VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2026, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 14, Number 3, July-September 2026 pp:89-108

HCER-ResNet50: A Deep Learning Framework for Accurate Cashew Nut Classification and Grading

S. Muthukumaran[1*], **B. Kamatchi**[1], **P. Arivazhagan**[1], **N. Kalaichelvi**[1]
R. Mahalakshmi[2]

[1]Assistant Professor, Department of Advanced Computing and Analytics, VELS Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai 600117, Tamil Nadu, India; [2]Associate Professor, Department of Advanced Computing and Analytics, VELS Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai 600117, Tamil Nadu, India

Keywords: *Cashewnuts Grading, Deep Learning, Image Compression, Image Enhancement, Convolutional Neural Network, Adam Optimizer*

Journal Info:
Submitted:
June 1, 2026
Accepted:
July 06, 2026
Published:
July 16, 2026

Abstract

Background: Cashew nuts are healthy and are commonly found in food products, confectionery, and foods, whereas cashew apples, shell derivatives are used in beverages and other industrial products. The rising commercial value of cashew has augmented the requirement of the effective and consistent grading mechanisms in order to safeguard the quality and market value. **Methodology:** In this paper, an automated cashew nut-grading system, Hybrid Compression and Enhancement based ResNet-50 (HCER-ResNet50), is suggested to overcome the weaknesses of manual and traditional methods of using machine-vision to grade cashew nuts. The structure combines the image compression, image enhancement and deep learning classification to produce reliable and high-quality grading across different imaging conditions. Noise removal, normalization and resizing prior to processing follow a pattern of consistent image quality, a hybrid compression scheme based on DWT, DFT, and DCT combined with other enhancement tools such as thresholding and contrast stretching enhances the clarity and separability of features. Based on this, a convolutional neural network with ResNet-50 is used to categorize cashew nuts into specific grades with a minimum human intervention. **Results:** In all image processing procedures, the classification accuracy of the HCER-ResNet50 was between 0.9598-0.9720, as the experimental data showed. The PNG Dataset achieve the best accuracy (0.9720) and this goes to show that both the transform and the wavelet-based pre-processing methods improve the performance of the grading. Most of the pre-processing methods including thresholding, DWT approximation, DCT and DFT, contrast enhancement, and DWT horizontal methods used approximately 18 to 22 minutes to train and categorize the model with a steady and computationally efficient performance. Furthermore, the Adam optimizer is used to enhance model parameters in the training stage and contributing to enhance the classification outcome. The proposed HCER-ResNet50 algorithm by accurately grading the Cashewnuts improves the revenue of the farmers and reduce post-harvest losses.

***Correspondence author email address:** muthumphil11@gmail.com
DOI: 10.21015/vtse.v14i3.2476

1 Introduction

Cashew Anacardium occidentale Anacardiaceae is a tropical perennial tree, indigenous to South America, and generally cultivated in steamy areas. It yields two economically valuable crops the cashew nut (a kidney-shaped seed, typically used as food or made into a variety of products such as butter and cheese) [4] and the cashew apple (an accessory fruit, used in beverages and fermented beverages). The tree usually reaches a height of 14 m, dwarf varieties grow to about 6 m, mature quicker, and yield better, which is more lucrative to have them in commercial cultivation. In 2023, Ivory Coast and India were the world front-runners of about 3.9 million tonnes of cashew production [5].

In the past, the crop found its way into other parts of the world due to the Portuguese trading routes linking Brazil with other places around the world like the Goa and the Southeast Asian and African regions. Cashew apple has a traditional medical application, such as in the treatment of dysentery and sore throat in Cuba and Brazil, and it is a possible food functional ingredient, as it contains dietary fiber, bioactive compounds, antioxidants and high levels of vitamin C. The juice contains high amounts of sugars, antioxidants, vitamin C, and has a large amount of essential minerals, including potassium, phosphorus, magnesium, calcium, iron, sodium, copper, and zinc. Potassium is important to the body because it is associated with the bone, whereas copper, sodium, zinc, and magnesium are important in the metabolism, enzyme activity, osmoregulation, and cell functions. Although cashew apple has a high nutritional value, that is, it comprises vitamins, minerals, and bioactive mixtures, it is not comparatively accepted as a food because of high levels of tannin and astringent taste. However, it has high prospects of value addition in functional food, nutraceuticals as well as child nutrition products particularly when it is used as a natural source of vitamins and minerals. It is seen that there are differences in nutrient and mineral compositions based on the location, and this means that cashew grown in southern India and analysed in terms of bioactive and nutritional properties [3]. Furthermore, cashew nut shell derivatives have significant industrial applications, including biochar production for soil amendment and other value-added products [1].

1.1 Challenges faced by Cashewnuts Cultivators in India

India The low productivity of cashew farmers is because the plantations based on seedlings are old, and the soil is not fertile enough and because they have not embraced the suggested cultivation methods [2]. Despite the availability of numerous processing units, the domestic production of raw nuts is not adequate and this has led to imports; supply shortage also occurs when the large processing nations such as Vietnam, Cambodia, import more nuts. The climate risks faced by farmers, attacks by pests and diseases, high input and labour expenses, fluctuation in price, and reliance on intermediaries are also experienced. Other limitations are the low levels of technical awareness, financial assistance and farming in marginalized land all of which decrease the yield and profitability.

1.2 Objective of this Research

Computer vision is a subspecialty of artificial intelligence that sanctions machines to observe painterly information in pictures and videos and use machine learning to means and deduce these images. It operates in three fundamental processes, namely recognition, which determines objects and patterns; reconstruction, which determines structural or spatial features; and reorganization, which determines relationship between visual elements. A standard computer vision system consists of data acquisition, pre-processing, model choice and model training. Images are initially collected and labelled and are further refined by using pre-processing techniques like resizing, contrast boost, and data augmentation to guarantee superior model execution. Computer vision is analogous to the human vision. The way the human eye takes images and sends them to the brain where it interprets the information and identifies objects, similarly, the computer vision systems consist of the camera that takes visual data and algorithms that determine the patterns and identify the objects. Nonetheless, these systems need to be trained on massive sets of labelled images prior to the identification of new objects in order to acquire relationships between patterns and classes.

1.3 Research Contribution

The research is a proposal of automated cashew nut grading system, which is called Hybrid Compression and Enhancement based ResNet-50 (HCER-ResNet50) of Automated Cashew Grading, to overcome the shortcomings of manual grading and current machine-vision schemes. The framework uses image compression, image enhancement, and a classification that uses deep learning to provide accurate and consistent grading when different imaging conditions apply. Pictures that are taken by an acquisition module are subjected to pre-processing processes of image noise, image normalization, and image size reduction to standardize the quality of inputs. A hybrid compression stage with Discrete Wavelet Transform (DWT), Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT) abolishes jobless information and keeps important spatial and frequency features. Thresholds, contrast stretching, and homomorphic filtering are additional methods to use to sharpen the images and increase the separation of features. The trained Convolutional Neural Network is based on the ResNet-50 architecture using the processed images to automatically classify cashew nuts to the respective grades. Finally, the Adam optimizer is used to enhance the stability and convergence of the ResNet-50 model.

