

Supplementary Materials

Title: Multispectral Imaging Combined with Tree-Based Ensemble Classifiers for Non-destructive Varietal Purity Assessment of KDML-105 Rice Seed

Table S1. Two-Way Analysis of Variance (ANOVA) Summary for Classification Performance Metrics

Type III Sum of Squares for the two-way ANOVA model (Classifier $\times$ Illumination Level) applied to each performance metric. Each experimental cell comprised five-fold cross-validation replicates ($n = 5$ per model–level combination), yielding 60 total observations per metric. Design: Intercept + Model + Level + Model $\times$ Level.

Dependent Variable: Accuracy

Source	Type III SS	df	MS	F	p	Partial η^2
Corrected Model	.584	11	.053	27.270	<.001	.862
Intercept	34.134	1	34.134	17519.446	<.001	.997
Model (Classifier)	.519	2	.259	133.092	<.001	.847
Illumination Level	.024	3	.008	4.073	.012	.203
Model $\times$ Level	.042	6	.007	3.594	.005	.310
Error	.094	48	.002			
Total	34.812	60				
Corrected Total	.678	59				

Dependent Variable: F1-score

Source	Type III SS	df	MS	F	p	Partial η^2
Corrected Model	.918	11	.083	8.336	<.001	.656
Intercept	32.120	1	32.120	3209.848	<.001	.985
Model (Classifier)	.665	2	.332	33.215	<.001	.581
Illumination Level	.065	3	.022	2.163	.105	.119
Model $\times$ Level	.188	6	.031	3.130	.011	.281
Error	.480	48	.010			
Total	33.517	60				
Corrected Total	1.398	59				

Dependent Variable: Precision

Source	Type III SS	df	MS	F	p	Partial η^2
Corrected Model	.392	11	.036	3.434	.001	.440
Intercept	40.155	1	40.155	3866.990	<.001	.988
Model (Classifier)	.329	2	.164	15.822	<.001	.397
Illumination Level	.033	3	.011	1.056	.376	.062
Model $\times$ Level	.031	6	.005	.493	.811	.058
Error	.498	48	.010			
Total	41.045	60				
Corrected Total	.891	59				

Dependent Variable: Recall

Source	Type III SS	df	MS	F	p	Partial η^2
Corrected Model	.760	11	.069	2.129	.036	.328
Intercept	30.626	1	30.626	943.514	<.001	.952
Model (Classifier)	.175	2	.087	2.693	.078	.101
Illumination Level	.142	3	.047	1.455	.239	.083
Model $\times$ Level	.444	6	.074	2.277	.052	.222
Error	1.558	48	.032			
Total	32.944	60				
Corrected Total	2.318	59				

Note: SS = Sum of Squares; df = degrees of freedom; MS = Mean Square; η^2 = partial eta squared. Bold rows denote the primary effects of interest (Model, Illumination Level, and their interaction). Error degrees of freedom (df = 48) and Mean Square Error are identical to those underlying the post-hoc comparisons in Table S2, confirming that both the Bonferroni-corrected pairwise comparisons and the Tukey's HSD test were computed from the same residual variance of the full two-way model.

Table S2. Post-Hoc Pairwise Comparisons of Classifier (Model) Performance

Pairwise comparisons among Extra Trees (ET), Random Forest (RF), and Support Vector Machine (SVM), based on estimated marginal means from the two-way ANOVA model. Bonferroni-corrected *p*-values are reported as the primary inferential statistic in the main text (Table 3); Tukey's HSD *p*-values, together with the associated homogeneous subset grouping, are reported here to support the descriptive classifier-grouping statement in Section 3.2.1. Both tests share the same error term (see Table S1).

Dependent Variable: Accuracy (Error MS = .002, df = 48)

Pairwise Comparison	Mean Difference	SE	Tukey HSD p	Bonferroni p
ET vs RF	.0144	.01396	.559	.917
ET vs SVM	.2040*	.01396	<.001	<.001
RF vs SVM	.1896*	.01396	<.001	<.001

Tukey HSD homogeneous subsets ($\alpha = .05$): Subset 1: SVM ($M = .623$); Subset 2: RF ($M = .813$), ET ($M = .827$) — Tukey HSD, $\alpha = .05$

Dependent Variable: F1-score (Error MS = .010, df = 48)

Pairwise Comparison	Mean Difference	SE	Tukey HSD p	Bonferroni p
ET vs RF	.0140	.03163	.898	1.000
ET vs SVM	.2300*	.03163	<.001	<.001
RF vs SVM	.2160*	.03163	<.001	<.001

