FUDMA Journal of Engineering and Technology
ISSN: 3092-9385
https://fjet.fudutsinma.edu.ng/index.php/fjet/
Volume 2, Issue 1, 2026, 910-922

Classification of Groundnut Pod Varieties using Histogram-Based Gradient Booster Classifier (HGBC) with Machine Learning Techniques

Benjamin BELLO[1*], Jibril D. JIYA[2], Aje TOKAN[3], Ahmed MOHAMMED[4]

[1*,2,4]Department of Electrical and Electronics Engineering, Abubakar Tafawa Balewa University, Bauchi, Nigeria
[3]Department of Mechanical and Production Engineering, Abubakar Tafawa Balewa University, Bauchi, Nigeria

[1*]benburgaper1@gmail.com, [2]jibjiya27@gmail.com, [3]ajetokan@atbu.edu.ng, [4]amohammed3@atbu.edu.ng

Abstract

Groundnut (Arachis hypogaea) is an important economic and food crop widely cultivated for human consumption, livestock feed and industrial applications. Groundnut pod varieties differ in size, shape, and surface characteristics (texture) which influence the threshing efficiency, seed damage, cleaning performance, and threshing machine parameter for post-harvest systems. However, conventional pod variety identification in most local processing environments is still performed manually, making the process labour-intensive, subjective, slow, and unsuitable for integration into intelligent threshing systems. This study developed a machine vision-based classification framework for automated classification of groundnut pod varieties using a Histogram Based Gradient Boosting Classifier (HGBC) that support intelligent threshing machine design and operation. A total of 360 groundnut pod images comprising Exdakar (116), Jarma (99), and Samnut26 (145) varieties were acquired under controlled imaging conditions. The methodology involved image acquisition, preprocessing, segmentation, geometric and texture feature extraction, dimensionality assessment using Principal Component Analysis (PCA), and supervised machine learning classification. Extracted engineering features included pod length, width, height, area, perimeter, aspect ratio, circularity, weight, and texture, which are critical parameters for machine interaction analysis and threshing component design. The dataset was divided into a 70:15:15 ratio i.e 70% training, 15% validation, and 15% testing for model development and evaluation. The HGBC model achieved an overall classification accuracy of 91.9%, with model performance of 95% precision, 95% recall, and 95% F1-score for Exdakar, 88% precision, 95% recall, and 91% F1-score for Jarma, and 93% precision, 88% recall, and 90% F1-score for Samnut26. The results demonstrate the technical feasibility of integrating machine vision techniques and intelligent classification into groundnut threshing systems for adaptive operations, improved threshing efficiency, and reduced seed damage.

***Keywords**: Groundnut, Classification, Intelligent, Accuracy, Threshing.*

1.0 Introduction

Agricultural produce plays a vital role in feeding both human and livestock populations globally. Groundnut (*Arachis hypogaea*) is an important economic oilseed crop that is cultivated worldwide for food, oil production, and industrial purposes [1]. In Nigeria and other developing countries, groundnut serves as a vital source of food, animal feed, and also contributes significantly to rural income generation [2]. Variations in physical characteristics of groundnut pods among different varieties influence the threshing processing efficiency, grading, and its market value. Traditionally, conventional methods of identification also classification of groundnut pods varieties relies on manual inspection by agricultural experts or operator, which is labour intensive, subjective, and prone to error and inconsistencies, especially when large volumes of groundnut pods must be evaluated [3]

Image processing technique refer to computerized manipulation and data analysis of the images of pods been captured using a wide array of sensors including visible light cameras, infra-red imaging devices and sensors operating in different resolution and pixels [4]. Recent advances in machine vision and artificial intelligence have enabled automated systems capable of performing agricultural inspection tasks with high accuracy and speed. Machine learning models can analyze digital images and extract discriminative features such as pods size, pods shape, pods texture, and colour for classification purposes [5]. These approaches reduce human dependency and improve reliability in agricultural decision-making for better operation. However, despite the advancement of machine learning in agriculture, only a few studies have focused on automated classification of crops. Most existing studies emphasize on disease detection or leaf classification rather than pods identification. Therefore, this study focuses on machine learning based framework and HGBC to accurately classify groundnut pods varieties. Therefore, HGBC is an optimize ensemble machine learning algorithm frame system, designed for efficiency and scalability, unlike traditional gradient boosting methods, HGBC describe its feature values into bins before training decision trees [6]. This process reduces the computational complexity, accelerates split finding, and lowers memory consumption [5].

2.0 Literature Review

In recent time, machine learning is arguably the most successful and widely used technique to address challenges that cannot be solved by hand crafted programs. When compared to conventional algorithms following

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

a predefined set of rules, a machine learning algorithm relies on a large amount of data that is observed in nature and generated by another algorithm [7]. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data [8]. Machine learning is a field of study that gives computers the ability to learn without being explicitly programmed [7]. Machine vision has become a fundamental component of precision agriculture, enabling automated monitoring, grading, and classification of agricultural products such as groundnuts. Image-based systems can analyze visual properties of crops and provide objective assessments faster than manual inspection [9]. The available data structured, datasets is important for training machine learning models, also studies involving groundnut leaf image datasets show that properly labeled and balanced data significantly improve classification accuracy and model generalization [10]. Many machine learning approaches have been applied to agricultural classification tasks such as Convolutional Neural Networks (CNN) have shown strong performance in plant recognition problems and the effectiveness of deep learning for agricultural image analysis [11]. Furthermore, combining feature extraction techniques such as Histogram of Oriented Gradients (HOG) with classical classifiers like K-Nearest Neighbor has achieved high accuracy for plant disease detection tasks [12]. Another deep learning model based on transfer learning using Visual Geometric Group 16- layer Network (VGG16) architecture achieved good result classification accuracy for plant disease identification, outperforming several baseline models [13]. According [14] an appropriate classification scheme and adequate amount of training samples are basics for an effective classification system, classification system is deliberate depends on users' requirements. [15] Investigated the engineering properties of groundnut pods for advanced pneumatic pod collection systems.