2 Literature Review

MD Tausif Mallick et al. [6] introduced an approach to disease and pest classification in mustard and mung bean plants based on a deep learning model that utilizes transfer learning and a MobileNetV3 CNN that is deployed to a System-on-Chip (SoC) platform. The research created a homemade RGB dataset that was gathered in a period of one year in farms, college fields, green houses, and laboratories under various environmental conditions, and images were manually purged, labelled and resized to train the model. There were several plant parts in the dataset and nine and eleven classes of mustard and mung bean respectively. The authors checked the comparison of MobileNetV3 to the InceptionV3 and VGG16/VGG19 when the experimental conditions were the same (they used 80:10:10 train-validation-test split and data augmentation). A custom classification head and dropout to improve regularisation were applied to the feature extractors, which were pre-trained and then trained by a two-stage training strategy consisting of freezing and fine-tuning convolutional layers. The experimental findings indicated that the quality of results of all models was high but MobileNetV3 had more stable learning, quicker convergence, and also a much lower number of parameters. It delivered classification precision of 96.14% in mung bean and 93.25% in mustard which is a little better than rival models.

Farhadi et al., [7] investigated to sort and grade hazelnuts by incorporating audio signal processing and ANN approaches. A method was deployed in the produced sound, owing to the impact of the hazelnut with a steel disk, led by the microphone located under the steel disk and relocated to a PC through a sound card. Afterwards, it was kept and processed by a program recorded in MATLAB software. Pham et al., [8] developed a module named SC3T that can be applied to incorporate into the backbone of the YOLOv8 framework. In the SC3T module, a Transformer block is efficiently combined together with the C3TR module. Most significantly, the classification is not only effectual but also compacted, which can be employed in an embedded tool of deployed cashew nut grading networks.

Jingwen Zheng et al. [9] examined the use of MobileNetV2 to classify the rice leaf disease in its optimized edge deployment form. The data were gathered by mobile phones and drones under conditions of different lightings, distances, and viewing angles and in eight disease characteristics with different degrees of severity. More open-source images were added to enhance background diversity and decrease domain bias whereas class imbalance was addressed through oversampling, class-balanced loss, and data augmentation. Several CNN models (MobileNetV2, AlexNet, ResNet18/50, and VGG16) were tested with 16-bit quantization to be run on an FPGA. The confusion matrix indicated steady classification performance with the best classified visual features being neck blast with its high accuracy of 97.43% and bacterial leaf blight with low accuracy of 91.89% as it had similar symptoms and different light conditions. There were misclassifications between the visually similar diseases like untrue smut, rice leaf binder and rice stem borer. The findings of grad-CAM analysis proved that the similarity in streak, lesions, and panicle damage resulted in classification errors. The present study has shown that model performance is influenced by the ambiguous symptoms and environmental differences and that more diverse and clearer dataset is required. Multiple platform experiments with images generated better accuracy and generalization, especially with neck blast. The inclusion of mobile captured images made the model robust and stable to facilitate accurate classification of rice disease in the dense environment.

Chen Jin et al. [10] employed near-infrared hyperspectral imaging and deep learning to identify six types of rice seeds with the help of 1D spectra, 2D pseudo-RGB images, and 3D hyperspectral data. Controlled conditions of capturing images, calibration, removal of noisy bands, and preprocessing measures, including smoothing and SNV normalization were done and the dataset was riven into training, validation, and testing sets. Conventional algorithms such as SVM and Logistic Regression used one dimensional spectral data whereas deep learning networks used one dimensional, two dimensional, and three dimensional data which underwent classification. In order to get round the problem of single-source data constraint, multi-dimensional fusion models (MDFF-CNN and MDDF-CNN) were trained to combine spectral and spatial information by use of feature or decision fusion and then Softmax classification. Experimental findings indicated that SVM surpassed the use of the Logistic Regression in the case of spectral data whereas fusion based CNN models exhibited better accuracy than their single dimensional counterparts. It was noted in the study that feature representation, robustness and overall classification accuracy are enhanced through the use of multi-source data fusion applying spectral characteristics (along with spatial information).

Lei Zhang et al. [11] suggest the multi-courtesy improved 3D Residual Convolutional Neural Network (3D-ResCNN) to primary detect mildew in tobacco shrubberies by depleting the hyperspectral imaging. The tobacco samples were taken in various areas and kept in controlled temperature and humidity to cause mildew and hyperspectral images (427.401069.75 nm) were obtained with black-white calibration to minimize noise. A sliding window was applied to extract spatial spectral patches to create a dataset of 1360 hyperspectral images having 250 spectral bands to train the model. The design incorporates the concept of residual learning in a 3D-CNN structure to enhance feature learning and stabilize deep network learning. It consists of a preliminary three-dimensional convolutional layer and a series of residual units with ReLU activation and batch normalization to study nonlinear spatial spectral representations and speed up the learning process. The extracted features are processed by the HCER-ResNet50 architecture, where deep residual feature learning is performed to classify the images into severe mildew, mild mildew, non-mildew, and background categories. In the experiments, it was revealed that the combination of residual connections with channel and spatial spectral attention mechanisms enhanced the classification performance considerably. The proposed model was more accurate (OA 99.12, AA 99.09, K 98.82) than outmoded machine learning models (SVM, RF), and baseline deep learning models (1D-CNN and 3D-CNN), and the visualization and the confusion matrix analysis showed that the proposed model produced fewer misclassification.

The study by Soo Jun Wei et al. [12] tested the pre-trained CNN models, i.e., AlexNet, VGG16, ResNet50, and DenseNet121 with the PlantVillage dataset, which has 55,446 images. Under-sampling and intensification were castoff to contract with the class unevenness, and the data was divided into training, validation, and testing groups. A transfer learning model was used to create a lightweight model that could be deployed as an edge. DenseNet121 was the most successful model and was chosen to be implemented. The optimized model was translated into ONNX and implemented on CPU and VPU hardware by OpenVINO, and the images were rescaled and normalized to ensure that they were inferred in a consistent manner. The classification layer was re-implemented using ReLU, Dropout and Softmax, and layer features extraction was frozen throughout the training. It was experimentally proven that DenseNet121 exhibited the highest accuracy of 96.4% and maintained similar performance in CPU and VPU settings (about 95.7%), which indicates that it is an effective and favourable component when it comes to using it in edge-based plant disease detection applications.

Momeny et al., [14] introduce a Convolutional Neural Network (CNN) system to identify the presence of cherries and offer an effective method for their grading. For the purpose of detecting cherry images on two categories (regular and irregular shaped) was organized. Later preprocessing the images, the proposed technique used its capability for improving generalization in the CNN via an incorporation of average pooling and max pooling system, to evaluate cherries. Abbaszadeh et al.,

[15] demonstrates an unsupervised learning methodology to categorize damaged and unpleasant pistachios depends on deep Auto-encoder NN.

