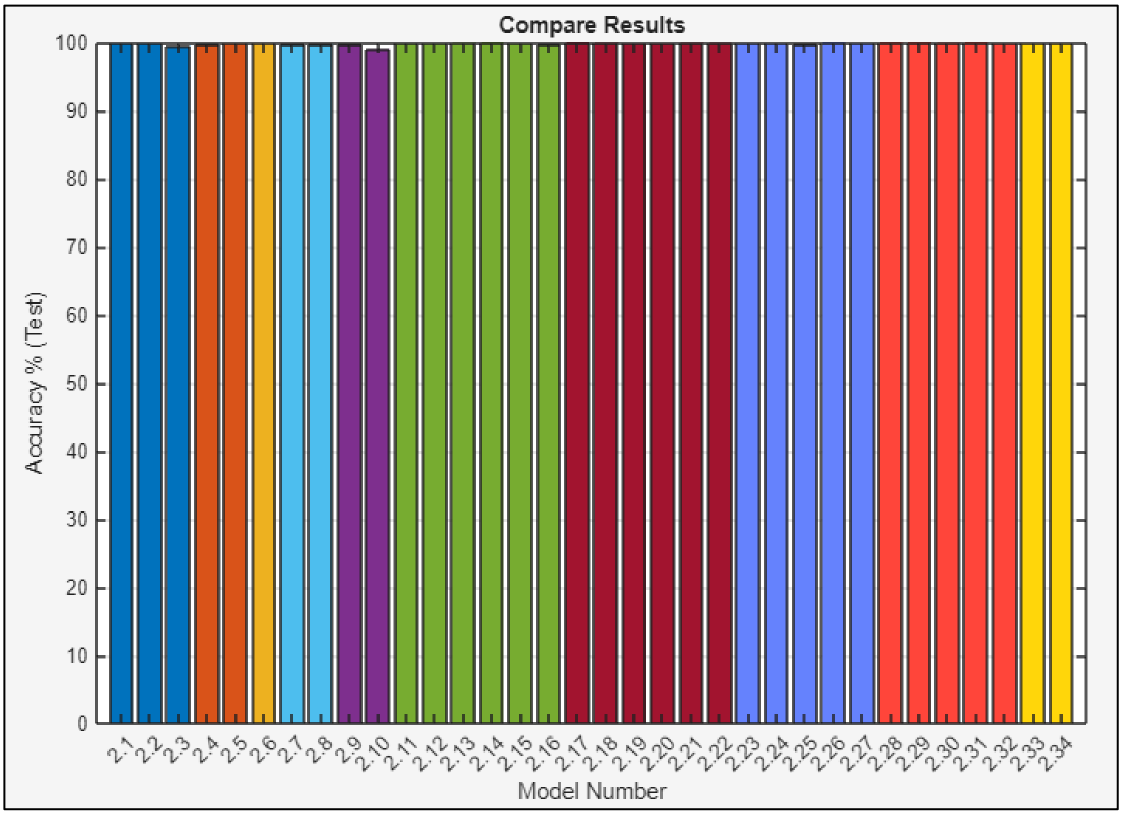


Fig. 6. Test accuracy of MATLAB CLA's machine learning classifiers

Fig. 8 shows the training time (in seconds) for each machine learning classifier using MATLAB's Classification Learner App. Each bar corresponds to a classifier, with taller bars representing longer training times. The training time provides insights into the computational demand of different classifiers. The chart shows that simpler models, such as Cosine KNN and Weighted KNN, require less time, while more complex models, such as the Subspace KNN, take significantly longer. For practical applications, particularly in resource-limited environments, models with shorter training times offer a significant advantage. However, for tasks where training occurs less frequently, the time requirement may be less critical if accuracy is prioritized. This balance between computational cost and performance should guide model selection.