The study considered eighteen (18) different groundnut varieties, selecting twenty (20) samples from each for analysis. A digital vernier caliper was used to measure the size of groundnut pods in terms of width (W), length (L), and thickness (T). The results showed that each variety had different average pod length, breadth, thickness, geometric mean, and sphericity. To evaluate the efficacy and efficiency of the pneumatic groundnut pod collector (GPC), certain engineering properties of groundnuts must be established. However, the digital vernier caliper measures only at specific points, which may not fully capture the true variability in groundnut, pod dimensions, particularly due to their irregular shapes. [16] Worked on handwritten Devanagari word recognition using a customized CNN. The model uses CNN for classification and consists of three convolutional layers. It achieved an accuracy of 94% on training data and 96% on testing data. However, using only 128 samples for training and testing is insufficient for a CNN model. To achieve better accuracy, a large, well-labeled dataset is required for classification. Therefore, these results cannot be considered fully validated for an accurate model evaluation. [17] worked on an intelligent classification model for peanut's varieties by colour and texture features. The research proposed a testing method according to image processing and computer vision the process is good, fast with high differences in rate. The machine learning method used were Support Vector Machine (SVM), Random Forest (RF), and Multilayer Perceptron (MLP) with an accuracy result of SVM had 86.07%, RF had 82.27% and MLP had 84.9%, thus the research has shortcoming of using only colour and texture for it classification which will not captured other important features that will influence the variety classification.

3.0 Methodology

The classification of the three varieties of groundnut pods such as Exdakar, Samnut26 and Jarma using HGBC is done under the following process as shown in Figure 1:

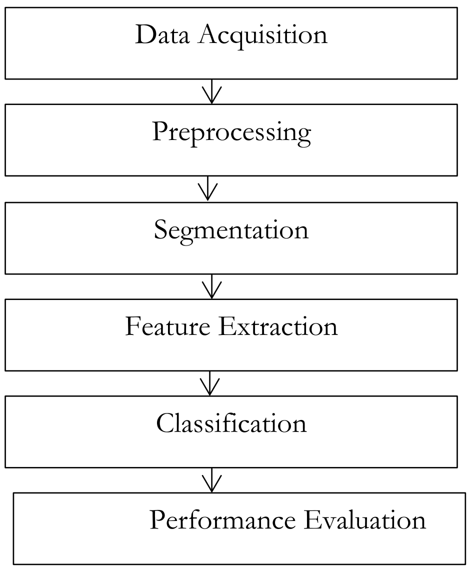

Figure1: HGBC Methodology Block Diagram

3.1 Data Acquisition

The groundnut pods were obtained from muda lawan market in Bauchi, Nigeria. However, the data acquisition is the initial stage in data acquisition process, the process begins by putting the groundnut pods on a white paper and captured the images using camera under consistent lighting conditions and the camera were set in a good position where a cleared and uniform pod images were acquired for further processes as shown in Figure 2. The camera has a good resolution of 20 megapixels was used to capture all the groundnut images at equal distance of 20cm under the same light intensities. For the purpose of journal research the following numbers of groundnut pods were selected Ex-dakar variety one hundred and sixteen (116) images was captured, Jarma variety ninety-nine (99) images was captured and Samnut26 variety one hundred and forty-five (145) images was also captured.

Figure 2: Varieties of Groundnut Pods

3.2 Preprocessing

Preprocessing is the process whereby the groundnut pod image quality is enhanced by removing the unwanted information. The processes include background removal, noise filtering, colour normalization, and resizing. These procedures ensure that segmentation, extracted features are accurately carried out on the groundnut pod as indicated in Figure 3

Figure 3: Preprocessed Groundnut Pod

3.3 Segmentation

Segmentation of groundnut pods is the process of dividing them into smaller sections for better assessment, edge detection that identify and separate specific regions of interest within the images. The technique helps in achieving the separation of groundnut pods from the background through setting a threshold value based on pixel intensity, edge detection i.e boundaries of the groundnut pods, the region-base and can be segmentation according to pixels base on their similarity in colour, textures, sizes and measurement of the groundnut pods whereby, the final output is a clean cropped image of size 1000x1000 pixels as shown in Figure 4 and the mathematical model for segmentation were adopted from [18]

$$S = \begin{cases} 1 \text{ if } I\,(x,y) \geq T \\ 0 \text{ if otherwise} \end{cases} \quad \ldots (1)$$

$$R = S(I) \quad \ldots (2)$$

Where I (x,y) = pixel intensity at location (x,y) of grayscale image
T = Threshold value chosen that separate the dark groundnut pod from the light background
S = Binary mask for groundnut pod region

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

R = Segmented region of the binary mask of the groundnut pod


Figure 4 Cropped Image

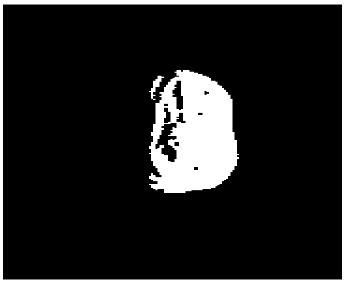
Figure 5: Initial Mask

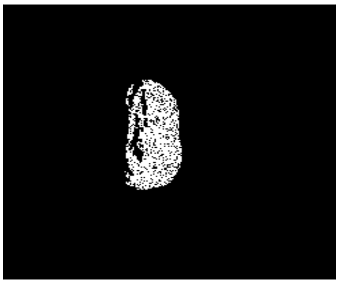
Figure 6: Refined Mask

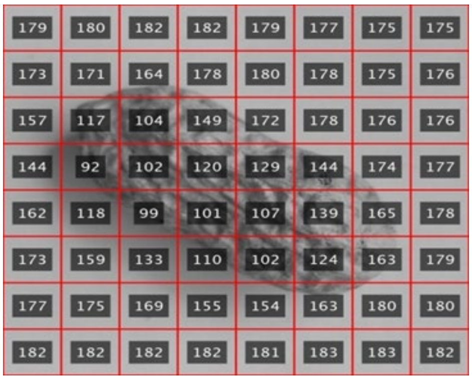

Figure 7: Labeled Intensity


Figure 8: Segmentation

However, the initial mask and refined mask in Figure 5 to 6 represents the cleaned binary segmentation output, while the labelled intensity image in Figure 7 assigns unique intensity values to segmented groundnut pods for accurate identification and analysis as indicated in Figure 8.