Furthermore, [13] presented a comparative analysis of VGG-16, ResNet50, and EfficientNet-B1 with optimization techniques for crop disease detection, demonstrating the effectiveness of ResNet50 in agricultural applications. [22] evaluated SE, ECA, and CA modules in ResNet-50 for grape leaf disease classification, further validating the adaptability of ResNet-50 for plant disease detection. [23] proposed a deep learning-based approach for classifying the maturity of cashew apples, which is directly relevant to the cashew domain.

3 Background Knowledge

A digital picture is a two-dimensional function f(x,y) where spatial coordinates represent location and intensity denotes gray level; when these values are discrete, the image is digital. Digital image processing involves manipulating such images using computers, with pixels as the basic elements. Imaging extends beyond human vision and covers the entire electromagnetic spectrum, including sources like ultrasound and electron microscopy. The boundaries between image processing, image analysis, and computer vision are not clearly defined but form a continuum [16]. Low-level processing focuses on pre-processing tasks, mid-level processing handles segmentation and recognition, and high-level processing interprets visual information. Overall, digital image processing includes both image-to-image operations and extraction of meaningful attributes for object recognition and applications.

3.1 Image Segmentation using Thresholding Operations

Image thresholding is an easy and effective process that can be applied in image segmentation to separate the objects and the background by the intensity value of the pixels [17]. In global, local and adaptive thresholding, a pixel value T is used to create the foreground and background pixels. In case of complex images, one can use several thresholds to divide the different object classes but this may not be easy. The performance of thresholding is contingent on the properties of histograms, the level of noise, and the uniformity of illumination and the comparative size of objects and ground, and reflectance. The equation for transforming a grayscale image $f(x,y)$ into binary image $g(x,y)$ with multi class thresholding is to have:

$$g(x,y) = \begin{cases} a, & \text{if } f(x,y) \leq T_1 \\ b, & \text{if } T_1 < f(x,y) \leq T_2 \\ c, & \text{if } f(x,y) > T_2 \end{cases} \quad (1)$$

Here $f(x,y)$ is the original image intensity at pixel (x,y), T is the verge value, $g(x,y)$ is the segmented output image.

3.2 Contrast Stretching

Poor illumination, insufficient sensor dynamic range or inappropriate lens parameter settings used when capturing images may result in low-contrast images [18]. Such images can be improved through contrast stretching which increases the range of intensity of the image to the maximum dynamic range of the display or recording medium. The degrees of mapping transformation are two points within the mapping curve that can be used to control the mapping operation and contrast strength. In case of equal transformation points, then the transformation is a linear mapping producing no change in intensity, but certain parameter choices can turn the transformation into some thresholding operation to create a binary image. The values of the intermediate parameters give different contrast enhancement levels without altering the order of intensity with the help of a monotonic transformation. The lowest and highest intensity values are usually stretched in a linear manner to the entire range $[0, L-1]$, enhancing the quality of the visual. In other situations, an average intensity may serve as a target to subdivide the image following the enhancement of contrast. The mathematical transformation used for contrast stretching is given by:

$$s = \begin{cases} \frac{s_1}{r_1} r, & 0 \leq r \leq r_1 \\ \frac{(s_2 - s_1)}{(r_2 - r_1)}(r - r_1) + s_1, & r_1 < r < r_2 \\ \frac{(L-1-s_2)}{(L-1-r_2)}(r - r_2) + s_2, & r_2 < r \leq L-1 \end{cases} \quad (2)$$

Where, r is the input pixel strength, s is the output pixel intensity, r_1, r_2 are the input intensity control units, s_1, s_2 were the output intensity control points, and L was the total number of intensity level.

3.3 Discrete Fourier Transform (DFT)

The discrete Fourier transform (DFT) is a variation of the continuous Fourier transform, based upon sampled bandlimited signals and the use of a finite number of uniformly spaced samples [19]. The DFT is a depiction of such trails in the frequency domain which fallouts in a collection of multipart coefficients that depict the frequency content of the signal. Based on the inverse DFT (IDFT) this allows the original discrete signal to be fully reconstructed and thus it is confirmed that the two are a transform pair that can be applied to all finite discrete data sets. The DFT, in one dimension, alters samples in the longitudinal domain, $f(x)$, into samples in the frequency domain, $F(u)$, and the IDFT rebuilds the original data with no loss of information. Both transforms are also periodic with period M and their properties enable the use of discrete signals to be analysed effectively. The DFT also presents the concept of circular convolution, in which the convolution of periodic discrete signals is performed over a fixed interval and it is the convolution-theorem, that is, product in the frequency domain is correlated with convolution in the space domain. These are the properties that render the DFT one of the essential instruments of digital signal and image processing. The equation for DFT is given by:

$$F(u) = \sum_{x=0}^{M-1} f(x)e^{-j\frac{2\pi}{M}ux}, \quad u = 0, 1, 2, \ldots, M-1 \tag{3}$$

The equation for IDFT is given by:

$$f(x) = \frac{1}{M}\sum_{u=0}^{M-1} F(u)e^{j\frac{2\pi}{M}ux}, \quad x = 0, 1, 2, \ldots, M-1 \tag{4}$$

3.4 Discrete Wavelet Transform (DWT)

Developed by Mallat (1987), the multiresolution theory offers a wavelet-based mechanism of analysing signals and images in multi-resolution form. It has been a combination of sub band coding, quadrature mirror filtering, and pyramid image processing [20]. In this method the scaling function synthesises the successive approximations of a signal at various resolutions and the wavelets synthesis captures the variation between adjacent resolutions. The discrete wavelet transform (DWT) is a representation of the signal as a linear sum of scaling and wavelet basis functions that can be orthonormal, orthonormal, and biorthogonal. In the 2-D Discrete Wavelet Transform (DWT), an image is divided into 4 sub bands at a given level i.e. approximation (LL), horizontal (LH), vertical (HL), and diagonal (HH). These are acquired by means of low-pass and the mathematical representation is given by:

Approximation (LL) is given by:

$$A(m,n) = \sum_x \sum_y f(x,y)g(2m-x)g(2n-y) \tag{5}$$

Horizontal detail (LH) is given by:

$$H(m,n) = \sum_x \sum_y f(x,y)g(2m-x)h(2n-y) \tag{6}$$

Vertical detail (HL) is given by:

$$V(m,n) = \sum_x \sum_y f(x,y)h(2m-x)g(2n-y) \tag{7}$$

Diagonal detail (HH) is given by:

$$D(m,n) = \sum_x \sum_y f(x,y)h(2m-x)h(2n-y) \tag{8}$$

Where, g is the low-pass scaling filter and h is the high-pass scaling filter.