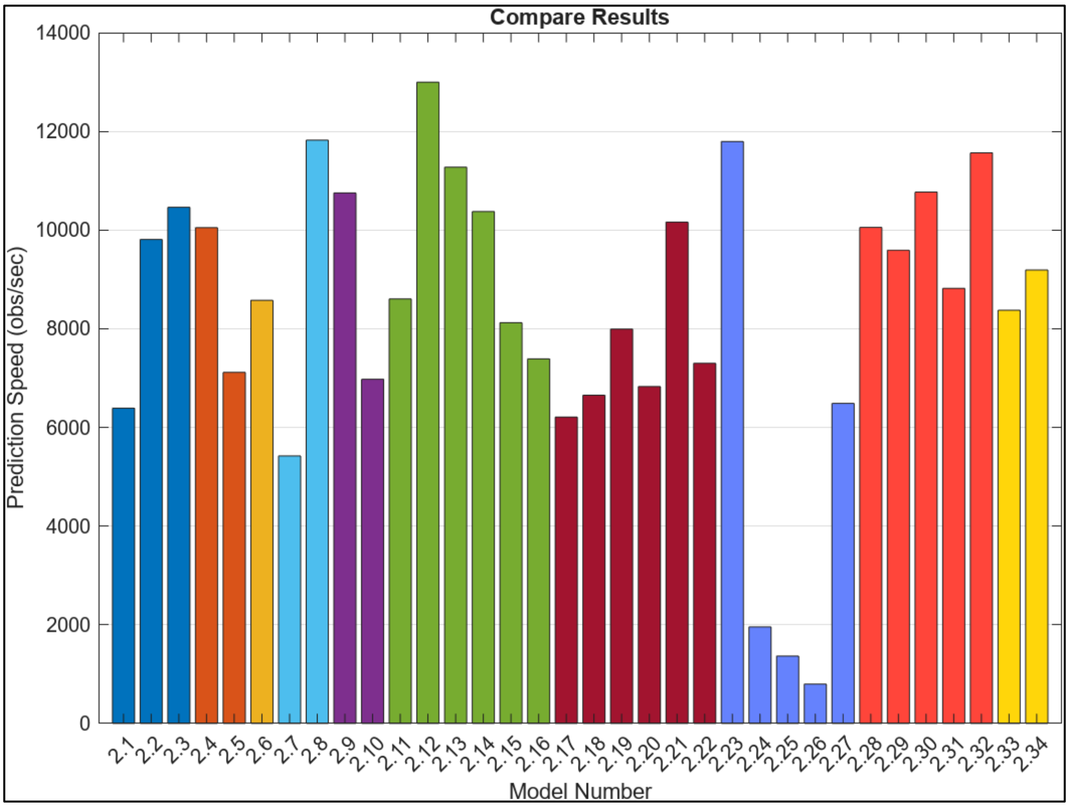


Fig. 7. Prediction speed of MATLAB CLA’s machine learning classifiers

4.4. Ranking of Models/Classifiers

Table 2 ranks MATLAB CLA classifiers based on their training and testing accuracy and training time. The Trilayered Neural Network stands out as the most efficient model, achieving perfect accuracy with the fastest training time of 1.03 seconds. Other models, such as Cubic SVM and Coarse KNN, also achieve 100% accuracy but require slightly longer training times. This ranking highlight MATLAB’s flexibility and computational efficiency, enabling users to choose models that best suit their needs. The Trilayered Neural Network’s performance demonstrates its suitability for both accuracy-critical and resource-constrained applications, particularly for GMO detection.

Table 3 presents the top-performing classifiers in WEKA, with Random Tree achieving perfect accuracy in both training and testing while requiring no measurable training time. Other models, like SMO Function and AdaBoost M1, also perform exceptionally well but with relatively longer training times, although almost negligible. These findings highlight WEKA’s computational efficiency, making it an excellent choice for resource-limited environments. With its negligible training time and high accuracy, Random Tree is a practical and scalable solution for field-level GMO detection.

4.5. Confusion Matrices of the Most Efficient ML Model of WEKA and MATLAB

Fig. 9 (Top) showcases the confusion matrix for MATLAB’s Trilayered Neural Network and reveals its 100% classification accuracy, with no misclassifications observed. This demonstrates the model's being not vulnerable to false positives in identifying both the GMO Golden Rice and the traditional rice. Such performance is essential in applications where false positives or negatives could have serious consequences, such as in regulatory compliance or consumer transparency. The summary metrics validate the Trilayered Neural Network as a reliable tool for GMO detection. Additionally, Trilayered Neural Network is a type of feedforward neural network with high flexibility because of its three-layered size settings, however it comes with difficult interpretability due its nonlinear functionality. MATLAB uses the algorithm of 'ficnet' wherein the first layer connects to the input data (predictor data), and each subsequent layer is connected to the previous one. Each layer multiplies the input by a weight matrix, adds a bias vector, and then applies an activation function. As output, the predicted labels are obtained alongside their classification scores (posterior probabilities) [41]. When it comes to practical application, Trilayered Neural Network could robustly handle complexities in rice features through nonlinear approach and still produce accurate prediction in fastest possible time, although relatively slower compared to WEKA's Random Tree.