3.4 Feature Extraction

Feature extraction identifies key characteristics relevant features in a groundnut pod image, such as shape, texture, and colour and texture for groundnut pod variety recognition as indicated in Figure 9. Furthermore, this feature extraction of groundnut pods is extracted geometric, pod properties such as area, perimeter, sphericity, and aspect ratio to characterize the shape also calculate texture descriptors that capture specific textural properties of groundnut pods. However, the mathematical model for feature extraction is expressed in equation 3 was adopted from [19].

$$F_{pod} = \Phi = f(A, P, AR, \mu, M) \quad \ldots(3)$$

F_{pod} = Segmented groundnut pod
f = feature extraction function that maps raw segmented image
A= Area of the pod
P= Perimeter of the pod
AR= Aspect ratio of width and height
μ = intensity colour value of segmented groundnut pod
M = Shape moment (Hu invariant moment)
Φ = Feature extraction

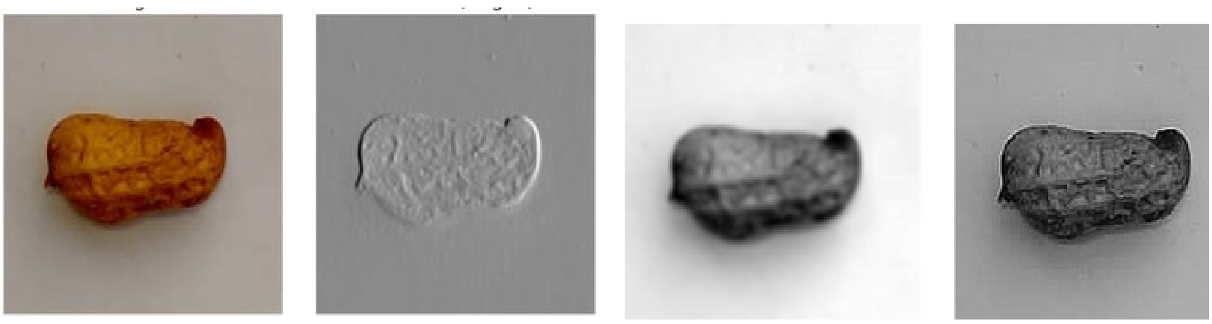

Figure 9: Feature Extraction

3.5 Classification

The classification of groundnut pod images involves using a trained machine learning model such as HGBC, the selected features are used to train the model by the built the algorithm iteratively builds decision trees that minimize prediction errors by optimizing a loss function. The features values obtained from the pod are discretized into histogram, which accelerates training and enables efficient handling of large datasets as the show in Figure 10, and the mathematical model is obtained from [19]

$$\dot{y} = C(F_{pod}) \quad \ldots (4)$$

Can also be written a $\dot{y} = C(\Phi(S(I(x, y))))$ … (5)

$\dot{y}$ = Predicted output class label (Exdakar, Jarma and Samnut26)
C = Classifier (HGBC layer)
Φ = Feature extraction
S = Segmentation function
I (x,y) = Input image of the groundnut pod

Figure 10: Classification of Groundnut Pods Varieties

3.6 Model Evaluation

Model performance was evaluated using standard performance metrics as adopted from [20]

i. Accuracy is the ratio of all the correct classified samples and to the total number of samples

$$\text{Accuracy} = \frac{TP}{TP+TN+FP+FN} \quad \ldots (6)$$

ii. Precision is the proportion of true positives out of the total predicted positives

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots (7)$$

iii. Recall is the proportion of positive samples of classified as true

$$\text{Recall} = \frac{TP}{TP+FN} \quad \ldots (8)$$

iv. F1-Score is the harmonic mean of recall and precision

$$F1 = 2 * \frac{\text{Precision} * Recall}{\text{Precision} + Recall} \quad \ldots (9)$$

4.0 Results and Discussion

This chapter presents the results obtained from the analysis of groundnut images for geometric, assessment of the machine learning model and it's the performance evaluation. The following results were obtained and presented.

4.1 Groundnut Geometric Properties

The groundnut pods true geometric dimensions, regardless of its orientation of the groundnut images as shown in table 11 to 13 which are in matric form. The physical and the engineering properties of different varieties of the groundnut such as Exdakar, Jarma, and Samnut26 were achieved using PCA implemented using from Python modules. This research targets the dimensional reduction of the groundnut pods to established its different sizes for the purposed of having a quality threshing efficiency of the groundnut seeds due to their difference in

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

size where by the software is embedded with this equation indicated in equation (10). This matrix in Figure 11 to 13 indicated the average different sizes of Exdakar, Samnut26 and Jarma varieties the purpose of identification, classification and for better threshing operation.

$$\text{Geometric Mean Diameter } (D_g) \text{ Size } = \sqrt[3]{L \times W \times T} \quad \ldots (10)$$

Table 1: Geometric of External Variety

1.55	1.56	1.58	1.59	1.59	1.59	1.61	1.62
1.57	1.6	1.56	1.59	1.61	1.6	1.6	1.61
1.46	1.12	1.08	1.19	1.27	1.44	1.61	1.6
1.24	0.73	0.92	0.94	1.02	1.09	1.45	1.65
1.33	0.76	0.62	0.69	0.77	0.8	1.35	1.67
1.42	1.19	0.9	0.76	0.72	1.05	1.55	1.64
1.47	1.43	1.38	1.29	1.27	1.43	1.59	1.63
1.51	1.5	1.5	1.5	1.52	1.57	1.62	1.64

Table 2: Geometric of Samnut26 Variety

1.64	1.63	1.63	1.64	1.65	1.63	1.61	1.59
1.62	1.64	1.68	1.68	1.66	1.6	1.62	1.65
1.65	1.63	1.51	1.51	1.57	1.37	1.06	1.37
1.7	1.42	1.24	1.42	1.51	1.27	0.69	0.5
1.48	1.04	1.02	1.38	1.38	0.96	0.48	0.28
1.59	0.91	0.55	0.82	0.94	0.49	0.31	0.29
1.64	1.69	1.18	0.5	0.59	0.51	0.41	0.43
1.63	1.64	1.71	1.32	0.75	0.66	0.61	0.54