3.5 Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) is a real orthogonal renovate with cosine tasks as its basis tasks whose frequencies are harmonically correlated with a supreme frequency demarcated by the length of the signal [21]. The DCT has an equal frequency range, but a better frequency resolution than the Fourier and Hartley transforms. The fact that the 2-D DCT is separable implies that the identical transformation matrix can be applied to one and two dimensional signals and images respectively. The DCT uses 2N-point periodicity and even symmetry unlike the DFT which uses N-point periodicity. This eliminates discontinuities in boundaries and features of high frequencies which are artificial and thus the DCT is especially effective in compressing images. The DFT of a symmetrically extended 2N-point signal will give the DCT of an N-point signal.

The equation for computing the cosine basis functions and weighting coefficients is given by:

$$C(u) = \sum_{x=0}^{N-1} f(x) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \tag{9}$$

The equation for computing the cosine basis for 2-D image is given by:

$$C(u,v) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \tag{10}$$

The DCT can also be imitative from the DFT of a proportionally prolonged signal, with weighting elements:

$$h(u) = e^{-j\pi u/(2N)}, \quad u = 0, 1, \ldots, N-1 \tag{11}$$

The final transformation of the input image was expressed mathematically by:

$$t_C = s \circ (h \circ t_1) \tag{12}$$

Here, the Hadamard element wise product is denoted by $\circ$.

3.6 ResNet50 based Architecture

ResNet-50 is a widely adopted deep residual network architecture designed to enable the effective training of very deep neural networks while mitigating the vanishing gradient problem through residual learning [22]. The core component of ResNet-50 is the residual block, which incorporates an identity shortcut connection that allows information to bypass one or more convolutional layers. This design facilitates efficient gradient propagation during training and enables the network to learn residual mappings instead of directly learning the desired transformations. For an input x, the output of a residual block is expressed as:

$$y = F(x) + x \tag{13}$$

where $F(x)$ represents the residual function learned by the convolutional layers within the block. This operation can be formulated as:

$$y = \sigma(W_2\sigma(W_1x) + x) \tag{14}$$

where W_1 and W_2 denote the convolutional weights, and $\sigma(\cdot)$ represents the activation function. The identity shortcut enables the network to learn the residual mapping:

$$F(x) = H(x) - x \tag{15}$$

rather than directly approximating the target mapping $H(x)$, thereby improving optimization and facilitating the training of deeper networks. [22] evaluated SE, ECA, and CA attention modules in ResNet-50 for grape leaf disease classification, demonstrating that attention mechanisms can further enhance ResNet-50's feature extraction capabilities in agricultural applications.

The ResNet-50 architecture begins with a 7×7 convolutional layer followed by a max-pooling layer. It then consists of four stages of residual learning, where each stage progressively increases the feature-map depth while reducing the spatial resolution. Each residual block adopts a bottleneck design comprising three convolutional layers arranged as 1×1, 3×3, and 1×1, together with a skip connection that adds the block input to its output. Overall, ResNet-50 contains 16 bottleneck residual blocks, along with the initial convolutional layer and the final fully connected classification layer, resulting in a total of 50 learnable layers.

3.7 Adam Optimizer

The Adam (Adaptive Moment Estimation) optimizer is utilized to update the ResNet-50 model parameters during the training period. It integrates momentum with adaptive learning rates to attain stable learning, improved convergence, and enhanced classification accuracy. The parameter update is computed as follows:

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{16}$$

Where θ_t presents the model parameters, α denotes the learning rate, $\hat{m}_t$ and $\hat{v}_t$ are the bias-corrected first and second moment estimates, and ε is a minor constant to avoid division by zero. Adam allows optimization training and improves the model performance of the proposed ResNet50 method.

4 Proposed Methodology

The data that was required in this study was gathered at Krishika Cashews located at Nadukuppam Village, Cuddalore District, Tamilnadu. The cashewnuts were picked and put in a white board and captured an image with

Redmi12 5G mobile in direct sunlight. The images taken were cropped and their background were removed and entire background were made transparent with white colour. There were two Cashewnuts datasets that were built, the first and the second datasets had 1200 and 7200 images respectively. The dataset comprised six distinct cashew nut categories: 240, Black Rotten, Raw Cashew, Split, Scorched Wholes (SW), and White Rotten. The datasets were divided into training and testing at the proportion of 70 to 30. The suggested Hybrid Compression and Enhancement based ResNet-50 of Automated Cashew Grading (HCER-ResNet50) Algorithm for Cashewnuts Classification shown in Figure 1 was based on an organized classification workflow. In the first step, raw images are gathered in the cashew nut database and obtained with the help of a camera module. Compression techniques such as Discrete Wavelet Transform (DWT), Discrete Fourier Transform (DFT), and Discrete Cosine Transform (DCT) were used for the images and each dataset is stored separately.

The algorithm then used image enhancement techniques such as Threshold Operation, Contrast Stretching were used and the enhanced images were stored separately as a database. The pre-processed image datasets were utilized to train the presented HCER-ResNet50 technique. Before being fed into the network, all images were resized to 180 × 180 × 3 to match the input requirements of ResNet-50. The model utilizes the deep residual learning architecture of ResNet-50, which consists of multiple residual blocks with skip connections that facilitate efficient feature extraction and improve gradient propagation during training. High-level features extracted by the backbone network are passed to the classification layers to predict the corresponding emotion class. To enhance model performance and convergence, the optimizer was applied during training, resulting in improved classification accuracy. Lastly, there is the output layer where the classification of cashew nut into six categories: 240, Black Rotten, Raw Cashew, Split, Scorched Wholes (SW), and White Rotten were done. Standard performance measures are used in the evaluation of the model performance and the findings facilitate the decision making in quality assessment and categorization of Cashewnuts. Lastly, the Adam optimizer is utilized to optimize the parameter values, securing effective training process and enhanced classification result. The proposed HCER-ResNet50 model offers an accurate and robust outcome for automated cashew nut grading.

5 Results and Discussion

The proposed Hybrid Compression and Enhancement based ResNet-50 of Automated Cashew Grading (HCER-ResNet50) algorithm was implemented using Python 3.8. The analysis and reconstruction of the image data were carried out in this HCER-ResNet50 algorithm with the help of a preprocessing pipeline based on the Discrete Cosine Transform (2D-DCT) algorithm (Python). The input images were loaded in a given directory, processed to grayscale, and processed to eliminate the alpha channels in case they occurred. Each of the images in the spatial domain was transformed to the frequency domain using a 2D-DCT, and then the inverse DCT (IDCT) was used to restore the images. Numerical comparison between the images that were reconstructed and the original ones were employed in the process of verifying the reconstruction accuracy.

Figure 2 shows samples of cashewnuts taken for classification across all six categories: 240, Black Rotten, Raw, Split, SW, and White Rotten.

Figure 3 shows the output of various image processing techniques applied to a 240 Cashewnut image, including (a) Original, (b) Contrast Stretching, (c) DCT, (d) DWT, (e) DWT Approx, (f) DWT Horizontal, (g) DWT Vertical, (h) DWT Diagonal, (i) DFT, and (j) Threshold.