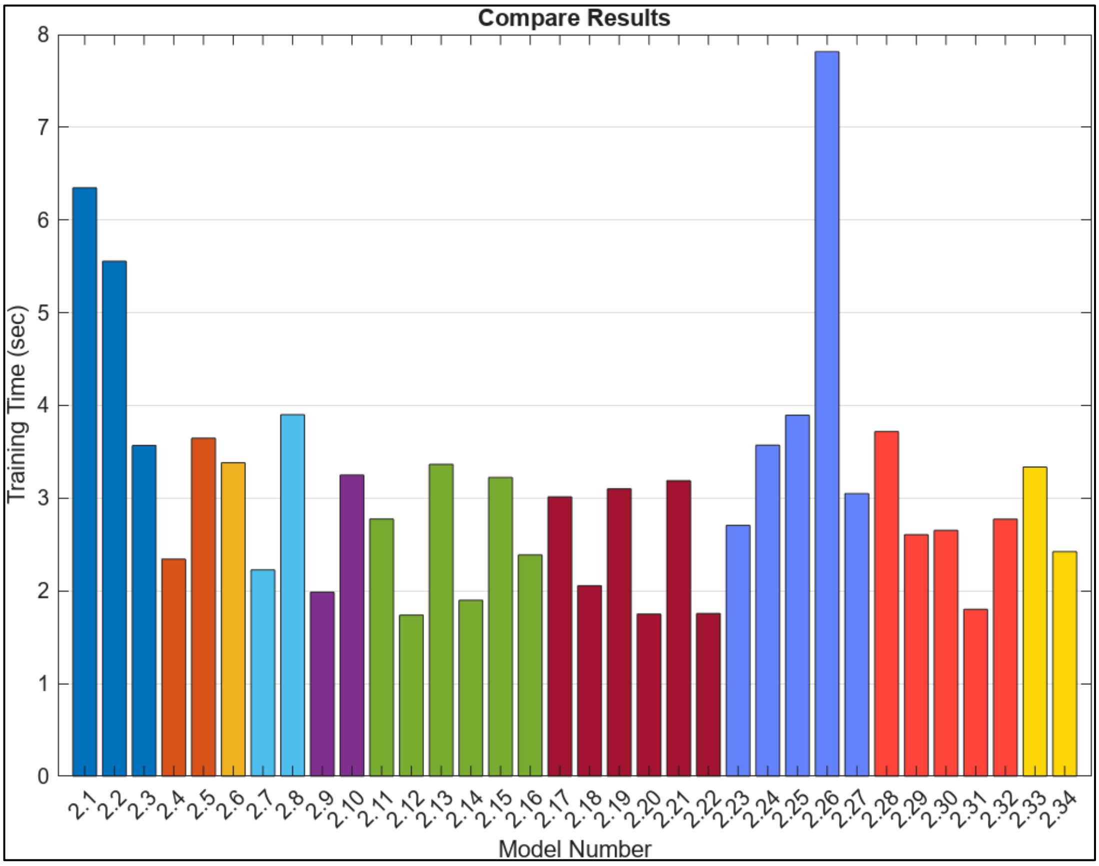


Fig. 8. Training time of MATLAB CLA's machine learning classifiers

Like in Fig. 9, WEKA's Random Tree confusion matrix in Fig. 10 (Top) shows perfect classification with no errors. This result mirrors MATLAB's best-performing model, suggesting that both platforms are equally capable of achieving high accuracy in this task. However, WEKA's faster computational efficiency, as evidenced by its consistent minimal training time, making it more suitable for real-time and large-scale applications. The confusion matrix solidifies Random Tree's role as a top contender for practical GMO rice detection. Additional metrics validated model robustness. Both the Trilayered Neural Network and Random Tree achieved Precision and Recall values of 1.0, confirming their ability to avoid false classifications. The F1-scores of all top models were similarly perfect, underscoring their suitability for sensitive applications like GMO detection. Furthermore, Random Tree, catered by WEKA and is under Abstract Classifiers, is a class that considers K randomly chosen attributes at each node [42]. It is characterized to have simpler interpretability. Random Tree can also effectively reduce risk of overfitting which can be beneficial in handling larger datasets especially that rice grains cannot just be counted in practical and on-site detection.

