

Deep learning–based multispectral seed classification for detecting adulteration in KDML-105 rice

Khunnithi Doungpueng¹⁾, Jirasin Prueksawan²⁾, Lalita Panduangnat²⁾, Anuradee Kirichee²⁾ and Jetsada Posom^{*2)}

¹⁾Department of Mechanical Technical Education, Faculty of Technical Education, Rajamangala University of Technology Isan Khon Kaen Campus, Khon Kaen 40000, Thailand

²⁾Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40000, Thailand

Received 30 December 2025

Revised 25 June 2026

Accepted 17 July 2026

Abstract

This study aimed to develop and evaluate Deep Learning models for detecting varietal contamination in KDML-105 jasmine rice seeds using multispectral imaging and seed morphological features. Four rice varieties—KDML-105 and three Non-KDML-105 varieties (Chainat-1, RD-15, and RD-6)—were used, with 1,000 seeds collected per variety (4,000 seeds in total). All samples were subjected to imaging, object segmentation, and statistical feature extraction across multiple spectral bands to create inputs for three Deep Learning models: a Bayesian Neural Network (BNN), a Convolutional Neural Network (CNN), and a Recurrent Neural Network (RNN). Experimental results showed that the BNN achieved the highest accuracy on multispectral datasets, reaching 0.99 for both the Physical Reflectance and Multi-Reflectance datasets. The CNN produced slightly lower accuracy, while the RNN consistently performed the poorest across all datasets, indicating that feed-forward architectures are better suited for non-sequential reflectance data. SHAP-based feature analysis identified *Min_nir* as the most influential variable for classifying both KDML-105 and Non-KDML-105 seeds, followed by *Mean_blue* and several statistical descriptors from the green and blue spectral bands. These findings highlight the strong discriminative capability of NIR reflectance for separating rice varieties with subtle morphological differences, supported by additional surface-related cues captured in the blue band. Overall, the results demonstrate that Deep Learning combined with multispectral imaging is a viable approach for detecting seed contamination. The identification of key spectral bands suggests the potential for designing a lower-cost inspection system using only NIR and blue channels, enabling the development of automated or real-time sorting solutions suitable for practical implementation in Thailand's rice industry.

Keywords: Spectral feature analysis, Bayesian neural networks, Agricultural quality inspection, Explainable AI (SHAP), Seed purity verification

1. Introduction

Khao Dawk Mali 105 (KDML-105) has long been recognized as one of Thailand's most distinguished rice varieties, valued for its distinctive aroma and flavor [1, 2]. In recent years, however, the intentional mixing of KDML-105 seeds with cheaper varieties to reduce production costs has become increasingly common. This practice lowers the purity of KDML-105, undermines international confidence in "Thai rice," and indirectly affects farmers' income and the competitiveness of Thailand's rice exports [3-6].

Contamination of rice varieties is typically detected only after the crop has grown in the field, when the mixed varieties become visually distinguishable from KDML-105—for example, through taller plant height or differing stem coloration [7-9]. In such cases, local agricultural officers must intervene in the field, providing a solution that addresses the problem only after it has occurred [10, 11]. A more effective and sustainable approach is to detect varietal contamination at the seed stage [12-16]. Currently, seed purity inspection is carried out at rice research centers by trained experts who rely on visual assessment of seed morphological characteristics [13, 17]. This process is time-consuming and susceptible to human error, which may arise from factors such as fatigue, emotional state, or environmental conditions [13, 18-21].

Advances in artificial intelligence—particularly deep learning—have created new opportunities for improving agricultural engineering applications, such as plant disease classification and seed quality assessment [22-27]. Deep learning offers strong capability in discerning subtle morphological features of rice seeds, including shape, texture, and spectral signatures [12, 27, 28]. These fine-grained distinctions can be challenging for human inspectors, especially when varieties share highly similar physical characteristics, as in the case of Kor Khor 15 (RD-15) mixed with KDML-105 [29, 30]. Moreover, integrating hyperspectral imaging (HSI) with deep learning has proven effective for discriminating among rice varieties [27, 31]. However, most existing studies have been conducted under laboratory conditions or require high-cost instruments, limiting their practical adoption [32]. While hyperspectral imaging offers continuous spectral coverage across hundreds of narrow bands, the associated acquisition hardware, data volumes, and preprocessing requirements impose practical constraints that limit its deployment in field conditions or community-level seed centers [32]. Fixed-band multispectral imaging, exemplified by the five-band MicaSense RedEdge-MX platform, represents an intermediate position

*Corresponding author.

Email address: jetspo@kku.ac.th

<https://doi.org/10.64960/easr.2026.265491>

between conventional RGB-based machine vision and full hyperspectral systems, offering spectral discriminative capacity at substantially lower hardware cost and data complexity, and is therefore adopted in the present study as a practically deployable alternative.

Despite this growing body of work, two gaps remain in the existing literature on multispectral seed classification. First, although prior studies have demonstrated that Deep Learning can extract subtle spectral and morphological cues for rice variety discrimination [12, 27, 28], few have systematically compared multiple Deep Learning architectures—feed-forward, convolutional, and recurrent—side-by-side on a common, non-sequential multispectral feature representation derived from a low-cost, fixed-band imaging platform, leaving open the question of which architectural family is best suited to this data structure. Second, although feature-importance analysis (e.g., SHAP) has been applied to hyperspectral and machine-learning classifiers in related work [33-41], its use for guiding the design of reduced-band, lower-cost multispectral inspection systems for varietal purity assessment remains comparatively unexplored. The present study addresses both gaps by developing and comparing three Deep Learning architectures (BNN, CNN, RNN) on a common multispectral feature set for distinguishing KDML-105 from morphologically similar contaminant varieties, and by using SHAP-based interpretability to identify the spectral bands most relevant to a cost-effective inspection system.

Given these challenges, early detection of varietal contamination at the seed stage is essential for preserving the purity of KDML-105. Deep learning provides a promising solution due to its ability to extract and classify subtle morphological differences in seed characteristics. Therefore, the objective of this study is to develop a deep learning model capable of detecting contamination of KDML-105 seeds with other varieties using multispectral imaging. The proposed approach aims to enhance inspection accuracy, minimize human error, and reduce the time required for seed purity assessment.