The frequency domain image preprocessing technique of Discrete Fourier Transform (DFT) was carried out in Python. The methodology involved image loading, DFT/IDFT, numerical computation and mask generation through the use of OpenCV (cv2), numerical computation, and mask generation through the use of NumPy, and saving of processed images through Matplotlib, and directory and file management with the use of the OS library. The input images were initially processed to grayscale and changed to the frequency domain with the help of DFT spatial domain. The frequency spectrum shifted to the middle and a circular high-pass filter mask was used to cut low-frequency details and keep high-frequency details. The filtered frequency representation was again changed back to the spatial domain by

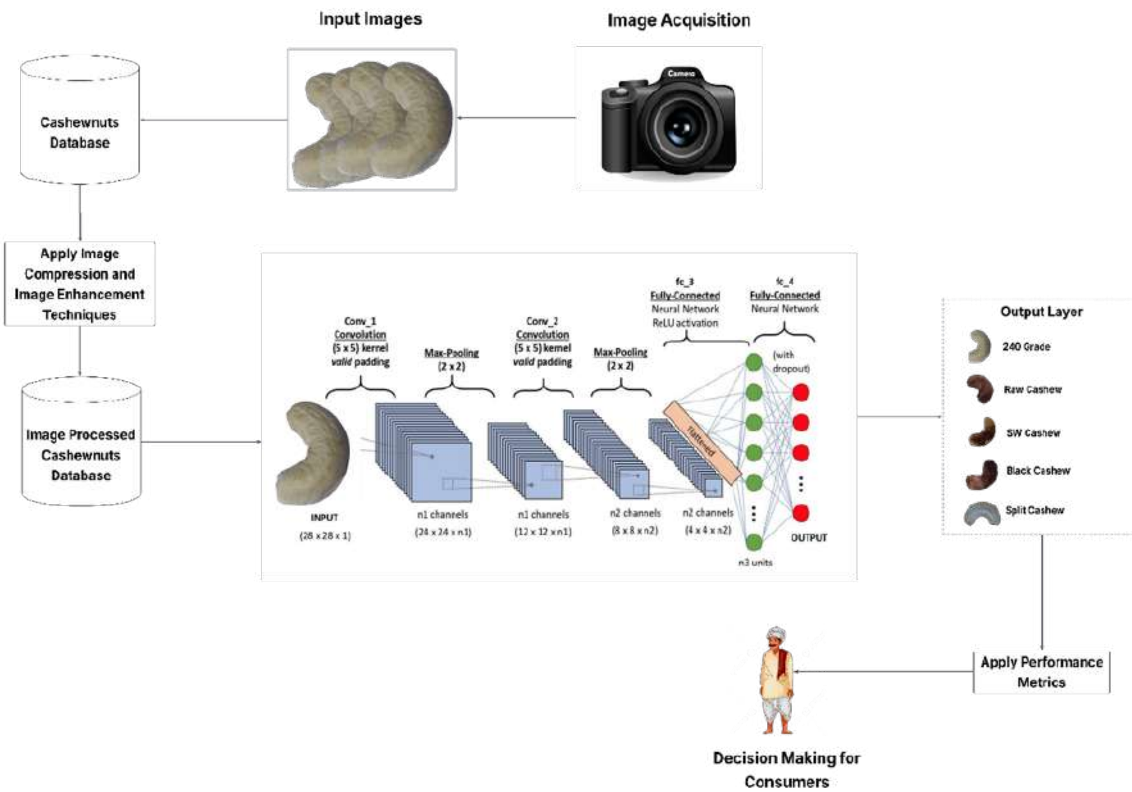

Figure 1. Proposed Hybrid Compression and Enhancement based ResNet-50 CNN of Automated Cashew Grading (HCER-ResNet50)

the inverse DFT to get improved images. The resulting images were stored in a given directory to be analyzed further. The given approach proves that Fourier-based preprocessing is effective in improving structural information and aiding feature description and analysis when applied in image analysis. This HCER-ResNet50 works with a multilevel Discrete Wavelet Transform (DWT)-type image preprocessing algorithm in Python. The method uses NumPy to execute arithmetic operations, PyWavelets (pywt) to do wavelet decomposition and reconstruction, Matplotlib to envisage, and scikit-image (skimage.io, skimage.color, skimage.exposure and skimage.transform) to read, convert colors and perform preprocessing of images. The dataset will contain the input image which is initially read and converted into grayscale to make the computation simpler. A two-level DWT is subsequently undertaken in the Daubechies wavelet (db5) with periodization mode in order to break down the picture into approximation and detailed frequency parts. Approximation coefficients describe the low frequency detail whereas horizontal, vertical, and diagonal detail coefficients describe the edge and texture details in different resolutions. Inverse wavelet transformation is then used to reconstruct the image to make sure that no information was lost. Lastly, the analyzing of the wavelet coefficients and the reconstructed image is displayed.

The Python Imaging Library (PIL) accessed through PIL.Image module allows the decomposition of images into separate RGB channels, pixel-wise point transformation to perform contrast stretching, and reuniting the fabricated bands to recover codified images. Application of HCER-ResNet50 uses various python libraries to do image preprocessing and automated thresholding. The computer vision framework selected as the main one is OpenCV library (cv2), which allows reading the image, turning it into a black and white and median filtering, as well as thresholding introduced by Otsu, which allows successfully eliminating noisy sources and segmenting an image into distinct parts. The os