Table 2. Top performing MATLAB CLA models

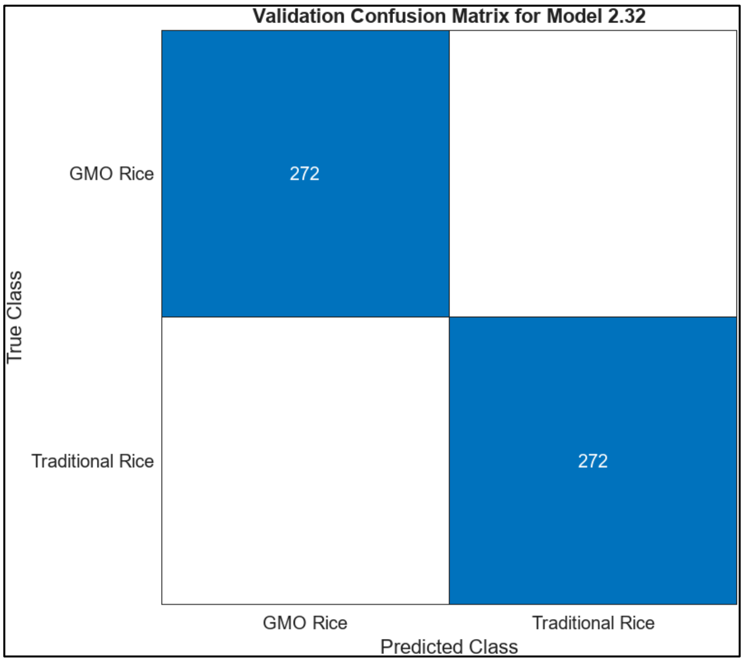
Top performing MATLAB CLA modelsMATLAB CLA Model/Classifier	Training Accuracy (%)	Testing Accuracy (%)	Training Time (sec)	Rank
Trilayered Neural Network	100.00	100.00	1.03	1
Cubic SVM	100.00	100.00	1.07	2
Coarse KNN	100.00	100.00	1.37	3
Linear Discriminant	100.00	100.00	1.42	4
Efficient Linear SVM	100.00	100.00	1.52	5
Linear SVM	100.00	100.00	1.54	6
Wide Neural Network	100.00	100.00	1.75	7
Efficient Logistic Regression	100.00	100.00	1.91	8
Bilayered Neural Network	100.00	100.00	2.16	9
Cosine KNN	100.00	100.00	2.32	10
Coarse Gaussian SVM	100.00	100.00	2.37	11
Quadratic Discriminant	100.00	100.00	3.35	12
Binary LGM Logistic Regression	100.00	100.00	4.66	13
Narrow Neural Network	100.00	100.00	11.24	14
Medium Neural Network	100.00	100.00	12.36	15
Subspace KNN Ensemble	100.00	100.00	14.04	16

Table 3. Top performing WEKA models

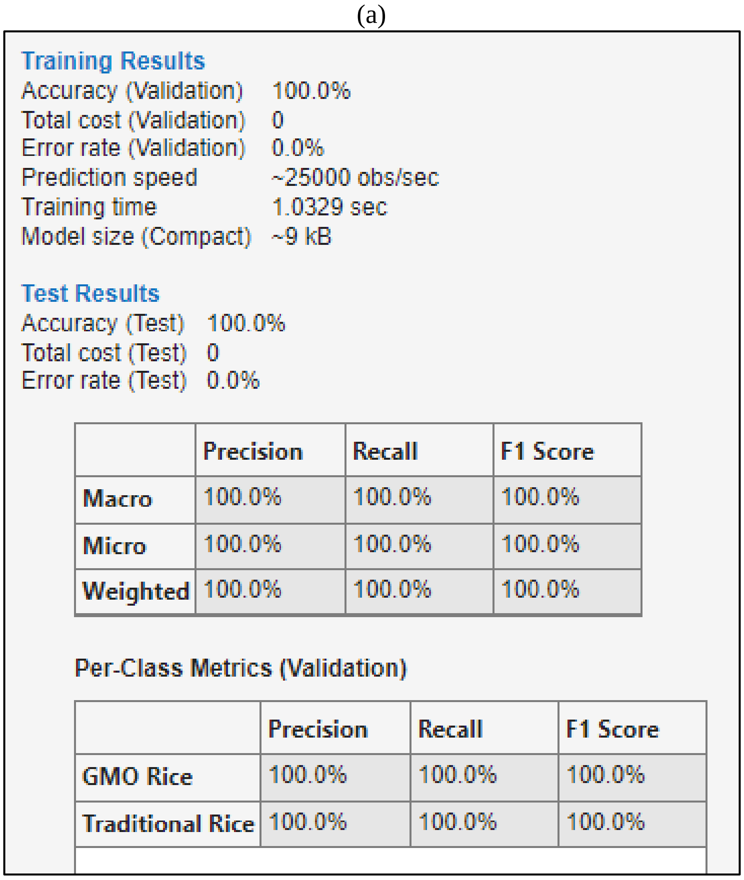
WEKA Model/Classifier	Training Accuracy (%)	Testing Accuracy (%)	Training Time (sec)	Rank
Random Tree	100.00	100.00	0.00	1
SMO Function	100.00	100.00	0.01	2
AdaBoost M1 Meta	100.00	100.00	0.01	2
Random Committee	100.00	100.00	0.01	2
Logistic Function	100.00	100.00	0.02	3
JRip Rule	100.00	100.00	0.02	3
Random Subspace	100.00	100.00	0.02	3
SGD Function	100.00	100.00	0.03	4
IBk Lazy	100.00	100.00	0.03	4
Randomizable Filtered Classifier	100.00	100.00	0.04	5
Random Forest Tree	100.00	100.00	0.05	6
Logit Boost	100.00	100.00	0.10	7
Multiclass Classifier	100.00	100.00	0.10	7
Multiclass Classifier Updateable	100.00	100.00	0.10	7
Simple Logic Function	100.00	100.00	0.13	8
Multilayer Perceptron Function	100.00	100.00	0.16	9
LMT Tree	100.00	100.00	0.16	9
Iterative Classifier Optimizer	100.00	100.00	0.20	10
LWL Lazy	100.00	100.00	0.34	11
Kstar Lazy	100.00	100.00	0.39	12