Table 3: Geometric of Jarma Variety

1.73	1.76	1.78	1.8	1.88	1.76	1.76	1.8
1.71	1.74	1.76	1.78	1.3	0.93	1.11	1.74
1.69	1.71	1.81	1.15	0.55	0.77	0.8	1.62
1.67	1.65	1.15	0.65	0.6	0.47	1.09	1.76
1.63	0.93	0.48	0.56	0.52	1.15	1.71	1.71
1.41	0.52	0.23	0.42	1.1	1.72	1.68	1.69
1.39	1.11	1.02	1.45	1.65	1.65	1.67	1.68
1.49	1.56	1.63	1.63	1.58	1.63	1.66	1.65

4.3 Classification of Groundnut Variety

The groundnut varieties were trained on an open-source of machine learning Python library module where an algorithm model was developed and were tested with the dataset as Ex-Dakar, Samnut26 and Jarma and the predictions results were displayed on the confusion matrix (Decision matrix) as shown in Figure 11 and also the Correlation map in Figure 12. The confusion matrix displayed the classification results of the validated data of each groundnut pods variety that are presented in a diagonal (Dark blue) and the misclassified results off the diagonal in light blue.

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

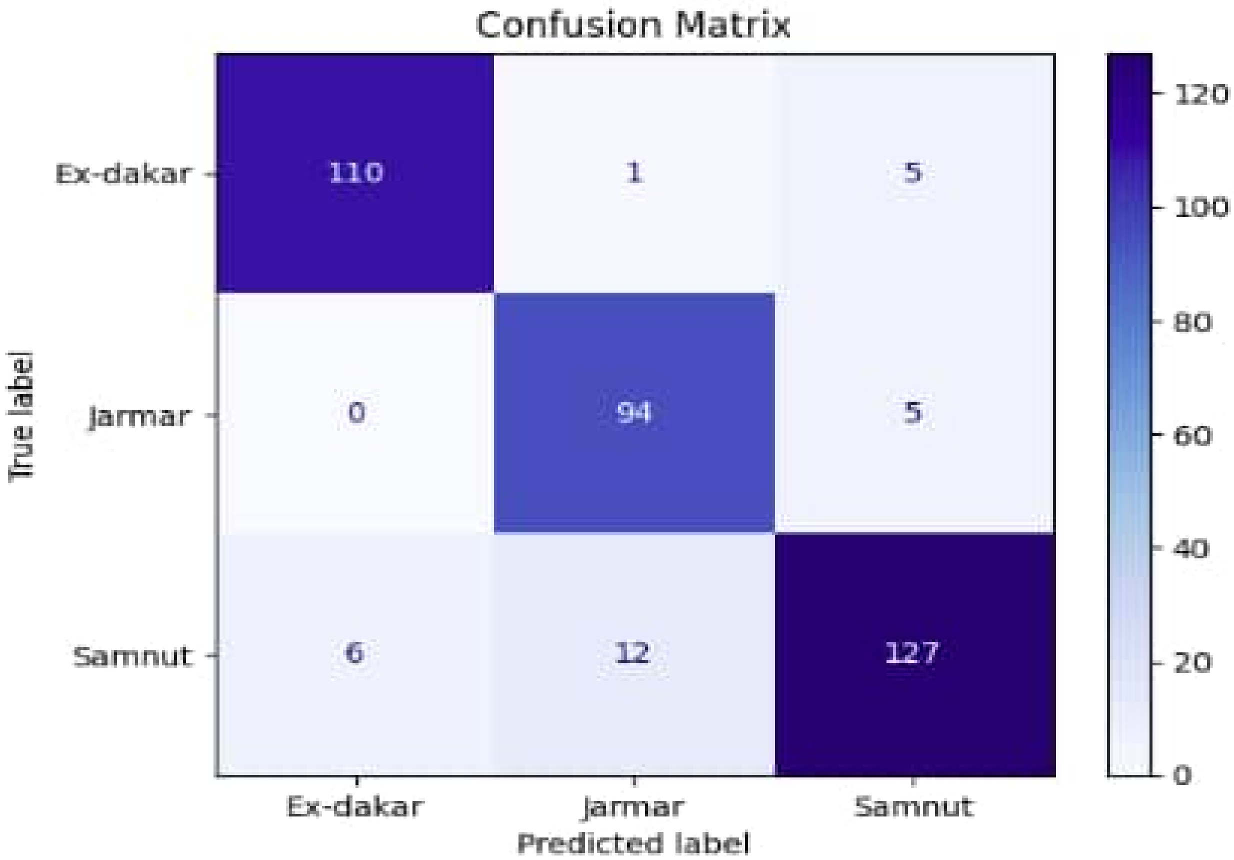

Figure 11: Confusion Matrix of HGBC

The confusion matrix presents a detailed assessment of model's classification performance across three distinct categories such as Exdakar, Jarma and Samnut26 in Figure 11. The matrix reveals a variation of the model's predictive capabilities, highlighting both its strengths and weak areas. The diagonal elements in dark blue represent correctly classified dataset where 110 of Ex-dakar, 94 of Jarma, and 127 of Samnut26 and the off-diagonal elements in light blue indicate misclassifications. The samples were correctly predicted for the Ex-dakar class, the model demonstrates high precision correctly classified 110 dataset, instead of 116 with only 1 misclassified as Jarma and 5 as Samnut26, and this implies that the model has a robust ability to identify Ex-dakar samples. Similarly, for Jarma out of 99 samples 94 was predicted correctly but had 5 samples misclassified as Samnut26 but none as Ex-dakar, these demonstrated relatively high classification accuracy. The Samnut26 class shows some high accuracy models prediction with 127 correct classifications out of 145, though it suffers from misclassifications across both the two (2) classes, particularly Jarma with 12 incorrect predictions, and 6 incorrect predictions as Ex-dakar.