2. Materials and methods

2.1 Image and reference data collection

The experiment used seed samples from four rice varieties: Khao Dawk Mali 105 (KDML-105) and three non-KDML-105 varieties—Chainat-1, RD-15 (Kor Khor 15), and RD-6 (Kor Khor 6). For each variety, 1,000 seeds were randomly sampled, giving a balanced dataset of 4,000 seeds across the four varieties. Within each variety, the 1,000 seeds were partitioned into five equal subsets of 200 seeds each; these subsets were used for cross-validation during model development and evaluation. For image acquisition, each subset of 200 seeds was arranged on two transparent acrylic seed-alignment plates (100 holes per plate) to ensure consistent spacing and orientation during imaging.

Imaging was performed using a purpose-built prototype imaging system comprising three subunits: a structural frame, a lighting unit, and an imaging unit. The enclosure was a cubic frame ($60 \times 60 \times 60$ cm) constructed from 30×30 mm aluminum profiles and secured with 20 mm rigid brackets. Matte black acrylic panels lined the interior to minimize stray reflections and ambient light interference. Illumination was provided by four 120 W GataPAR-38 halogen lamps mounted at the frame corners and positioned approximately 40 cm above the base. Image capture was conducted with a MicaSense RedEdge-MX multispectral camera capable of recording five discrete spectral bands: Red (668 nm), Green (560 nm), Blue (475 nm), Red-edge (717 nm), and a single Near-Infrared band (NIR, 840 nm); the camera was mounted 40 cm above the base. A transparent acrylic seed alignment plate (100 holes) was used to hold seeds in an ordered grid for each capture session. Camera exposure parameters (shutter speed, lens aperture, and white balance) and the camera-to-sample distance (40 cm) were held constant across all acquisition sessions to ensure cross-session comparability of the extracted reflectance values. Prior to each imaging session, the camera was calibrated using a certified reflectance reference panel under the prevailing illumination conditions, and raw 16-bit imagery was subsequently converted to calibrated reflectance values. The overall setup is illustrated in Figure 1.

2.2 Image enhancement and segmentation

Raw multispectral images were preprocessed to isolate individual seeds and remove background artifacts. The preprocessing pipeline included region-of-interest (ROI) extraction and object segmentation to separate seed pixels from the background. Segmentation was performed using a two-stage procedure: first, Otsu's adaptive thresholding was applied to the NIR band—selected for its high contrast between seed tissue and background—to generate an initial binary mask; second, where spatial illumination gradients compromised threshold-based separation, a supplementary HSV colour-space segmentation was applied to the RGB composite image to refine the mask boundaries. The two outputs were combined via a logical OR operation, followed by morphological closing to fill internal voids and consolidate seed contours. Background removal produced binary mask layers for each seed, which were then applied to the multispectral images to extract seed-specific data. From these masked regions, per-seed spectral and spatial statistics were computed for each band, including mean, maximum, minimum, standard deviation, and variance. The resulting mask layers were also used to compute morphological measurements (see Section 2.3).

2.3 Data preparation and feature extraction

From the segmented and masked seed images, geometric features—seed length and seed width—were measured after calibration from pixels to physical units. In addition to the geometric measurements, the per-band statistical descriptors computed in Section 2.2 (mean, max, min, standard deviation, variance) were used as spectral features. The feature set thus combined spectral descriptors and simple morphological features for each seed. For classification, samples were labeled into two classes: KDML-105 and Non-KDML-105. The resulting balanced dataset of 4,000 seeds (1,000 seeds per variety), each organized into five subsets of 200 seeds, enabled five-fold cross-validation to evaluate model generalizability. Within each fold, seeds from all four varieties were proportionally represented, ensuring that both classes (KDML-105 and Non-KDML-105) were balanced across training and validation partitions.

2.4 Model development and evaluation

Three model families were implemented and compared: Bayesian Neural Networks (BNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). The selection of these three families was guided by the structure of the extracted feature set, which consists of per-band statistical descriptors (mean, maximum, minimum, standard deviation, and variance) together

with simple morphological measurements, rather than raw image arrays. These architectures were selected to capture different aspects of the data: CNNs for spatial/pattern recognition in the multispectral images, RNNs for sequential/spectral relationships across bands, and BNNs to provide probabilistic prediction and uncertainty estimates. The BNN was implemented as a multilayer perceptron with one or two fully connected hidden layers using ReLU activation, in which dropout was applied to approximate Bayesian inference over the network weights, yielding both class predictions and associated uncertainty estimates. The CNN consisted of one-dimensional convolutional layers applied to the ordered feature vector, followed by pooling and fully connected layers, allowing the model to learn local patterns across adjacent spectral and morphological descriptors. The RNN was implemented using recurrent (e.g., LSTM/GRU-type) layers that processed the per-band statistical descriptors as an ordered sequence across spectral bands, followed by a fully connected output layer. Although multispectral reflectance features are not inherently sequential, the RNN was included as a contrasting architecture to empirically evaluate whether sequence-modelling mechanisms confer any advantage for this feature representation; as detailed in Section 3, its consistently lower accuracy across all datasets confirms that feed-forward architectures (BNN, CNN) are better suited to this non-sequential reflectance-based feature space. Classical machine learning classifiers (e.g., Support Vector Machine, Random Forest) have been widely applied to multispectral and hyperspectral feature sets for rice variety classification [12, 27, 31]; the present study focuses on a systematic comparison among Deep Learning architectures on the same feature representation, with comparison against classical machine learning baselines identified as a direction for future work (see Section 4). Model hyperparameters—such as the number of hidden layers/filters, learning rate, batch size, and number of epochs—were tuned empirically by repeated experiments to mitigate overfitting. Standard regularization strategies (for example, early stopping and dropout) and data-level augmentation were employed where appropriate to improve generalization. Each model type was trained on the same training folds and evaluated under the same validation/testing protocol to ensure fair comparison. Performance was quantified using accuracy, precision, recall, and F1-score. Accuracy represents the overall proportion of correctly classified seeds; precision represents the proportion of seeds predicted as a given class that truly belong to that class; recall (sensitivity) represents the proportion of seeds of a given class that are correctly identified; and the F1-score is the harmonic mean of precision and recall, providing a balanced measure when class distributions are uneven. In addition, the area under the receiver operating characteristic curve (AUC), Cohen's Kappa, and the Matthews correlation coefficient (MCC) were computed as supplementary measures of overall classification agreement and robustness to chance agreement. To assess model stability, each training procedure was repeated across multiple runs and folds, and the metric distributions were recorded. Statistical analysis of the algorithms' performance was conducted using analysis of variance (ANOVA) on accuracy to test for significant differences among the three model families; where appropriate, post-hoc pairwise comparisons were performed to identify which methods differed significantly. Accuracy was selected as the primary variable for formal statistical testing because it provides a single, readily interpretable summary of overall classification performance and was consistently the metric exhibiting the clearest separation among models across all datasets; precision, recall, and F1-score were reported descriptively across the same model groupings to corroborate the accuracy-based ranking and were found to follow consistent trends. The model exhibiting the highest, statistically supported accuracy (i.e., also ranking consistently highest on precision, recall, and F1-score) was selected as the optimal model. The experimental workflow and evaluation procedure are summarized in Figure 1.