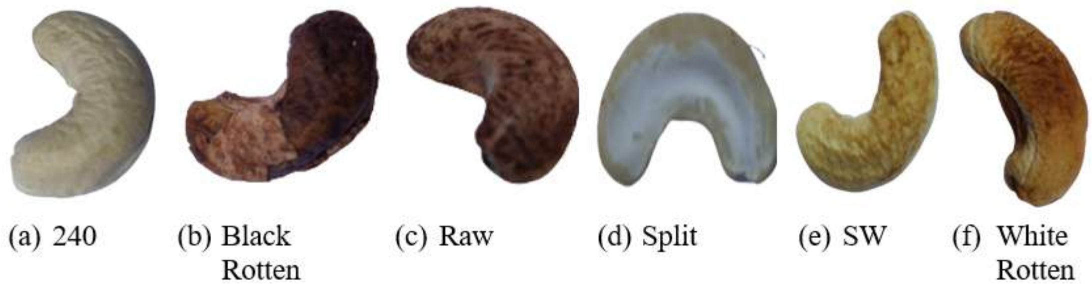

Figure 2. Samples of Cashewnuts taken for Classification

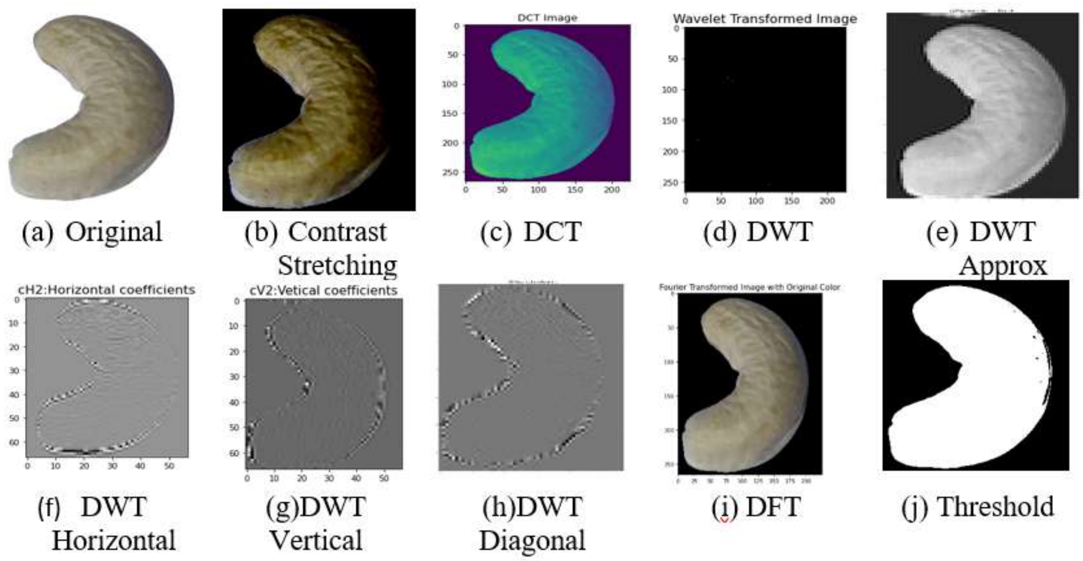

Figure 3. Output of the Various Image Processing Techniques if 240 Cashewnut Image

library also has file system functions like directory creation, path functions and batch iteration over image files. skimage.io sub-module of scikit-image package is added that enables image input/output functions to be performed in scientific image processing processes, and matplotlib.pyplot library is added because it may be needed to perform visualization and analysis of processed images in experimentation.

The proposed HCER-ResNet50 is a method, which is based on the use of TensorFlow and Keras to classify cashew images into six categories automatically. The ImageDataGenerator is used to perform image pre-processing and augmentation the training images are rescaled (1/255), have shear range of 0.2, zoom range of 0.2, and horizontal flipping, and the validation split is 0.1. The images are loaded in directory structures with target size of 180x180 and a batch of 30. Similarly, test data are augmented without rescaling. The HCER-ResNet50 architecture is constructed on a sequential model which has three convolutional blocks with a classification layer. The original convolutional layer entails of 64 3x3 kernel convolutional filters, a stride of 2, a same padding, a 3x3 kernel, and ReLU activation, and then a 2x2 max-pooling layer with a stride of 2. The second convolutional layer has 32 filters, the kernel size is 3x3, the stride is 2, the padding is same and the activation is ReLU, and another 2x2 max-pooling layer is introduced. And the third convolutional layer consists of 32 filters with a 3 x 3 kernel, stride 2, a same-padding, and ReLU activation, and a layer of max-pooling. The feature maps are flattened and the output layer is a fully connected dense layer with 6 neurons that represent the six classes and a softmax activation to perform multiple classes of classification. The compilation of the model is done with Adam optimizer, with categorical cross-entropy loss, and accuracy as the evaluation measure. The training is done in 100 epochs where it is validated on the test dataset. The performance of the model is measured by a confusion matrix and classification report and other measures of performance and given in Table 1.

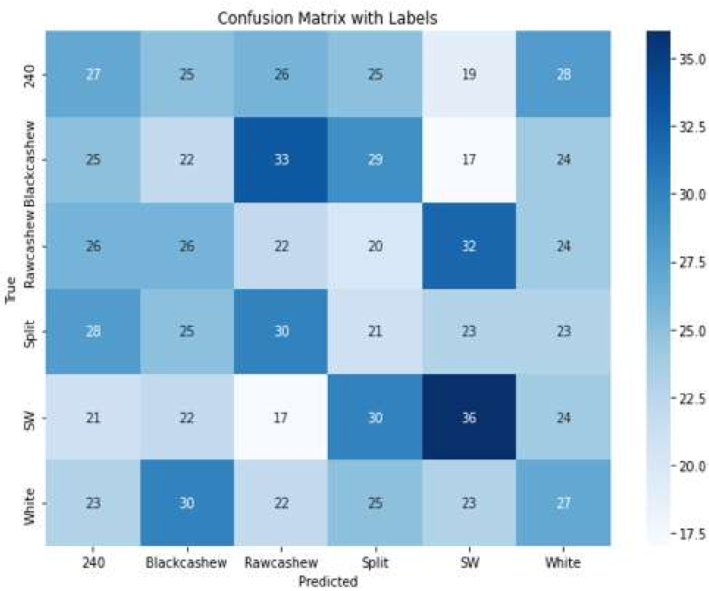

Figure 4. Confusion Matrix for the Original Cashewnuts dataset containing 1200 images

Figure 4 shows the confusion matrix for the original cashewnuts dataset containing 1200 images.

Figure 5 shows the confusion matrix for the DWT image dataset containing 1200 images.

Figure 6 presents the confusion matrix for thresholding operation dataset containing 1200 images.

Figure 7 presents the confusion matrix for contrast stretching dataset containing 1200 images.

The proposed HCER-ResNet50 was tested with the help of various preprocessing methods (applied to the cashew nut image data) such as contrast enhancement, Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) components (approximation, diagonal, vertical, and horizontal) and threshold-based segmentation. The analysis was conducted in terms of conventional