At the time of writing, a similar rice classification study focusing on RGB values reported 100% accuracy using MATLAB's Coarse Tree, effectively distinguishing between healthy and unhealthy rice seed plants [43]. However, the model took 0.32189 seconds to build, which is relatively faster than the performance result of the MATLAB models obtained in this study. In another local study, WEKA's Random Tree also achieved 100% accuracy when using the same dataset from the aforementioned study, but with a focus on WEKA's models [44]. In contrast to the previous study, it recorded a training time of 0.02 seconds, showing that Random Tree performs better in classifying GMO Golden Rice compared to traditional rice—rather than just distinguishing between healthy and unhealthy rice.

Existing studies have not yet explored the classification and detection of genetically modified rice and other biofortified crops using similar methodologies. Most studies have focused on comparing traditional white rice variants based on features such as morphology, texture, nitrogen content, internal broken, and other rice grader parameters [45]-[50]. This highlights a gap in the current body of knowledge, calling for further research and improvements in GMO rice classification. Future studies are expected to validate the results presented in this paper and explore the potential for improving the models by considering additional GMO rice variants, hybrid varieties, and incorporating other relevant features.



(a)



Training Results

Accuracy (Validation)	100.0%
Total cost (Validation)	0
Error rate (Validation)	0.0%
Prediction speed	~25000 obs/sec
Training time	1.0329 sec
Model size (Compact)	~9 kB

Test Results

Accuracy (Test)	100.0%
Total cost (Test)	0
Error rate (Test)	0.0%

	Precision	Recall	F1 Score
Macro	100.0%	100.0%	100.0%
Micro	100.0%	100.0%	100.0%
Weighted	100.0%	100.0%	100.0%

Per-Class Metrics (Validation)

	Precision	Recall	F1 Score
GMO Rice	100.0%	100.0%	100.0%
Traditional Rice	100.0%	100.0%	100.0%

(b)

Fig. 9. Confusion matrix (a) and summary (b) of MATLAB's most efficient model

Vol. 5, No. 2, 2025, pp. 661-677

```
=== Confusion Matrix ===

   a   b   <-- classified as
 272   0 |   a = gmo_golden_rice
   0 272 |   b = traditional_rice
```

(a)

```
Correctly Classified Instances         544              100      %
Incorrectly Classified Instances         0                0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances              544

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC
                 1.000    0.000    1.000      1.000    1.000      1.000
                 1.000    0.000    1.000      1.000    1.000      1.000
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000
```

(b)

Fig. 10. Confusion matrix (a) and summary (b) of WEKA's most efficient model

5. Conclusion

This study successfully conducted a comprehensive comparative analysis of the 80 ML classifiers in assessing their effectiveness in distinguishing GMO Golden Rice from traditional rice, using RGB histogram data. Most MATLAB and WEKA platforms showed perfect accuracy in both training and testing phases. The Trilayered Neural Network emerged as the most efficient model in MATLAB, with its ability to handle complexities in rice color feature in just 1.03 sec. In WEKA, the Random Tree stood out for its negligible training time 0.00 sec, this is crucial in examining large dataset as Random Tree can effectively reduce risk of overfitting. The integration of Precision, Recall, and F1-scores further confirms the robustness of the proposed models. Moreover, this study noted a significant gap in current research—existing studies primarily focus on classifying traditional rice based on features like morphology, texture, and color. The application of machine learning to genetically modified rice, as explored here, offers new possibilities and a much-needed expansion in the field of GMO detection and biofortified crop monitoring. Future research should seek to validate these findings and explore ways to improve the models further by considering additional GMO rice variants, such as the IRRI's newly released low-sugar GMO rice. Future work may also deal with hybrid variants, as well as adding extra features like Morphology and Texture, which can still be done within MATLAB. Ultimately, this approach has the potential to revolutionize GMO monitoring, providing a scalable, field-deployable solution that empowers stakeholders across the agricultural value chain, from farmers to policymakers, enabling more efficient GMO rice detection, enhancing food safety, and promoting precision agriculture.

Author Contributions: All authors contributed equally, in all forms, in the conduct of this study. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgment: The principal author expresses his utmost gratitude to Mrs. Nerie Rose M. Veniegas for generously providing the GMO Golden Rice all the way from Pinamalayan, Mindoro. Special thanks to his research advisor, Dr. Edwin R. Arboleda, for his invaluable advices and support throughout the study. The researcher also extends his sincerest appreciation to Engr. Nemilyn A. Fadchar for her expertise and exemplary contributions to this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] FAO, " Rice is Life: International Year of Rice 2004 and Its Implementation," *Food and Agriculture Organization*, 2004, https://books.google.co.id/books/about/Rice_is_Life.html?id=es0gYn9maugC&redir_esc=y.