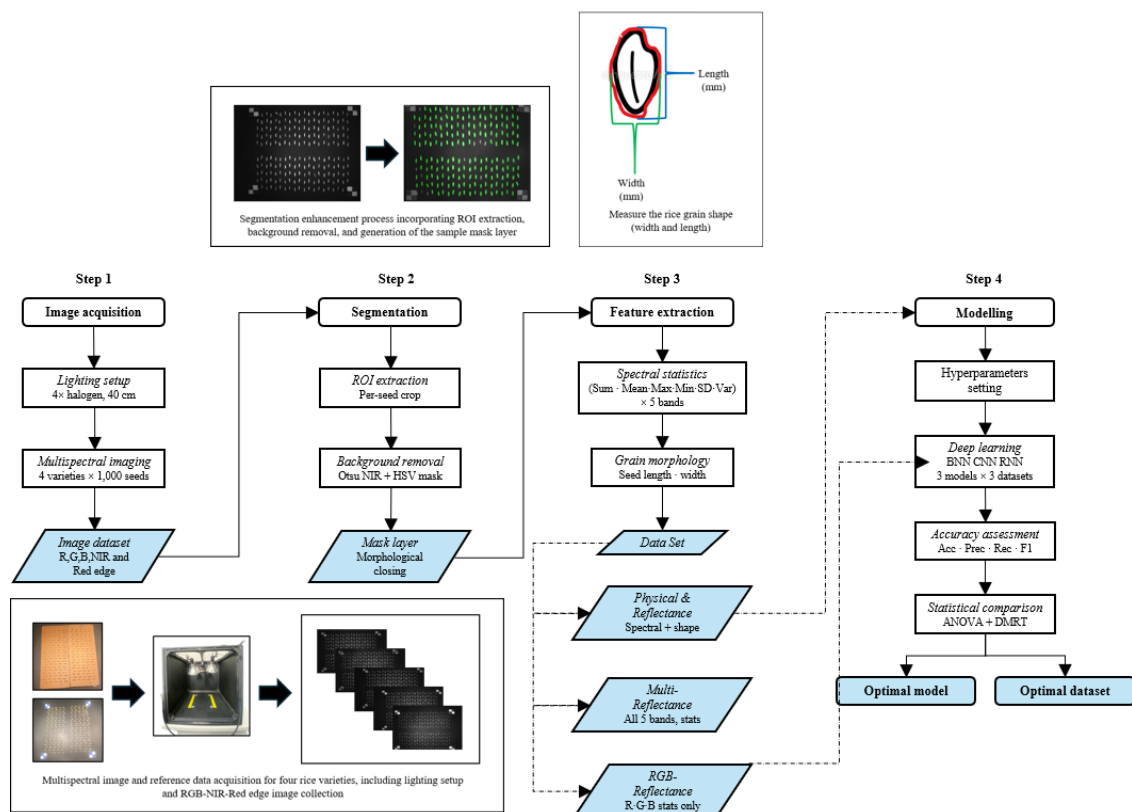


Figure 1 Experimental workflow for multispectral-based rice seed authentication across four steps: (1) multispectral image acquisition (Red, Green, Blue, Red-edge, and NIR bands) of four rice varieties (1,000 seeds per variety), yielding the raw Image dataset; (2) segmentation enhancement via ROI extraction, background removal, and mask layer generation; (3) feature extraction producing three datasets — the Physical and Reflectance dataset (per-band spectral statistics with grain length and width), the Multi-Reflectance dataset (spectral statistics from all five bands), and the RGB-Reflectance dataset (spectral statistics from Red, Green, and Blue bands only); and (4) Deep Learning model training (BNN, CNN, RNN) on each dataset, with accuracy assessment and statistical comparison (ANOVA + DMRT) to determine the optimal model and optimal dataset.

3. Results and discussion

3.1 Performance evaluation of Deep Learning models across different datasets

This study evaluated the performance of three Deep Learning models—Bayesian Neural Network (BNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN)—using three distinct datasets: (1) the Physical and Reflectance dataset, (2) the Multi-Reflectance dataset, and (3) the RGB-Reflectance dataset. Model performance was assessed using four standard evaluation metrics: accuracy, precision, recall, and F1-score (described in detail in Section 3.4). The results obtained from these experiments are summarized as follows.

3.1.1 Performance of Deep Learning models on the Physical–Reflectance dataset

Evaluation on the Physical–Reflectance dataset showed that the BNN model achieved the highest performance, with an accuracy of 0.99 and AUC, precision, recall, and F1-score values ranging from 0.99 to 1.00. These results indicate that the BNN model was highly effective in distinguishing between KDML-105 and Non-KDML-105 seeds. The Kappa and MCC values were both 0.98, further demonstrating strong agreement between the model predictions and the actual class labels.

The CNN model produced the second-highest accuracy at 0.98. In contrast, the RNN model achieved a noticeably lower accuracy of 0.87 and required the longest processing time at 20.75 seconds. Given that seed data are non-sequential, the RNN architecture is less suitable for this application, making it a less practical choice for seed purity assessment. The complete performance comparison is presented in Table 1.

3.1.2 Performance of Deep Learning models on the Multi-Reflectance dataset

On the Multi-Reflectance dataset, the BNN model achieved the highest performance, with an accuracy of 0.99 and an F1-score of 0.98. The CNN model followed closely, attaining an accuracy of 0.98. Examination of the Kappa and MCC values showed that BNN and CNN performed at comparable levels. In contrast, the RNN model yielded the lowest accuracy at 0.86, consistent with the expectation that recurrent architectures are not well suited for image-based data lacking temporal dependencies.

Processing time (TT) for the BNN and CNN models was 2.75 and 4.30 seconds, respectively—both acceptable for applications requiring near–real-time response. A summary of the performance metrics is presented in Table 1.