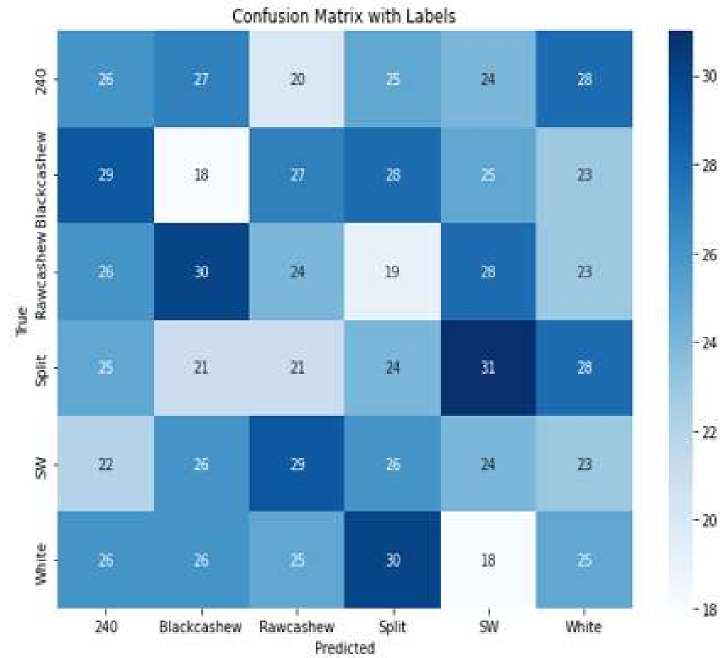

Figure 5. Confusion Matrix for the DWT image dataset containing 1200 images

performance measures, including accuracy, error rate, recall, specificity, precision, negative predictive value (NPV), false positive rate (FPR), false negative rate (FNR) and false discovery rate (FDR) and computational time. The experimental findings demonstrate that the HCER-ResNet50 had the same level of classification of all the preprocessing algorithms with the value of accuracy of between 0.714 and 0.726. The greatest accuracy of 0.726 was achieved on the DWT horizontal component with the DCT and DWT approximation methods close behind at 0.725. Competitive performance was also shown by contrast-based preprocessing and thresholding with the results of 0.724. The DFT based dataset yielded marginally lower results with more accurate representation of features at 0.71.

The error levels were quite low at all datasets ranging between 0.274 and 0.285, which confirms the strength of the proposed HCER-ResNet50 architecture regardless of the preprocessing technique applied. The mean recall was between 0.143 and 0.178 with the biggest recall in the DWT horizontal and the DCT datasets implying that it is able to identify the positive cashew classes better. All the preprocessing methods had a high specificity of between 0.828 and 0.837, which means that they per-

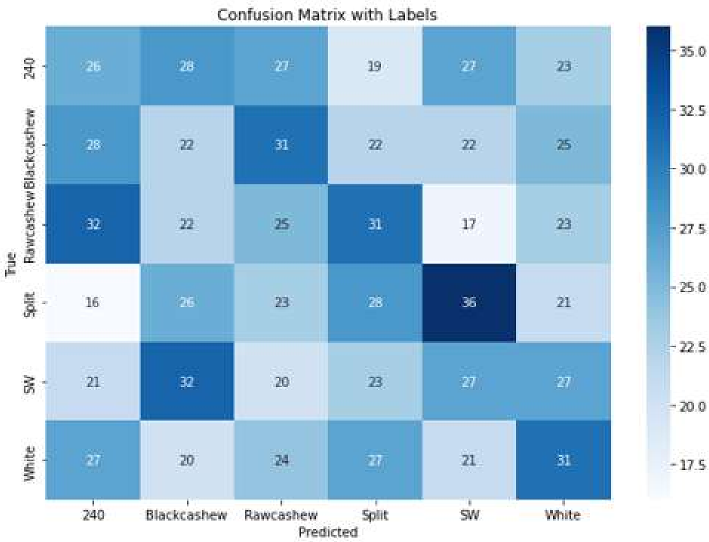

Figure 6. Confusion matrix for Thresholding Operation dataset containing 1200 images

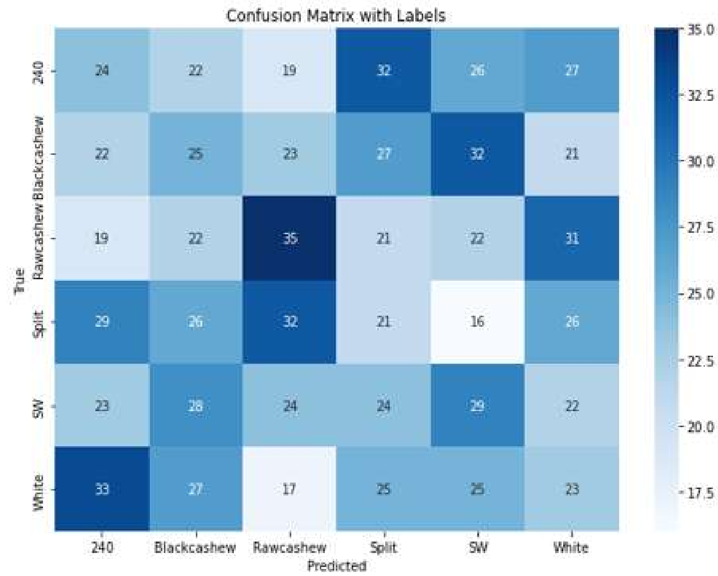

Figure 7. Confusion matrix for Contrast Stretching Dataset Containing 1200 Images

formed well on the correct identification of negative samples. Recall and Precision values showed similar trends with the values being 0.143 to 0.178 which is stable positive prediction. The negative predictive value was always high (around 0.829-0.835), which proved the reliable classification of negative cases. The false positive was also low in all datasets (around 0.164-0.171), which means that there is not much false negative samples being classified as positive.

Nevertheless, the false negative value was comparatively bigger (0.821-0.856) which indicates that a few positive cases were missed, which can be explained by intra-class similarity of cashew kernels. Equally, false discovery rates also traversed the same trend as false negative rates, which implies the difficulty in identifying a slight visual difference between cashew categories. Computational wise, the training time was different based on the preprocessing method. The longest training time was shown to be experienced by the DWT vertical component (around 55 min) whereas the other preprocessing models (DCT, contrast enhancement, and thresholding) took 18-22 minutes to reach model convergence. On the whole, it can be concluded that the proposed HCER-ResNet50 model is nearly consistent in terms of performance with various preprocessing methods, although the features derived with the DWT method, especially the horizontal component, exhibit a little bit higher accuracy in classification. The high specificity, constant accuracy, and average recall verify the efficiency of the HCER-ResNet50 in the cashew nut classification, as well as point to the chances of the further enhancement of its sensitivity, provided through the improved feature extraction and model optimization

The classification of the Convolutional Neural Network on the Cashewnuts was again tested using the dataset of completely 7200 images. The dataset was further subdivided as the training set, which consisted of 5400 images with six target classes (240, Blackcashew, Rawcashew, Split, SW, White) all of which had 900 images each. In the training set, a total of 1,800 images were used across the six target classes, with 300 images per class. This experiment used the same hyperparameters as those employed for training the HCER-ResNet50 model with 1,200 images. The model performance was evaluated using the performance metrics presented in Figures 8 and 9 and the results of the performance metrics were tabulated in Table 2.

Figure 8 shows the confusion matrix for the original cashewnuts dataset containing 7200 images.

Figure 9 shows the confusion matrix for the DCT cashewnuts dataset containing 7200 images.

Figure 10 presents the confusion matrix for contrast stretching dataset containing 7200 images.