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

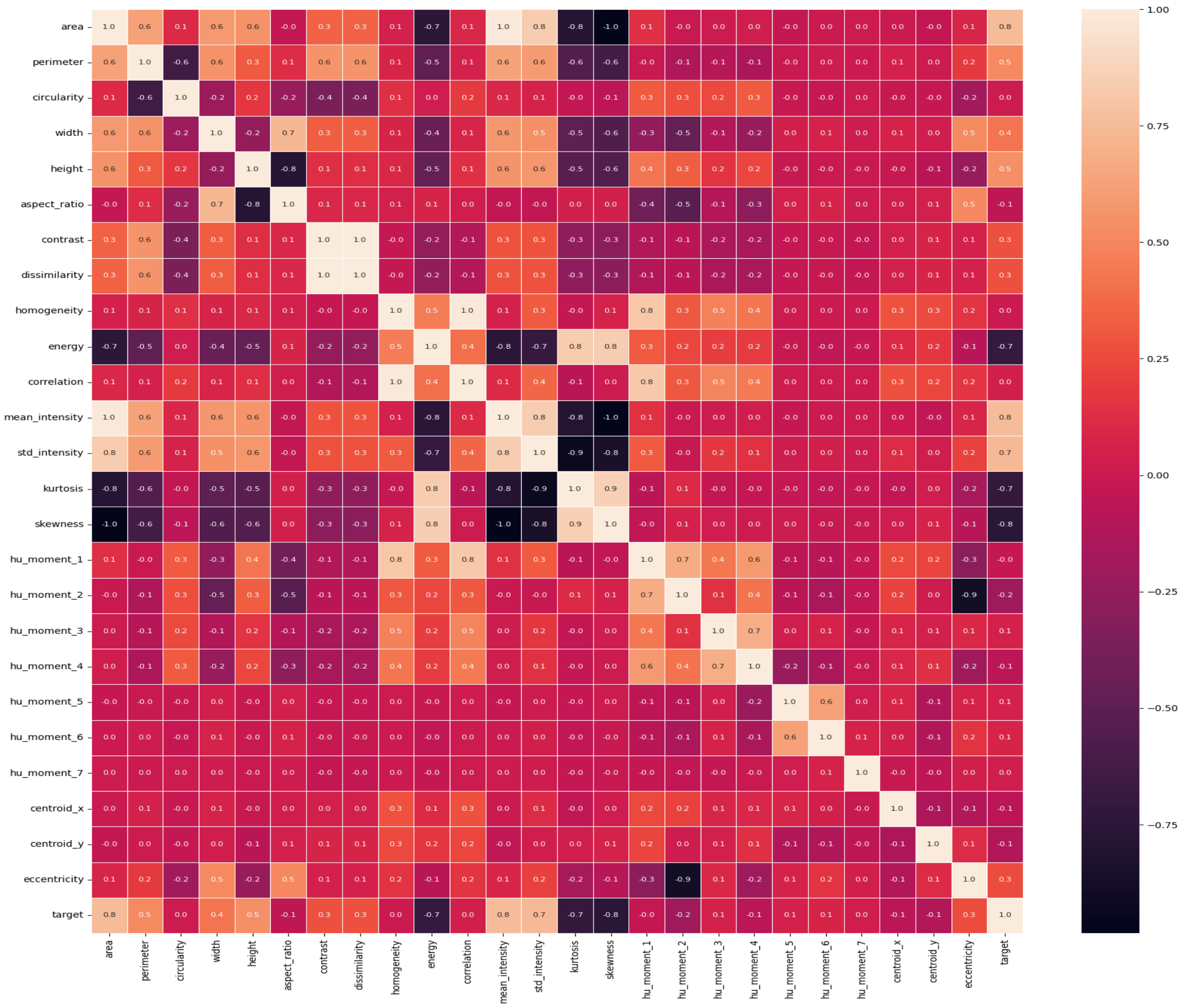

Figure 12: Correlation Map

The correlation map in Figure12 shows the strength and direction of relationships among variable best for the model such as pod length, width, weight, aspect ratio, dissimilarity, height, circularity, area, and perimeter. This correlation is important in engineering application in identify key traits for improved groundnut varieties for optimized machine thresher design. The correlations map results were obtained to have a range value of results i.e area, perimeter, width, height, mean intensity, and standard intensity (0.5 to 1.0), this indicated that these features are very significant for differentiating pod size and shape characteristics.

4.3 Training

In this research about 360 samples of groundnut pods as dataset was used to train the machine learning model, which comprise of three varieties of groundnut pods such as Exdakar, Jarma, and Samnut26, with image counts of 116, 99, and 145 respectively. The dataset was divided into three distinct subsets as 70:15:15 ratios for training, validation, and testing purposes as indicated in table 4

Table 4: Data Training Set

Classification	Training data	Validation data	Testing data
Exdakar	82	17	17
Jarma	69	15	15
Samnut26	101	22	22

4.3.1 Training and Testing (Validation) of HGBC Model

The training and validation accuracy test trends over 50 epochs for a machine learning model of HGBC. The training accuracy (blue line) exhibits a sharp increase within the first few epochs, quickly risen at approximately 0.945, indicating that the model rapidly learns the training dataset. The test accuracy (orange line) also rises initially, stabilizing around 0.905, but remains consistently lower than the training accuracy. The presence of a gap between the two curves suggests a generalization gap, which implies that the model performs better on the training data

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

than on unseen test data. This divergence between training and testing accuracy is a key indicator of potential overfitting. The model, while achieving high accuracy on the data it was trained on, demonstrates a limited ability to generalize these learned patterns to new. The persistent gap between the training and testing accuracy, even after the training accuracy risen, signifies that the model might be memorizing the training data rather than learning generalizable features this explain better as shown in figure 13.

However, the training and validation loss test exhibit a sharp decline of loss progression over 50 epochs, indicating that the model is effectively learning to minimize the prediction error on both the training and unseen data. This rapid initial decrease suggests that the model's parameters are quickly adjusting to capture the underlying patterns in the data. However, after a few epochs, the training loss continues to decrease and stopped at a significantly lower value compared to the testing loss which stopped at the higher level i.e beyond this point, both curves stopped, suggesting convergence, with the training loss stabilizing around 0.20 and the test loss around 0.27, this divergence between the training and testing loss is a hallmark of overfitting. While the model becomes increasingly proficient at fitting the training data, achieving a low training loss, its ability to generalize to new, unseen data deteriorates, as evidenced by the higher and relatively stable testing loss as indicated in figure 14.