3.1.3 Performance of Deep Learning models on the RGB–Reflectance dataset

Evaluation using the RGB–Reflectance dataset showed a noticeable performance decline across all models due to the reduced number of spectral channels. The BNN and CNN models achieved accuracies of 0.89 and 0.91, respectively, indicating that RGB information alone may be insufficient for accurately discriminating rice seeds with subtle morphological similarities. Nonetheless, the CNN model outperformed the BNN slightly in terms of precision and F1-score, likely due to its inherent ability to capture spatial features more effectively from RGB images.

The RNN model once again yielded the lowest accuracy at 0.85 and exhibited a relatively high processing time of 15.55 seconds, making it unsuitable for practical deployment. Full performance details are summarized in Table 1.

3.1.4 Cross-dataset comparison

A cross-dataset comparison revealed that datasets containing multiple spectral bands—namely the Physical Reflectance and Multi-Reflectance datasets—consistently produced higher overall accuracy than the RGB–Reflectance dataset. This finding underscores the importance of spectral information, suggesting that spectral reflectance features provide more discriminative power for seed variety classification than color features alone.

In terms of processing time, the Physical Reflectance and RGB–Reflectance datasets yielded similar average computation times at 8.98 and 7.28 seconds, respectively. By contrast, the Multi-Reflectance dataset required the longest processing time, with an average of 10.00 seconds (Table 1). This increase may be attributed to the higher dimensionality of the multispectral inputs, which demand more computational resources during both preprocessing and model evaluation.

Considering the results presented above, an overall examination of model accuracy and computational efficiency indicates that the BNN consistently achieved the highest accuracy when multispectral information was included. This suggests a strong compatibility between spectral-based features and the feed-forward architecture, which can directly learn variations in reflectance intensity without relying on convolutional or recurrent operations. Consequently, the BNN is particularly suitable for datasets in which discriminative cues arise primarily from spectral rather than spatial characteristics. This observation aligns with the findings of Weng et al. [23], who reported that multispectral imaging effectively distinguishes rice varieties even when morphological differences are minimal, and that BNN models can leverage individual spectral bands with high efficiency.

In contrast, the CNN outperformed the BNN when evaluated on the RGB–Reflectance dataset. This is expected, as convolutional layers excel at extracting spatial features such as edges, surface textures, and color gradients—attributes that become more critical when spectral information is limited to only three channels. These results are consistent with studies by Feng et al. [22] and Izquierdo et al. [33], which demonstrated that CNNs provide substantial advantages in seed and plant disease classification from RGB imagery due to their ability to capture subtle visual patterns that may be difficult to discern through manual inspection.

The RNN, however, exhibited the lowest accuracy and the longest processing time across all datasets. This finding is consistent with the well-established understanding that recurrent architectures are optimized for sequential data, such as sensor signals or time-series measurements, rather than static images. Similar observations have been reported by Zhao and Xie [32] and Śliwińska-Bartel et al. [34], who noted that RNNs typically perform poorly on image-based tasks lacking temporal dependencies, resulting in reduced classification accuracy and limited practical applicability.

In terms of computational efficiency, the BNN model required the shortest processing time across all datasets, ranging from 2.55 to 2.87 seconds. This efficiency further supports its suitability for real-world deployment in seed contamination detection systems. Although the CNN required slightly more processing time than the BNN, its computational cost remained within an acceptable range

for practical implementation. By contrast, the RNN's substantially higher processing time renders it unsuitable for applications requiring timely responses or near-real-time operation.

3.2 Comparison of model accuracy using Duncan's Multiple Range Test (DMRT)

The comparison of model accuracy using Duncan's Multiple Range Test (DMRT) revealed statistically significant differences in performance among the models at the 95% confidence level. The letter groupings (a, b, c) presented in the results table indicate the DMRT-based classification of model performance, where models assigned different letters differ significantly in mean accuracy. These groupings highlight the relative ranking of the models and provide a clear statistical basis for distinguishing their effectiveness.

3.2.1 DMRT comparison of model accuracy for the Physical and Reflectance dataset

The Duncan's Multiple Range Test (DMRT) results for the Physical and Reflectance dataset showed that the BNN and CNN models did not differ significantly in terms of accuracy at the 95% confidence level. This indicates that, when both spectral and physical features are incorporated, feed-forward neural networks and convolutional architectures achieve similarly high levels of accuracy. In contrast, the RNN model performed significantly worse, reinforcing that recurrent structures are not well suited for data lacking temporal dependencies (Table 1).

3.2.2 DMRT comparison of model accuracy for the Multi-Reflectance dataset

The Duncan's Multiple Range Test (DMRT) results for the Multi-Reflectance dataset indicated no statistically significant difference in accuracy between the BNN and CNN models. This finding suggests that multispectral information, when paired with an appropriate network architecture, provides strong discriminative capability for seed classification. In contrast, the recurrent architecture did not offer additional benefits in this context, further confirming its limited suitability for non-sequential imaging data (Table 1).

3.2.3 DMRT comparison of model accuracy for the RGB-Reflectance dataset

The Duncan's Multiple Range Test (DMRT) results for the RGB-Reflectance dataset showed that the accuracy of the CNN model differed significantly from that of the BNN and RNN models. This significant difference highlights the advantage of CNNs in extracting spatial features—such as edges, textures, and color gradients—from RGB images. These spatial representations enable CNNs to perform more effectively than feed-forward and recurrent architectures when only limited spectral information is available (Table 1).

A cross-dataset comparison of accuracy using Duncan's Multiple Range Test (DMRT) showed that the BNN and CNN models did not differ significantly when multispectral data were used. This outcome is consistent with the findings of Weng et al. [23], who demonstrated that multispectral information effectively discriminates rice varieties even when morphological differences are subtle. The results therefore indicate that both BNN and CNN architectures can learn spectral variability and are suitable for seed contamination detection. In support of this interpretation, Śliwińska-Bartel et al. [34] also reported that seed-purity classification accuracy improves markedly when multiple wavelength bands are incorporated.

Across all datasets, however, the RNN model consistently produced the lowest accuracy, with statistically significant differences compared with both BNN and CNN. This aligns with the observations of Zhao and Xie [32], who emphasized that recurrent architectures are inherently optimized for sequential or time-dependent data—such as sensor streams—rather than static imagery. Because rice seed images do not contain temporal structure, the RNN is unable to extract discriminative features effectively, resulting in substantially poorer performance relative to feed-forward and convolutional models.