Tukey HSD homogeneous subsets ($\alpha = .05$): Subset 1: SVM ($M = .583$); Subset 2: RF ($M = .799$), ET ($M = .813$) — Tukey HSD, $\alpha = .05$

Dependent Variable: Precision (Error MS = .010, df = 48)

Pairwise Comparison	Mean Difference	SE	Tukey HSD p	Bonferroni p
ET vs RF	.0240	.03222	.739	1.000
ET vs SVM	.1676*	.03222	<.001	<.001
RF vs SVM	.1436*	.03222	<.001	<.001

Tukey HSD homogeneous subsets ($\alpha = .05$): Subset 1: SVM ($M = .714$); Subset 2: RF ($M = .858$), ET ($M = .882$) — Tukey HSD, $\alpha = .05$

Dependent Variable: Recall (Error MS = .032, df = 48)

Pairwise Comparison	Mean Difference	SE	Tukey HSD p	Bonferroni p
ET vs RF	.0059	.05697	.994	1.000
ET vs SVM	.1174	.05697	.109	.135
RF vs SVM	.1114	.05697	.134	.169

Tukey HSD homogeneous subsets ($\alpha = .05$): Single homogeneous subset (all three classifiers): SVM ($M = .638$), RF ($M = .750$), ET ($M = .756$) — Tukey HSD, $\alpha = .05$; not significant, consistent with the non-significant main effect of model on recall ($F(2,48) = 2.693$, $p = .078$)

* The mean difference is significant at the .05 level. Full SPSS output, including Levene's test of equality of error variances, estimated marginal means, simple main-effects tests, and profile plots for all four dependent variables, is available from the corresponding author upon request and has been retained by the authors for the purpose of editorial verification.

Table S3. Experiment Setup Configuration (PyCaret setup())

Parameter	Value	Description
train_size	0.80	Proportion of data allocated to the training set (80/20 split)
fold	5	Number of cross-validation folds
fold_strategy	stratifiedkfold (default)	Stratified k-fold cross-validation, preserving class proportions in each fold
normalize	False	Feature normalisation was not applied prior to model fitting
session_id	3934	Fixed random seed applied across data splitting and all classifiers, ensuring reproducibility
n_jobs	-1 (default)	All available CPU cores used for model fitting

Table S4. Classifier Hyperparameter Configuration

Explicitly set parameters were configured by the PyCaret model container; all remaining parameters retain scikit-learn (version compatible with PyCaret 3.3.0, >1.4.0) default values for the respective estimator class.

Classifier	scikit-learn Class	Category	Parameter = Value
Extra Trees (ET)	<i>sklearn.ensemble. ExtraTreesClassifier</i>	Explicitly set (PyCaret)	random_state = 3934
			n_jobs = -1
		scikit-learn default (unchanged)	n_estimators = 100
			criterion = 'gini'
			max_depth = None (unconstrained)
			min_samples_split = 2
			min_samples_leaf = 1
			max_features = 'sqrt'
			bootstrap = False
			class_weight = None

Classifier	scikit-learn Class	Category	Parameter = Value
Random Forest (RF)	<i>sklearn.ensemble. RandomForestClassifier</i>	Explicitly set (PyCaret)	random_state = 3934
			n_jobs = -1
		scikit-learn default (unchanged)	n_estimators = 100
			criterion = 'gini'
			max_depth = None (unconstrained)
			min_samples_split = 2
			min_samples_leaf = 1
			max_features = 'sqrt'
			bootstrap = True
			class_weight = None

Classifier	scikit-learn Class	Category	Parameter = Value
SVM (linear kernel)	<i>sklearn.linear_model. SGDClassifier</i>	Explicitly set (PyCaret)	loss = 'hinge'
			penalty = 'l2'
			tol = 0.001
			eta0 = 0.001
			random_state = 3934
			n_jobs = -1
		scikit-learn default (unchanged)	alpha = 0.0001
			l1_ratio = 0.15
			fit_intercept = True
			max_iter = 1000
			learning_rate = 'optimal'
			class_weight = None

Note: PyCaret's 'svm' model identifier corresponds to "SVM – Linear Kernel", implemented internally via `sklearn.linear_model.SGDClassifier` (hinge loss, approximating a linear support vector classifier through stochastic gradient descent) rather than `sklearn.svm.SVC`. All parameter values were confirmed against the official PyCaret 3.3.0 package (`pycaret/containers/models/classification.py`) obtained from PyPI, and cross-checked against the scikit-learn (>1.4.0) API reference for each estimator class.