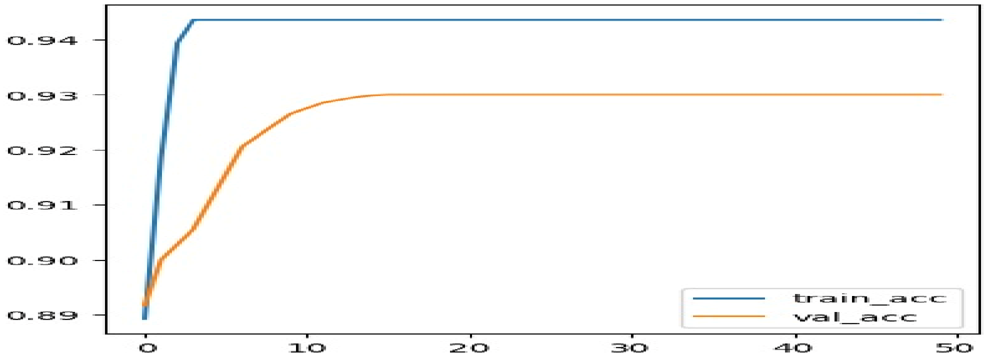

Figure 13: The Training and Validation Accuracy Test

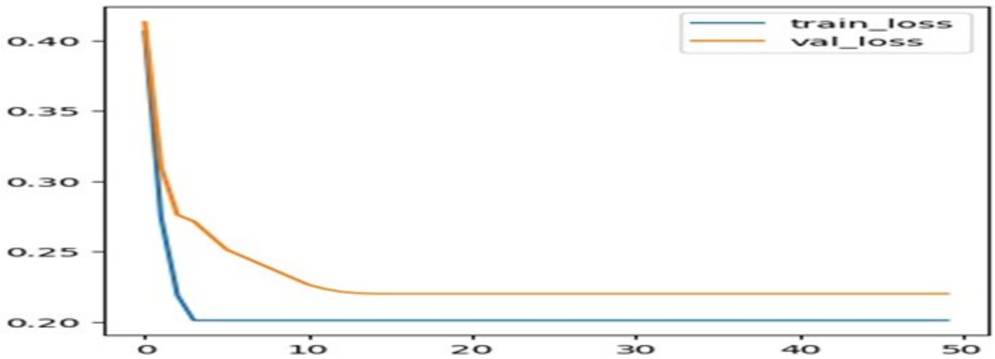

Figure14: The Training and Validation Loss Test

4.4 Performance Evaluation

$$.TP_{Exd} = 110,\ TP_{Jar} = 94, TP_{Sam} = 127,$$

$$Accuracy = \frac{TP_{Exd}+TP_{Jar}+TP_{Sam}}{Total\ Sample} = \frac{110 + 94 + 127}{110 + 6 + 6 + 94 + 1 + 12 + 5 + 5 + 127} = 0.919 \approx 91.9\%$$

Exdakar

$$.FP = 0 + 6 = 6,\ FN = 1 + 5 = 6$$

$$Precision = \frac{\text{TP}}{\text{TP+FP}} = \frac{110}{110+6} = 0.948 \approx 0.95 = 95\%$$

$$Recall = \frac{\text{TP}}{\text{TP+FN}} = \frac{110}{110+6} = 0.948 \approx 0.95 = 95\%$$

$$F1 = 2*\frac{\text{Precision} * Recall}{\text{Precision} + Recall} = 2*\frac{0.948 * 0.948}{0.948 + 0.948} = 0.948 \approx 0.95 = 95\%$$

Jarma

$$.FP = 1 + 12 = 13,\ FN = 0 + 5 = 5$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} = \frac{94}{94+13} = 0.878 \approx 0.88 = 88\%$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} = \frac{94}{94+5} = 0.949 \approx 0.95 = 95\%$$

$$F1 = 2*\frac{\mathrm{Precision} * Recall}{\mathrm{Precision} + Recall} = 2*\frac{0.88 * 0.95}{0.88 + 0.95} = 0.913 \approx 0.91 = 91\%$$

Samnut26

$.FP = 5 + 5 = 10,\ FN = 6 + 12 = 18$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} = \frac{127}{127+10} = 0.927 \approx 0.93 = 93\%$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} = \frac{127}{127+18} = 0.876 \approx 0.88 = 88\%$$

$$F1\text{- }Score = 2*\frac{\mathrm{Precision} * Recall}{\mathrm{Precision} + Recall} = 2*\frac{0.93 * 0.88}{0.93 + 0.88} = 0.901 \approx 0.90 = 90\%$$

The performance matrix of the HGBC model across all the varieties were summarize in table 5

Table 5: Performance Machine Learning Model of HGBC

Models	**Groundnut Pod Varieties**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
HGBC	Exdakar		95	95	95
	Jarma	91.9	88	95	91
	Samnut26		93	88	90

Other model performance are presented in table 6 which was adopted from [20]

Table 6: Performance of other Models

Models	**Groundnut Pod Varieties**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Svm	Chandra	86.07	86.4	84.2	87
Rf	Chandra	82.27	82	82	82
Mlp	Chandra	84.9	84	84	84