When considering the RGB-Reflectance dataset specifically, the CNN achieved significantly higher accuracy than the BNN. This advantage stems from the convolutional layers' ability to extract spatial cues—including texture patterns, surface characteristics, and edge information—more effectively than fully connected layers. Similar trends were reported by Izquierdo et al. [33], who used a CNN to classify European rice varieties from visible-spectrum images and found that even minor texture differences could be leveraged to improve classification accuracy.

From an application standpoint, the BNN model demonstrated the best performance when multispectral data were available, highlighting its practical relevance for seed-purity inspection systems. As low-cost multispectral imaging solutions are increasingly adopted by rice mills and seed-processing centers [6, 36], the BNN offers a compelling advantage due to its faster processing speed and stable performance. These characteristics suggest that, in operational environments, a BNN-based system may be more feasible and efficient than a CNN-based alternative.

Table 1 Performance comparison of BNN, CNN, and RNN models across three image datasets using multiple evaluation metrics, highlighting the superior accuracy and consistency of the BNN model

Data set	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
Physical and Reflectance data set	BNN	0.99 ^a	1.00	0.99	0.99	0.99	0.98	0.98	2.55
	CNN	0.98 ^a	1.00	0.98	0.97	0.97	0.95	0.95	3.65
	RNN	0.87 ^c	0.94	0.85	0.81	0.83	0.66	0.66	20.75
Muti-Reflectance data set	BNN	0.99 ^a	1.00	0.99	0.98	0.98	0.97	0.97	2.75
	CNN	0.98 ^a	1.00	0.98	0.97	0.98	0.95	0.95	4.3
	RNN	0.86 ^c	0.92	0.81	0.80	0.81	0.61	0.61	22.95
RGB-Reflectance data set	BNN	0.89 ^c	0.94	0.84	0.86	0.85	0.70	0.70	2.87
	CNN	0.91 ^b	0.95	0.85	0.89	0.87	0.73	0.74	3.41
	RNN	0.85 ^c	0.90	0.77	0.80	0.78	0.57	0.57	15.55

Remark: Different superscript letters (a–c) within each dataset indicate statistically significant differences among models according to ANOVA followed by Duncan's Multiple Range Test (DMRT) at $p < 0.05$.

3.3 Analysis of feature importance (Top 10 SHAP Features) for KDML-105 and Non-KDML-105 classification using the Bayesian Neural Network (BNN)

3.3.1 Feature importance analysis for KDML-105 using the BNN model

Figure 2 presents the SHAP-based feature importance for classifying KDML-105 using the BNN model. The results show that Min_nir was the most influential variable, with a SHAP value of approximately 0.27—substantially higher than all other features. This indicates that the model relies heavily on the near-infrared (NIR) reflectance minimum to differentiate KDML-105 seeds, highlighting the critical role of NIR spectral information.

The second most important feature was Mean_blue, with SHAP values ranging from 0.06 to 0.07. Following this were Min_green, Variance_green, and Mean_green, each showing SHAP values of roughly 0.03–0.04. Together, these features reflect the contribution of blue–green spectral characteristics in supporting discrimination within this class, suggesting that subtle variations in reflectance intensity across these wavelengths are informative for KDML-105 identification.

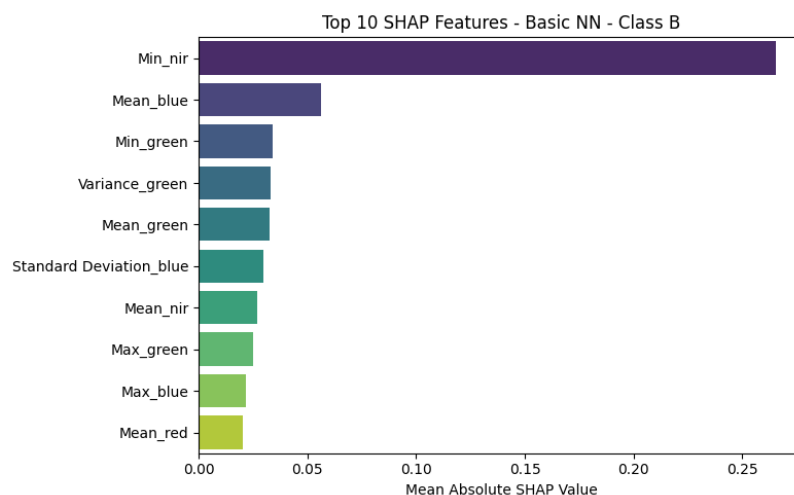


Figure 2 Top ten SHAP features contributing to the Bayesian Neural Network's classification of KDML-105, highlighting the dominant influence of Min_nir followed by spectral features from the blue and green bands

3.3.2 Feature importance analysis for Non-KDML-105 using the BNN model

Figure 3 illustrates the SHAP feature importance results for classifying Non-KDML-105 seeds using the BNN model. Similar to the KDML-105 analysis, Min_nir emerged as the most influential feature, with a SHAP value of approximately 0.27—substantially higher than all other variables. This underscores the critical role of NIR reflectance minima in distinguishing Non-KDML-105 seeds from KDML-105.

The second most informative feature was Mean_blue, which exhibited a SHAP value of around 0.14, indicating that variations in the blue spectral band contribute significantly to the classification. Other important features included Max_green, StandardDeviation_nir, and StandardDeviation_blue, each with SHAP values ranging from 0.06 to 0.08. Additional features such as Mean_nir, Mean_green, Count_green, and Sum_blue had slightly lower SHAP values (approximately 0.04–0.05), but still played a supportive role in improving classification performance by providing complementary spectral cues.

Overall, these results suggest that the BNN model relies predominantly on a small subset of spectral features—particularly those in the NIR and blue wavelength regions—to achieve high discriminative accuracy when classifying Non-KDML-105 seeds.