[2] PSA, "Philippine Rice Industry Statistics," *Philippine Statistics Authority*, 2023, https://psa.gov.ph/statistics/crops/palay-rice-production-prices.

[3] WHO, "Micronutrient Deficiencies: Vitamin A Deficiency," *World Health Organization*, 2020, https://www.who.int/data/nutrition/nlis/info/vitamin-a-deficiency#:~:text=Vitamin%20A%20deficiency%20results%20from,rarely%20seen%20in%20developed%20countries.

[4] G. A. Stevens *et al*., "Trends and mortality effects of vitamin A deficiency in children in 138 low-income and middle-income countries between 1991 and 2013: a pooled analysis of population-based surveys," *The Lancet Global Health*, vol. 3, no. 9, pp. e528–e536, 2015, https://doi.org/10.1016/S2214-109X(15)00039-X.

[5] A. C. Dubock, J. Wesseler, R. M. Russell, C. Chen, and D. Zilberman, "Golden Rice, VAD, Covid and Public Health: Saving Lives and Money," *Integrative Advances in Rice* Research, 2022, https://doi.org/10.5772/intechopen.101535.

[6] DOST-FNRI, "Philippine Nutrition Facts and Figures: 2018-2019 Expanded National Nutrition Survey (ENNS)," *Department of Science and Technology - Food and Nutrition Research Institute*, 2022, https://enutrition.fnri.dost.gov.ph/uploads/2018-2019%20ENNS%20FACTS%20AND%20FIGURES_JULY182023.pdf.

[7] UNICEF, "Direct and indirect effects of the COVID-19 pandemic and response in South Asia," *United Nations Children's Fund*, 2021, https://www.unicef.org/rosa/reports/direct-and-indirect-effects-covid-19-pandemic-and-response-south-asia.

[8] B. P. M. Swamy *et al*., "Development and characterization of GR2E Golden rice introgression lines," *Scientific Reports*, vol. 11, no. 1, p. 2496, 2021, https://doi.org/10.1038/s41598-021-82001-0.

[9] IRRI, "Golden Rice Project," *International Rice Research Institute*, 2021, https://www.irri.org/golden-rice-project.

[10] IRRI, "Philippines Authorizes Commercial Golden Rice Cultivation," *International Rice Research Institute*, 2021, https://www.irri.org/golden-rice-cultivation.

[11] DA, "First Large-Scale Harvest of Golden Rice in the Philippines," *Department of Agriculture*, 2022, https://phys.org/news/2022-11-farmers-philippines-cultivated-golden-rice.html.

[12] GLP, "What is nutritionally enhanced Golden Rice and why is it controversial?," *Genetic Literacy Project*, 2024, https://geneticliteracyproject.org/gmo-faq/what-is-nutritionally-enhanced-golden-rice-and-why-is-it-controversial/.

[13] VIB, "Golden Rice," *VIB*, 2024, https://vib.be/sites/vib.sites.vib.be/files/2022-05/vib_fact_GoldenRice_EN.pdf.

[14] ISAAA, "GMO Approval Database," *International Service for the Acquisition of Agri-biotech Applications*, 2024, https://www.isaaa.org/gmapprovaldatabase/event/default.asp?EventID=528.

[15] A. Sarkar, B. Madhavidevi, S. Nandi, A. Mukherjee and M. Botlagunta, "Comparative Machine Learning Approaches to Identify the Rice Cultivars," *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pp. 945-950, 2024, https://doi.org/10.1109/ICSCSS60660.2024.10625290.

[16] M. Śliwińska-Bartel, D. T. Burns, and C. Elliott, “Rice fraud a global problem: A review of analytical tools to detect species, country of origin and adulterations,” *Trends in Food Science & Technology*, vol. 116, pp. 36-46, 2021, https://doi.org/10.1016/j.tifs.2021.06.042.

[17] S. J. Smyth, W. A. Kerr, and P. W. B. Phillips, "Biotechnology Regulation and Trade," *Natural Resource Management and Policy*, vol. 51, 2017, https://doi.org/10.1007/978-3-319-53295-0.

[18] P. Safaei, E. M. Aghaee, G. J. Khaniki, S. A. K. Afshari, and S. Rezaie, “A simple and accurate PCR method for detection of genetically modified rice,” *Journal of Environmental Health Science and Engineering*, vol. 17, no. 2, pp. 847-851, 2019, https://doi.org/10.1007/s40201-019-00401-x.