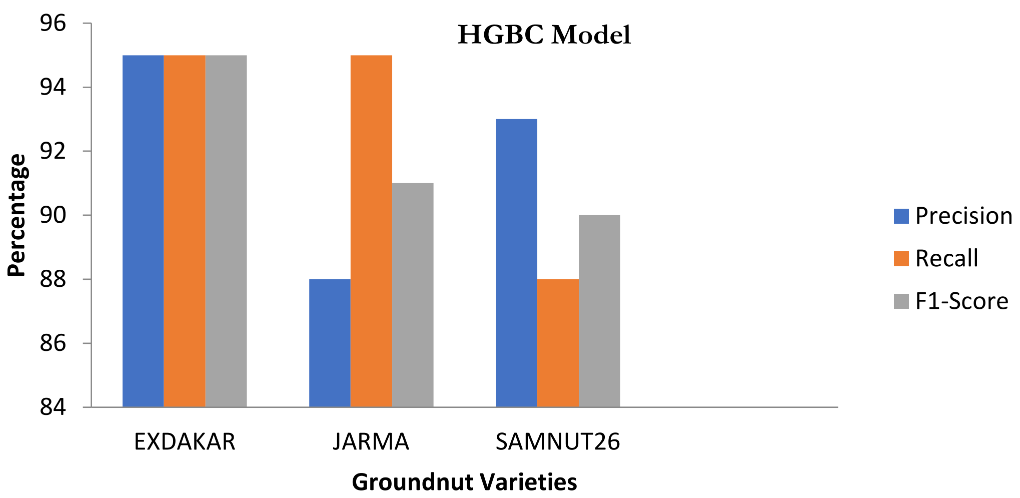

Figure 15: Performance of HGBC Model on Groundnut Varieties

The performance evaluation of the HGBC classifier for three groundnut pods varieties Ex-dakar, Jarma, and Samnut26 using Precision, Recall, and F1-Score metrics as shown in Figure 15. Ex-dakar achieved the best performance with both 95% in precision, recall, and F1-Score, indicating accurate prediction. Jarma recorded it best result as recall with value 95% this means that the model identified its correctly, precision was lower 88% due to higher false positive value, and the resulting F1-Score of 91% while Samnut26 achieved its highest results as precision with 93%, indicating fewer false positive predictions, but its recall 88%, with the F1-Score of 90%. Therefore, the HGBC model demonstrated strong classification capability across all groundnut varieties, but Ex-dakar performed the best.

4.4.1 Models performance comparison

The models accuracy comparison results shown in Figure 16 indicate that the HGBC model achieved the highest classification accuracy of 91% performance, which outstand all other machine learning models. However Support Vector Machine (SVM) followed with an accuracy of 86%, demonstrating good classification performance. Multilayer Perceptron (MLP) achieved 85% accuracy, this performance is slightly lower. Random Forest (RF) recorded the lowest accuracy of approximately 82%, with the lowest classification performance. The results indicate that the HGBC model provided the most reliable and accurate prediction performance among the evaluated classifiers models.

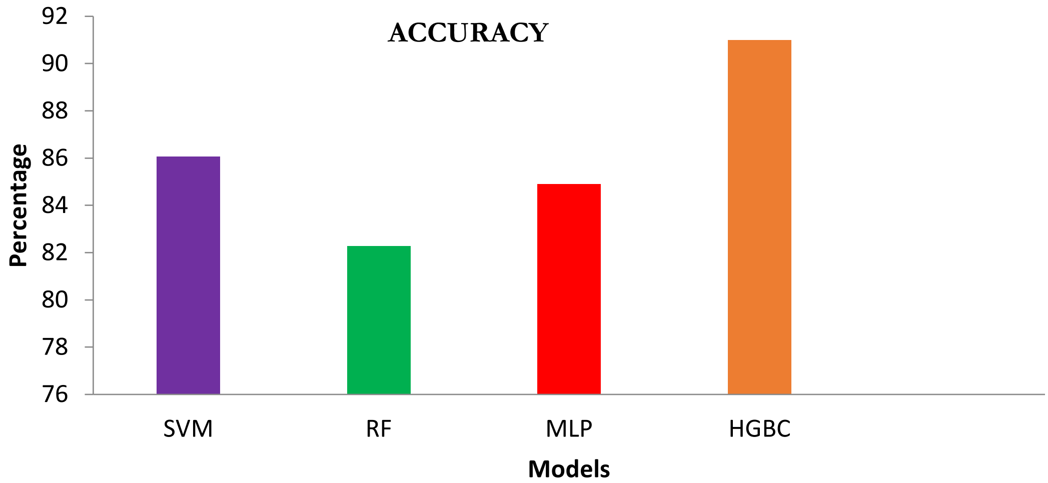

Figure 16: Comparison of Accuracy between HGBC Model and other Model

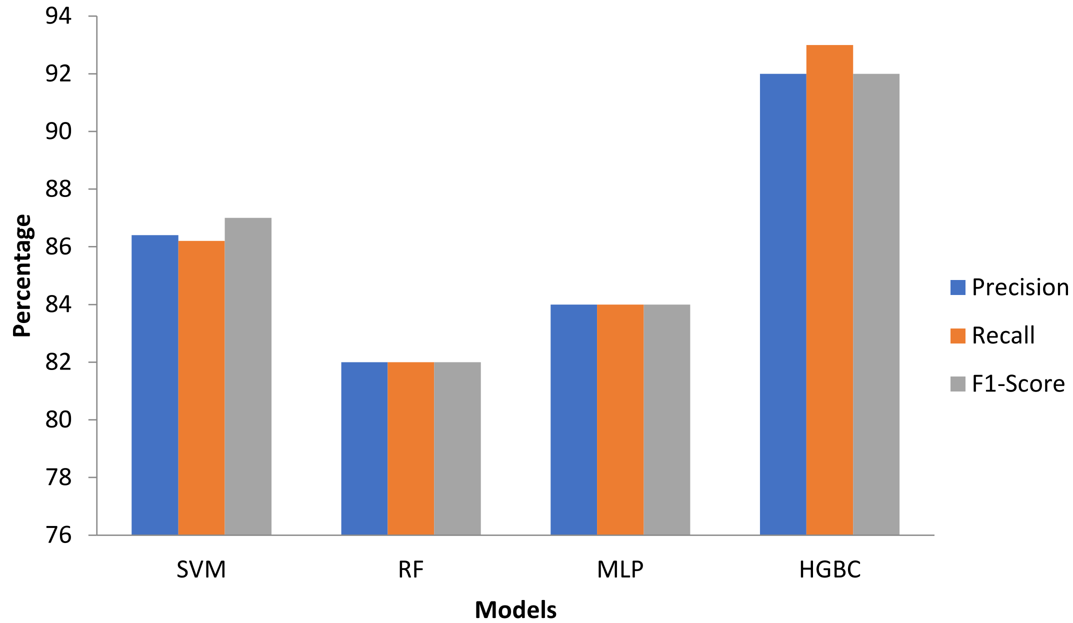

Figure 17: Performance Comparison between HGBC Model and other Model

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

The comparison performance of the models shown in Figure 17 indicate that the HGBC model achieved the best performance with the highest precision 92%, recall 93%, and F1-score 92%, indicating superior classification accuracy and consistency for groundnut pod identification. SVM demonstrated moderate performance with values range between 86 to 87%, showing good predictions result. MLP achieved slightly lower performance with uniform values of 84% across all evaluation metrics, while RF recorded the lowest performance with 82% precision, recall, and F1-score. Furthermore, HGBC outperformed the other entire machine learning models.