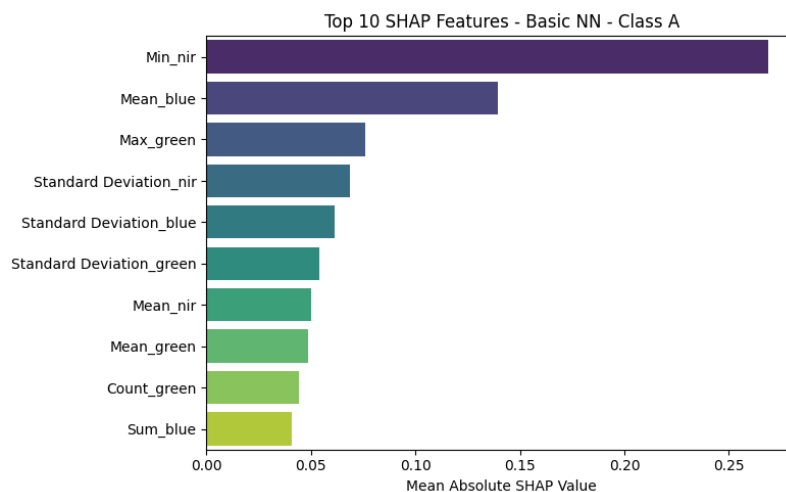


Figure 3 Top ten SHAP features contributing to the Bayesian Neural Network's classification of Non KDML-105, emphasizing the dominant role of Min_nir followed by blue- and green-band statistical attributes

An examination of the SHAP analysis for the BNN model revealed that both KDML-105 and Non-KDML-105 classes were influenced by the same two dominant features: Min_nir and Mean_blue. Among these, Min_nir had the highest contribution in both classes, underscoring the critical role of near-infrared (NIR) reflectance in distinguishing rice seeds that exhibit only subtle morphological differences [12, 36-38]. The NIR region is known to be sensitive to internal seed properties such as moisture content, surface structure, and biochemical composition, which likely explains its strong discriminative power in this study [39-41].

The second most influential feature, Mean_blue, reflects the importance of the blue spectral band, which is highly responsive to surface brightness and physical characteristics such as texture smoothness or roughness, specular highlights, and surface cleanliness [42, 43]. Despite the visual similarity between seed samples from the two groups, these blue-band variations provided meaningful cues that supported effective classification [44, 45].

These findings indicate that rice-seed discrimination is driven not merely by color differences but by finer spectral characteristics that are imperceptible to human vision [46, 47]. The model therefore relies primarily on NIR information, complemented by blue and green spectral features, to capture subtle distinctions within and between classes [46, 48]. From an engineering perspective, the SHAP results suggest that it may not be necessary to use a full multispectral imaging system for this application [36, 49]. Instead, a simplified sensing configuration that captures only NIR and blue channels could potentially achieve comparable performance while reducing hardware complexity [50, 51]. Reducing the number of input features would also decrease model complexity and improve processing speed, thereby enhancing the suitability of the system for real-time classification [52].

From a physical and biochemical perspective, the dominance of Min_nir as a discriminative feature is consistent with the established sensitivity of near-infrared reflectance to internal seed properties such as moisture content, starch granule structure, and endosperm density—characteristics that vary subtly among rice varieties as a function of grain composition and that are not accessible through visible-light inspection alone [39-41]. The minimum NIR reflectance value, in particular, is influenced by localized regions of the seed surface where internal structural or compositional differences produce the strongest absorption, making it a sensitive indicator of inter-variety variation even when overall seed morphology is similar. The secondary importance of Mean_blue is consistent with the sensitivity of the blue spectral region to surface-level optical properties, including surface texture (smoothness versus roughness), specular reflection from the seed hull, and surface cleanliness, all of which can differ subtly between KDML-105 and the non-target varieties as a consequence of differences in hull thickness, pericarp characteristics, and post-harvest handling [42, 43]. Taken together, these two features capture complementary information—internal compositional variation (NIR) and external surface characteristics (blue)—that jointly enable the BNN to distinguish seeds whose gross morphology is otherwise highly similar.

4. Conclusions

This study developed and evaluated Deep Learning models for detecting varietal contamination in KDML-105 jasmine rice seeds using multispectral imaging combined with morphological features. The results demonstrated that the BNN achieved the highest accuracy across all datasets containing multiple spectral bands, reaching 0.99 in both the Physical Reflectance and Multi-Reflectance datasets. Although the CNN delivered comparable performance under multispectral conditions, it showed a relative advantage when the data were reduced to RGB only, owing to its strength in extracting spatial image features. In contrast, the RNN consistently produced the lowest accuracy across all datasets, confirming that recurrent architectures are not well suited for static image classification tasks such as rice seed discrimination.

The SHAP analysis further revealed that the most influential features were Min_nir, Mean_blue, and several statistical descriptors in the green-blue spectral range. These findings highlight the dominant role of NIR reflectance in distinguishing KDML-105 from Non-KDML-105 seeds, supported by additional cues from the blue channel related to surface texture characteristics such as smoothness and roughness. Collectively, these results suggest that only a subset of key spectral bands is required for effective classification, even when the varieties exhibit minimal morphological differences.

Overall, the findings indicate that a BNN model combined with multispectral data offers the most suitable approach for developing a seed-contamination detection system. Its high accuracy, stability, and low computational cost make it well suited for real-world deployment in rice mills, seed production centers, or real-time sorting systems. Moreover, insights from the SHAP analysis provide practical benefits for engineering design, suggesting that a reduced-cost imaging system—such as one employing only NIR and blue bands—could serve as an effective alternative to a full multispectral camera, thereby lowering system costs and improving feasibility for industrial-scale implementation.

With respect to generalizability, the five-fold cross-validation protocol described in Section 2.1 ensured that each model was evaluated on seeds not seen during training within every fold, and the consistency of the BNN's accuracy across folds and across the Physical-Reflectance and Multi-Reflectance datasets indicates that its performance is not an artefact of a particular data partition. Nevertheless, all seeds in the present dataset were collected, imaged, and processed within the same experimental campaign, and the spectrally proximate RD-15 variety—owing to its genetic similarity to KDML-105 noted in Section 1—is expected to remain the most challenging case for any classifier operating on this feature representation. A full external validation of the Deep Learning models on independently collected seed batches, additional harvest seasons, and different growing locations is therefore identified as a priority for future work to confirm transferability beyond the conditions examined here.