[19] H. Liu *et al.*, “Rapid detection of P–35S and T-nos in genetically modified organisms by recombinase polymerase amplification combined with a lateral flow strip,” *Food Control*, vol. 107, p. 106775, 2020, https://doi.org/10.1016/j.foodcont.2019.106775.

[20] H. Zhao *et al.*, "In Situ Collection and Rapid Detection of Pathogenic Bacteria Using a Flexible SERS Platform Combined with a Portable Raman Spectrometer," *International Journal of Molecular Sciences*, vol. 23, no. 13, p. 7340, 2022, https://doi.org/10.3390/ijms23137340.

[21] İ. Cinar and M. Koklu, "Identification of Rice Varieties Using Machine Learning Algorithms," *Journal of Agricultural Sciences*, vol. 28, no. 2, pp. 307-325, 2021, https://doi.org/10.15832/ankutbd.862482.

[22] Dubey and Petruzzello, “Golden Rice,” *Britannica*, 2024, https://www.britannica.com/technology/golden-rice.

[23] N. K. Naik, M. V. Subbarao, P. K. Sethy, S. K. Behera, and G. R. Panigrahi, "Machine Learning with ANOVA-Based Method for Identifying Rice Varieties," *Journal of Agriculture and Food Research*, vol. 18, p. 101397, 2024, https://doi.org/10.1016/j.jafr.2024.101397.

[24] P. Saxena, K. Priya, and S. Goel, "Rice Varieties Classification Using Machine Learning Algorithms," *Journal of Pharmaceutical Negative Results*, vol. 13, no. 7, pp. 3762-3772, 2022, https://doi.org/10.47750/pnr.2022.13.S07.479.

[25] S. J. M. Rad, F. A. Tab and K. Mollazade, "Classification of Rice Varieties Using Optimal Color and Texture Features and BP Neural Networks," *2011 7th Iranian Conference on Machine Vision and Image Processing*, pp. 1-5, 2011, https://doi.org/10.1109/IranianMVIP.2011.6121583.

[26] S. Qadri *et al.*, "Machine Vision Approach for Classification of Rice Varieties Using Texture Features," *International Journal of Food Properties*, vol. 24, no. 1, pp. 1615-1630, 2021, https://doi.org/10.1080/10942912.2021.1986523.

[27] L. Kamelia, E. A. Z. Hamidi, and R. M. Fadilla, “Rice quality classification system using convolutional neural network and an adaptive neuro-fuzzy inference system,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 4, pp. 4113-4120, 2024, http://doi.org/10.11591/ijai.v13.i4.pp4113-4120.

[28] S. B. Ahmed, S. F. Ali and A. Z. Khan, "On the Frontiers of Rice Grain Analysis, Classification and Quality Grading: A Review," *IEEE Access*, vol. 9, pp. 160779-160796, 2021, https://doi.org/10.1109/ACCESS.2021.3130472.

[29] B. Arora, N. Bhagat, L. R. Saritha and S. Arcot, "Rice Grain Classification using Image Processing & Machine Learning Techniques," *2020 International Conference on Inventive Computation Technologies (ICICT)*, pp. 205-208, 2020, https://doi.org/10.1109/ICICT48043.2020.9112418.

[30] A. Sharma, A. Jain, P. Gupta and V. Chowdary, "Machine Learning Applications for Precision Agriculture: A Comprehensive Review," *IEEE Access*, vol. 9, pp. 4843-4873, 2021, https://doi.org/10.1109/ACCESS.2020.3048415.

[31] N. M. U. Din *et al.*, "RiceNet: A Deep Convolutional Neural Network Approach for Classification of Rice Varieties," *Expert Systems with Applications*, vol. 235, pp. 121214, 2024, https://doi.org/10.1016/j.eswa.2023.121214.

[32] V. T. Hoang, D. P. van Hoai, T. Surinwarangkoon, H. Duong, and K. Meethongjan, "A Comparative Study of Rice Variety Classification Based on Deep Learning and Hand-Crafted Features," *ECTI Transactions on Computer and Information Technology*, vol. 14, no. 1, pp. 1–10, 2020, https://doi.org/10.37936/ecti-cit.2020141.204170.

[33] Y. Qian, Q. Xu, Y. Yang, H. Lu, H. Li, X. Feng, and W. Yin, "Classification of Rice Seed Variety Using Point Cloud Data Combined with Deep Learning," *International Journal of Agricultural and Biological Engineering*, vol. 14, no. 5, pp. 206-212, 2021, https://doi.org/10.25165/j.ijabe.20211405.5902.

[34] Y. Meng, Z. Ma, Z. Ji, R. Gao, and Z. Su, "Fine Hyperspectral Classification of Rice Varieties Based on Attention Module 3D-2DCNN," *Computers and Electronics in Agriculture*, vol. 203, pp. 107474, 2022, https://doi.org/10.1016/j.compag.2022.107474.