4.5 Conclusion

The developed HGBC was effective in classifying three groundnut pod varieties Exdakar, Jarma, and Samnut26 with an overall classification accuracy of 91.9**%**. The model demonstrated strong performance, where Exdakar achieved 95% precision, 95% recall, and 95**%** F1-Score, Jarma recorded 88**%** precision, 95**%** recall, and 91% F1-Score, while Samnut26 had 93% precision, 88% recall, and 90% F1-Score. Using a dataset of 360 groundnut pod samples, the extracted geometric and texture features proved the effectiveness of different varieties. These findings confirmed that machine vision combined with HGBC can automate groundnut pod classification. The HGBC model provided the most reliable and accurate prediction performance among the evaluated classifiers models such as SVM, RF, and MLP and has a good precision, recall and F1-Score. Therefore, the research established the classification system for an intelligent groundnut threshing machine that effectively reduced seed damage.

References

[1] B. Bello, A. Tokan, and D. J. Jiya, “Design and fabrication of an automated groundnut threshing machine,” *International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)*, *Universal Scientific Organization, 2019*.

[2] J. C. Lokossou, H. D. Affognon, A. Singbo, M. B. Vabi, A. Ogunbayo, P. Tanzubil, A. C. Segnon, G. Muricho, H. Desmae, and H. Ajeigbe, “Welfare impacts of improved groundnut varieties adoption and food security implications in the semiarid areas of West Africa,” *2019*.

[3] B. Bello, A. Tokan, D. Jiya, and A. Musa, “Design model of automated groundnut threshing machine,” *Majlesi Journal of Mechatronic Systems*, *vol. 8, no. 1, pp. 19–24, 2021*.

[4] M. Sabry, M. Osman, A. Hussein, M. W. Mehrez, S. Jeon, and W. Melek, “A generic image processing pipeline for enhancing accuracy and robustness of visual odometry,” *2022. DOI: 10.3390/s22228967*.

[5] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, *pp. 785–794*.

[6] J. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, *vol. 147, pp. 70–90, 2018*.

[7] S. Kollmannberger, “Fundamental concept of machine learning deep learning in computational mechanics,” *Studies in Computational Intelligence*, *vol. 977, 2021. DOI: 10.1007/978-3-030-76587-3_2*.

[8] M. Z. M. Sabri, N. A. A. Majid, S. A. Hanawi, N. I. M. Talib, and A. I. A. Yatim, “Prediction model based on continuous data for student performance using principal component analysis and support vector machine,” *TEM Journal*, *vol. 12, no. 2, pp. 1201–1210, 2023, DOI: 10.18421/TEM122-66*.

[9] L. Prokhorenkova, G. Gusev, A. Vorobev, A. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” in *Advances in Neural Information Processing Systems*, *2018*.

[10] H. Durmuş, E. Güneş, and M. Kurci, “Disease detection on the leaves of the tomato plants by using deep learning,” in *6th International Conference on Agro-Geoinformatics*, 2017.

[11] A. Taner, M. T. Mengstu, K. C. Selvi, H. Duran, İ. Gür, and N. Ungureanu, “Apple varieties classification using deep features and machine learning,” 2024. *DOI: 10.3390/agriculture14020252*.

[12] P. K. Sethy and B. Behera, “Detection of coronavirus disease using deep features,” *Preprints*, 2020.

[13] A. Camargo and J. Smith, “Image pattern classification for the identification of disease causing agents in plants,” *Computers and Electronics in Agriculture*, vol. 66, no. 2*, pp. 121–125, 2009*.

[14] A. Wang, W. Zhang, and X. Wei, “A review on weed detection using ground-based machine vision and image processing techniques,” *Computers and Electronics in Agriculture*, *vol. 158, pp. 226–240, 2019. DOI: 10.1016/j.compag.2019.02.005*.

[15] K. Verma, R. K. Kathiria, T. D. Mehta, P. M. Chauhan, S. Yadav, and A. L. Lakhani, “Investigation of engineering properties of groundnut pods for advanced pneumatic pod collection systems,” *Plant Archives*, *2024. DOI: 10.51470*.

[16] D. Salunke, V. Shivanag, P. Sabne, P. Jadhav, and H. Saini, “Handwritten Devanagari word recognition using customized convolution neural network,” *2023*.

[17] V. G. Narendra and A. S. Kini, “An intelligent classification model for peanut varieties by color and texture features,” *International Journal of Engineering & Technology*, *vol. 7, no. 2.27, pp. 250–254, 2018*.

[18] Y. Wang, F. Ye, J. Zeng, J. Cai, and W. Huang, “Adherent peanut image segmentation based on multi-modal fusion,” *Sensors*, *vol. 24, no. 14, p. 4434, 2024. DOI: 10.3390/s24144434*.

Received: 05/04/2026; Revised: 23/05/226; Accepted: 27/05/2026; Published: 01/06/2026

[19] Y. Zhang, L. Wu, and S. Wang, "Feature extraction and classification for image-based plant recognition using machine learning techniques," *Computers and Electronics in Agriculture*, *vol. 162, pp. 112–121, 2019. DOI: 10.1016/j.compag.2019.04.018*.

[20] A. Taner, M. T. Mengstu, K. Ç. Selvi, H. Duran, İ. Gür, and N. Ungureanu, "Apple varieties classification using deep features and machine learning," *Agriculture, vol. 14, no. 2, 2024. DOI: 10.3390/agriculture14020252*.

[21] V. G. Narendra and A. S. Kini, "An intelligent classification model for peanut varieties by color and texture features," *International Journal of Engineering and Technology*, *vol. 7, no. 2.27, pp. 250–254, 2018*.