From a practical implementation standpoint, the BNN's combination of high accuracy and the lowest processing time among the three architectures (Table 1) is consistent with deployment as an inline inspection module within an automated seed-sorting line: each seed (or seed batch, depending on the singulation mechanism) would be imaged by the multispectral camera under controlled illumination, passed through the segmentation and feature-extraction pipeline described in Sections 2.2–2.3, and classified by the trained BNN, with seeds flagged as Non-KDML-105 diverted for rejection or further manual inspection. The per-seed processing times reported in Table 1 (on the order of a few seconds for the BNN under multispectral conditions) indicate that, with appropriate pipelining of image acquisition and inference, throughput compatible with mechanical seed-handling rates is achievable. The hardware requirements correspond to the imaging system described in Section 2.1—a MicaSense RedEdge-MX multispectral camera, controlled illumination enclosure, and a standard computing unit for inference—representing a substantially lower capital cost than full hyperspectral imaging systems [32]. The SHAP-based identification of NIR and blue bands as the dominant discriminative features further suggests that a reduced-band sensor configuration could lower hardware cost still further; however, the cost, accuracy trade-offs, and engineering design of such a reduced-band system were not directly evaluated in this study and constitute a concrete direction for subsequent applied research. A formal techno-economic assessment, encompassing total system cost, throughput, and comparison with incumbent manual inspection methods, is identified as a priority for future work.

5. Acknowledgements

This work was supported by the National Research Council of Thailand (NRCT), Contract No. N42A670923. The authors gratefully acknowledge the Khon Kaen Rice Research Center for kindly providing the rice seed material used in the experiments, and the Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, for granting access to laboratory facilities.

6. References

- [1] Vanavichit A, Kamolsukyeunyong W, Siangliw M, Siangliw JL, Traprab S, Ruengphayak S, et al. Thai Hom Mali rice: origin and breeding for subsistence rainfed lowland rice system. *Rice*. 2018;11(1):20.
- [2] Kongpun A, Pusadee T, Jaksomsak P, Chinachanta K, Tuiwong P, Chan-In P, et al. Abiotic and biotic factors controlling grain aroma along value chain of fragrant rice: a review. *Rice Sci*. 2024;31(2):142-58.
- [3] Onmankhong J, Ma T, Inagaki T, Sirisomboon P, Tsuchikawa S. Cognitive spectroscopy for the classification of rice varieties: a comparison of machine learning and deep learning approaches in analysing long-wave near-infrared hyperspectral images of brown and milled samples. *Infrared Phys Technol*. 2022;123:104100.
- [4] Chen M, Song J, He H, Yu Y, Wang R, Huang Y, et al. Quantitative analysis of high-price rice adulteration based on near-infrared spectroscopy combined with chemometrics. *Foods*. 2024;13(20):3241.
- [5] Pechdin W, Naruepatr J. Toward rice export competitiveness, community development, and farmer well-being: evidence from Thailand. *Multidiscip Sci J*. 2025;7(9):e2025443.
- [6] Pitiphunpong S, Champangern S, Suwannaporn P. The Jasmine rice (KDML 105 variety) adulteration detection using physico-chemical properties. *Chiang Mai J Sci*. 2011;38(1):105-15.
- [7] Bist DR, Chapagae P, Rawal R, Bist G, Bohara R. Multivariate analysis of yield and phenotypic diversity in rainfed rice (*Oryza sativa* L.) landraces from Gokuleshwar, Baitadi, Nepal. *Discov Agric*. 2025;3:167.
- [8] Aznan A, Gonzalez Viejo C, Pang A, Fuentes S. Computer vision and machine learning analysis of commercial rice grains: a potential digital approach for consumer perception studies. *Sensors*. 2021;21(19):6354.
- [9] Van Hoai DP, Surinwarangkoon T, Hoang VT, Duong HT, Meethongjan K. A comparative study of rice variety classification based on deep learning and hand-crafted features. *ECTI-CIT Trans*. 2020;14(1):1-10.
- [10] Melesse MB, Maredia MK, Wossen T, Odeny D, Spielman DJ, Michelson H, et al. The role of genotyping in measuring improved variety adoption and impact: advances, challenges, and policy directions. *Food Policy*. 2025;137:102984.
- [11] Rahajaharilaza K, Muller B, Violle C, Brocke K, Ramavovololona, Morel JB, et al. Upland rice varietal mixtures in Madagascar: evaluating the effects of varietal interaction on crop performance. *Front Plant Sci*. 2023;14:1266704.
- [12] Jin B, Zhang C, Jia L, Tang Q, Gao L, Zhao G, et al. Identification of rice seed varieties based on near-infrared hyperspectral imaging technology combined with deep learning. *ACS Omega*. 2022;7(6):4735-49.
- [13] Liu W, Zeng S, Wu G, Li H, Chen F. Rice seed purity identification technology using hyperspectral image with LASSO logistic regression model. *Sensors*. 2021;21(13):4384.
- [14] Phan TTH, Vo QT, Nguyen HD. A novel method for identifying rice seed purity using hybrid machine learning algorithms. *Heliyon*. 2024;10(14):e33941.
- [15] Saha D, Manickavasagan A. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: a review. *Curr Res Food Sci*. 2021;4:28-44.
- [16] Ghaffari A. Precision seed certification through machine learning. *Technol Agron*. 2024;4:e019.
- [17] Ruslan R, Khairunniza-Bejo S, Jahari M, Ibrahim MF. Weedy rice classification using image processing and a machine learning approach. *Agriculture*. 2022;12(5):645.
- [18] He Y, Fan B, Sun L, Fan X, Zhang J, Li Y, et al. Rapid appearance quality of rice based on machine vision and convolutional neural network research on automatic detection system. *Front Plant Sci*. 2023;14:1190591.
- [19] Mittal S, Dutta MK, Issac A. Non-destructive image processing based system for assessment of rice quality and defects for classification according to inferred commercial value. *Measurement*. 2019;148:106969.
- [20] Singh SK, Jha R, Pandey S, Mohan C, Chetna, Ghosh S, et al. Artificial intelligence-based tools for next-generation seed quality analysis. *Crop Des*. 2025;4(1):100094.
- [21] Brutas MJB, Fajardo AL, Quilloy EP, Manuel LJR, Borja AA. Enhancing seed germination test classification for pole sitao (*Vigna unguiculata* (L.) Walp.) using SSD MobileNet and faster R-CNN models. *Appl Sci*. 2024;14(13):5572.
- [22] Feng Q, Wu X, He Y, Zhang L. Hyperspectral imaging combined with deep transfer learning for rice disease detection. *Front Plant Sci*. 2021;12:693521.
- [23] Weng S, Tang P, Yuan H, Guo B, Yu S, Huang L, et al. Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion. *Spectrochim Acta A Mol Biomol Spectrosc*. 2020;234:118237.
- [24] Wang S, Xu D, Liang H, Bai Y, Li X, Zhou J, et al. Advances in deep learning applications for plant disease and pest detection: a review. *Remote Sens*. 2025;17(4):698.
- [25] Upadhyay A, Chandel NS, Singh KP, Chakraborty SK, Nandede BM, Kumar M, et al. Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Artif Intell Rev*. 2025;58(3):92.
- [26] Das PK, Rupa SS, Pumrin S, Das UC, Hossen MK. Deep learning for plant disease detection and classification: a systematic analysis and review. *Curr Appl Sci Technol*. 2024;24(4):e0259016.
- [27] Jin C, Zhou L, Pu Y, Zhang C, Qi H, Zhao Y. Application of deep learning for high-throughput phenotyping of seed: a review. *Artif Intell Rev*. 2025;58:76.
- [28] Díaz-Martínez V, Orozco-Sandoval J, Manian V, Dhatt BK, Walia H. A deep learning framework for processing and classification of hyperspectral rice seed images grown under high day and night temperatures. *Sensors*. 2023;23(9):4370.
- [29] Phan TTH, Nguyen LHB. Enhancing rice seed purity recognition accuracy based on optimal feature selection. *Ecol Inform*. 2025;86:103044.
- [30] Islam MM, Himel GMS, Moazzam MG, Uddin MS. Artificial intelligence-based rice variety classification: a state-of-the-art review and future directions. *Smart Agric Technol*. 2025;10:100788.