[35] G. Shi, X. Zhang, G. Qu, and Z. Chen, "Classification of Rice Varieties Using SIMCA Applied to NIR Spectroscopic Data," *ACS Omega*, vol. 7, no. 50, pp. 46623–46628, 2022, https://doi.org/10.1021/acsomega.2c05561.

[36] J. Onmankhong, T. Ma, T. Inagaki, P. Sirisomboon, and S. Tsuchikawa, "Cognitive Spectroscopy for the Classification of Rice Varieties: A Comparison of Machine Learning and Deep Learning Approaches in Analyzing Long-Wave Near-Infrared," *Infrared Physics & Technology*, vol. 123, pp. 104100, 2022, https://doi.org/10.1016/j.infrared.2022.104100.

[37] N. Rahmani and A. Mani-Varnosfaderani, "Quality Control, Classification, and Authentication of Iranian Rice Varieties Using FT-IR Spectroscopy and Sparse Chemometric Methods," *Journal of Food Composition and Analysis*, vol. 112, p. 104650, 2022, https://doi.org/10.1016/j.jfca.2022.104650.

[38] Y. Meng *et al.*, "Fine Hyperspectral Classification of Rice Varieties Based on Self-Attention Mechanism," *Ecological Informatics*, vol. 75, p. 102035, 2023, https://doi.org/10.1016/j.ecoinf.2023.102035.

[39] U. Rauf et al., "A New Method for Pixel Classification for Rice Variety Identification Using Spectral and Time Series Data from Sentinel-2 Satellite Imagery," *Computers and Electronics in Agriculture*, vol. 193, p. 106731, 2022, https://doi.org/10.1016/j.compag.2022.106731.

[40] L. Zhu *et al.*, "Identification of Rice Varieties and Determination of Their Geographical Origin in China Using Raman Spectroscopy," *Journal of Cereal Science*, vol. 82, p. 175-182, 2018, https://doi.org/10.1016/j.jcs.2018.06.010.

[41] A. R. Jalalvand, M. Roushani, H. C. Goicoechea, D. N. Rutledge, H. W. Gu, "MATLAB in electrochemistry: A review," *Talanta*, vol. 194, pp. 205-225, 2019, https://doi.org/10.1016/j.talanta.2018.10.041.

[42] WEKA, "RandomTree Classifier," *WEKA*, 2024, https://weka.sourceforge.io/doc.dev/weka/classifiers/trees/RandomTree.html.

[43] K. W. V. Geollegue, E. Romeroso Arboleda, and A. Agustin Dizon, "Seed of rice plant classification using coarse tree classifier," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 2, pp. 727-735, 2022, http://doi.org/10.11591/ijai.v11.i2.pp727-735.

[44] E. R. Arboleda, A. R. Bautista, D. G. M. Arboleda, "Classification of Rice Seeds Using Random Tree Algorithm," *Gongcheng Kexue Yu Jishu/Advanced Engineering Science*, 2024, https://www.gkyj-aes-20963246.com/article/classification-of-rice-seeds-using-random-tree-algorithm.

[45] V. Sonawane, N. Gaikwad, H. Mandekar, K. Baradkar, and C. Gunjal, "Rice Quality Analysis and Classification Using Image Processing Techniques," *International Journal of Computer Science and Mobile Computing*, vol. 10, no. 6, pp. 79–82, 2021, https://doi.org/10.47760/ijcsmc.2021.v10i06.008.

[46] M. K. Khan, M. Mohan, P. Jaiswal, P. Kumar, and P. Srivastava, "A Review Paper on Rice Quality Analysis Using Image Processing Technique," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 5, pp. 292–294, 2022, https://doi.org/10.22214/ijraset.2022.42118.

[47] M. Muliady, L. Tien Sze, K. Voon Chet, and S. Patra, "Classification of rice plant nitrogen nutrient status using k-nearest neighbors (k-NN) with light intensity data," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 1, pp. 179-186, 2021, http://doi.org/10.11591/ijeecs.v22.i1.pp179-186.

[48] Bhupendra, K. Moses, A. Miglani, and P. K. Kankar, "Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset," *Computers and Electronics in Agriculture*, vol. 195, p. 106811, 2022, https://doi.org/10.1016/j.compag.2022.106811.

[49] H. Zia, H. S. Fatima, M. Khurram, I. U. Hassan, and M. Ghazal, "Rapid Testing System for Rice Quality Control through Comprehensive Feature and Kernel-Type Detection," *Foods*, vol. 11, no. 18, p. 2723, 2022, https://doi.org/10.3390/foods11182723.

[50] J. Ye, Z. Hu, Y. Chen, D. Fu, and J. Zhang, "Identification of Broken Rice Rate Based on Grading and Morphological Classification," *LWT*, p. 117175, 2024, https://doi.org/10.2139/ssrn.4979045.