- [31] Tang Z, Ma S, Qi H, Zhang X, Zhang C. Nondestructive detection of rice milling quality using hyperspectral imaging with machine and deep learning regression. *Foods*. 2025;14(11):1977.
- [32] Zhao Y, Xie Q. Review of deep learning applications for detecting special components in agricultural products. *Computers*. 2025;14(8):309.
- [33] Izquierdo M, Lastra-Mejías M, González-Flores E, Pradana-López S, Cancilla JC, Torrecilla JS. Visible imaging to convolutionally discern and authenticate varieties of rice and their derived flours. *Food Control*. 2020;110:106971.
- [34] Śliwińska-Bartel M, Burns DT, Elliott C. Rice fraud a global problem: a review of analytical tools to detect species, country of origin and adulterations. *Trends Food Sci Technol*. 2021;116:36-46.
- [35] Srinuttrakul W, Mihailova A, Islam MD, Liebis B, Maxwell F, Kelly SD, et al. Geographical differentiation of Hom Mali rice cultivated in different regions of Thailand using FTIR-ATR and NIR spectroscopy. *Foods*. 2021;10(8):1951.
- [36] Ahmed MT, Monjur O, Khaliduzzaman A, Kamruzzaman M. A comprehensive review of deep learning-based hyperspectral image reconstruction for agri-food quality appraisal. *Artif Intell Rev*. 2025;58:96.
- [37] Fabiyi SD, Vu H, Tachtatzis C, Murray P, Harle D, Dao TK, et al. Varietal classification of rice seeds using RGB and hyperspectral images. *IEEE Access*. 2020;8:22493-505.
- [38] Qi H, Huang Z, Sun Z, Tang Q, Zhao G, Zhu X, et al. Rice seed vigor detection based on near-infrared hyperspectral imaging and deep transfer learning. *Front Plant Sci*. 2023;14:1283921.
- [39] Sun H, Zhang L, Rao Z, Ji H. Determination of moisture content in barley seeds based on hyperspectral imaging technology. *Spectrosc Lett*. 2020;53(10):751-62.
- [40] Saha D, Senthilkumar T, Sharma S, Singh CB, Manickavasagan A. Application of near-infrared hyperspectral imaging coupled with chemometrics for rapid and non-destructive prediction of protein content in single chickpea seed. *J Food Compos Anal*. 2023;115:104938.
- [41] Reddy P, Guthridge KM, Panozzo J, Ludlow EJ, Spangenberg GC, Rochfort SJ. Near-infrared hyperspectral imaging pipelines for pasture seed quality evaluation: An overview. *Sensors*. 2022;22(5):1981.
- [42] Zhang X, Liu J, Meng L, Qin C, An Z, Wang Y, et al. Enhanced blue band vegetation index (The re-modified anthocyanin reflectance index (RMARI)) for accurate farmland shelterbelt extraction. *Remote Sens*. 2024;16(19):3680.
- [43] Kior A, Sukhov V, Sukhova E. Application of reflectance indices for remote sensing of plants and revealing actions of stressors. *Photonics*. 2021;8(12):582.
- [44] Hennessy A, Clarke K, Lewis M. Hyperspectral classification of plants: a review of waveband selection generalisability. *Remote Sens*. 2020;12(1):113.
- [45] Falcioni R, Antunes WC, Demattê JAM, Nanni MR. Reflectance spectroscopy for the classification and prediction of pigments in agronomic crops. *Plants*. 2023;12(12):2347.
- [46] Bhargava A, Sachdeva A, Sharma K, Alsharif MH, Uthansakul P, Uthansakul M. Hyperspectral imaging and its applications: a review. *Heliyon*. 2024;10(12):e33208.
- [47] Jeong SW, Lyu JI, Jeong HW, Baek J, Moon JK, Lee C, et al. SUnSeT: Spectral unmixing of hyperspectral images for phenotyping soybean seed traits. *Plant Cell Rep*. 2024;43:164.
- [48] Kaushik M, Nidamanuri RR, Aparna B. Hyperspectral discrimination of vegetable crops grown under organic and conventional cultivation practices: a machine learning approach. *Sci Rep*. 2025;15:7897.
- [49] Yan YJ, Wong WK, Chen CJ, Huang CC, Chien JT, Ou-Yang M. Hyperspectral signature-band extraction and learning: an example of sugar content prediction of *Syzygium samarangense*. *Sci Rep*. 2023;13:15100.
- [50] Tan Y, Gu J, Lu L, Zhang L, Huang J, Pan L, et al. Hyperspectral band selection for crop identification and mapping of agriculture. *Remote Sens*. 2025;17(4):663.
- [51] Lazar A, Meir I, Lati RN, Bechar A. Weed detection: innovative hyperspectral image analysis for classification and band selection of site-specific and selective weeding robot. *Agronomy*. 2025;15(11):2576.
- [52] Salhi A, Alshamrani R, Althbiti A, Ismail A, Abd-ElRahman M, Hassan BM. Optimizing high dimensional data classification with a hybrid AI driven feature selection framework and machine learning schema. *Sci Rep*. 2025;15(1):35038.