Table 1. Performance of HCER-ResNet50 for 1200 Images

	PNG Dataset	Contrast	DCT	DFT	DWT	DWT Approx	DWT Diagonal	DWT Vertical	DWT Horizon	Threshold
Time	22m:49s:29ms	19m:30s:48ms	21m:45s:79ms	20m:53s:62ms	17m:58s:69ms	55m:31s	18m:18s:67ms	18m:21s:35ms	18m:30s:37ms	19m:41s:9ms
Accuracy	0.724	0.725	0.714	0.725	0.720	0.716	0.722	0.715	0.726	0.724
Error Rate	0.276	0.275	0.285	0.275	0.280	0.283	0.277	0.284	0.274	-
Recall	0.172	0.176	0.143	0.176	0.156	0.150	0.166	0.146	0.178	0.174
Specificity	0.834	0.835	0.828	0.835	0.831	0.830	0.833	0.829	0.837	0.834
Precision	0.172	0.176	0.143	0.176	0.156	0.150	0.166	0.146	0.178	0.174
Negative Predicted Value	0.834	0.835	0.828	0.835	0.831	0.830	0.833	0.829	0.835	0.834
False Positive Rate	0.165	0.164	0.171	0.164	0.170	0.170	0.166	0.170	0.164	0.165
False Negative Rate	0.827	0.823	0.856	0.823	0.843	0.850	0.833	0.853	0.821	0.825
False Discovery Rate	0.827	0.823	0.856	0.823	0.843	0.850	0.833	0.853	0.821	0.825

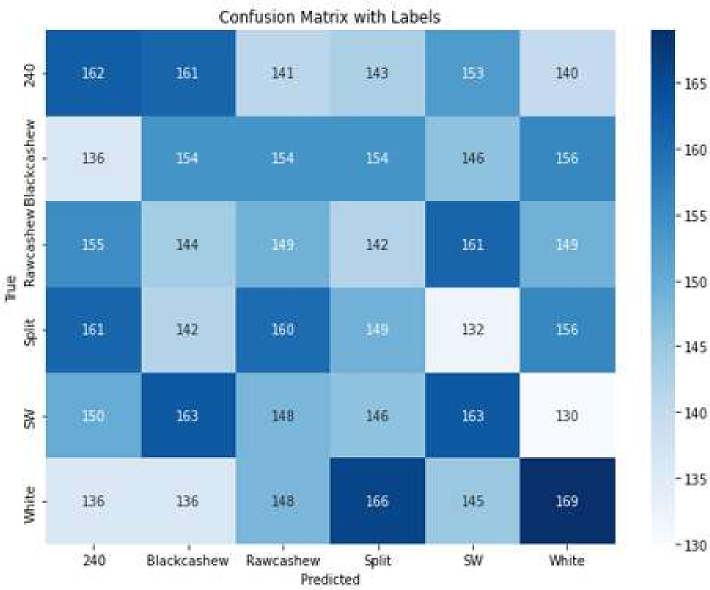

Figure 8. Confusion Matrix for the Original Cashewnuts dataset containing 7200 images

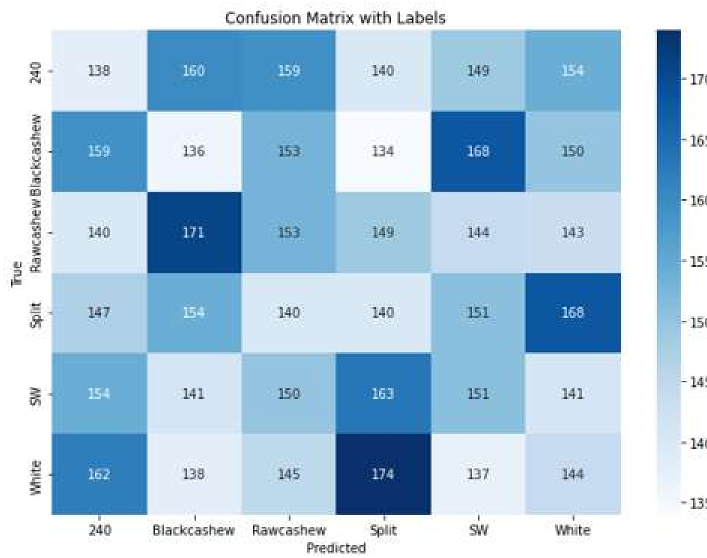

Figure 9. Confusion Matrix for the DCT Cashewnuts dataset containing 7200 images

Figure 11 presents the confusion matrix for thresholding operations dataset containing 7200 images.

The efficiency of the proposed HCER-ResNet50 given in Table 2 was tested on a large-scale image dataset of cashew nuts comprising 7200 images based on various preprocessing methods such as contrast enhancement, Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) components (approximation, diagonal, vertical and horizontal) and threshold-based segmentation. The assessment was conducted on the basis of typical classification measures, that is, accuracy, error, recall, specificity, precision, negative predictive value (NPV), false positive rate (FPR), false negative rate (FNR), false discovery rate (FDR), and computational time. The overall accuracy was between 0.718 and 0.725 with highest accuracy of 0.725 with contrast-enhanced images, second was threshold-based preprocessing of 0.723 and thirdly, the DWT-based representations with the highest accuracy of about 0.720-0.722.

The DFT generated dataset was slightly worse at 0.719, which means that there was relatively low discrimination of features than with the space-domain enhancement methods. The error was also stable and low in all datasets (between 0.275 and 0.282), which indicates the strength and the generalization ability of the HCER-ResNet50 model. The recall values were between 0.154 and 0.175 with a maximum value of recall observed to be contrast-based preprocessing and thresholding which demonstrated higher detection of true positive cashew classes in such circumstances. Specificity values were also quite high in all preprocessing techniques (0.831-0.835) indicating that the HCER-ResNet50 has a high ability to detect negative sample correctly. The trends of precision values were

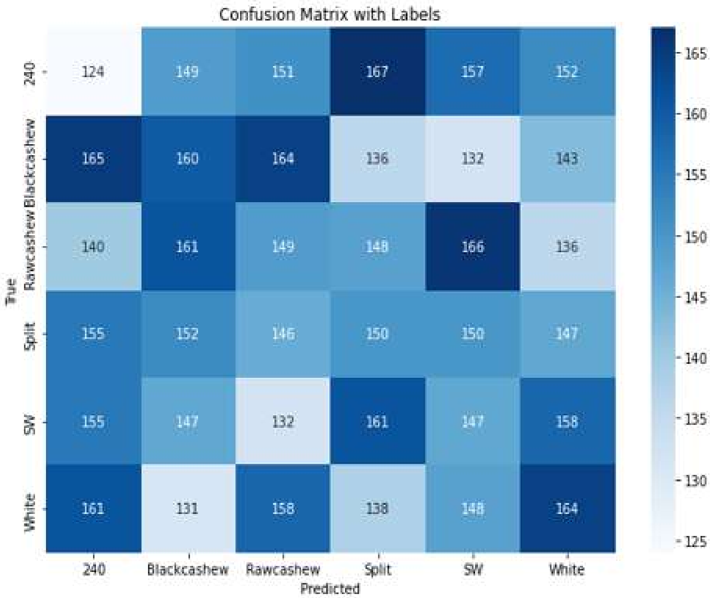

Figure 10. Confusion Matrix for Contrast Streching Dataset Containing 7200 images

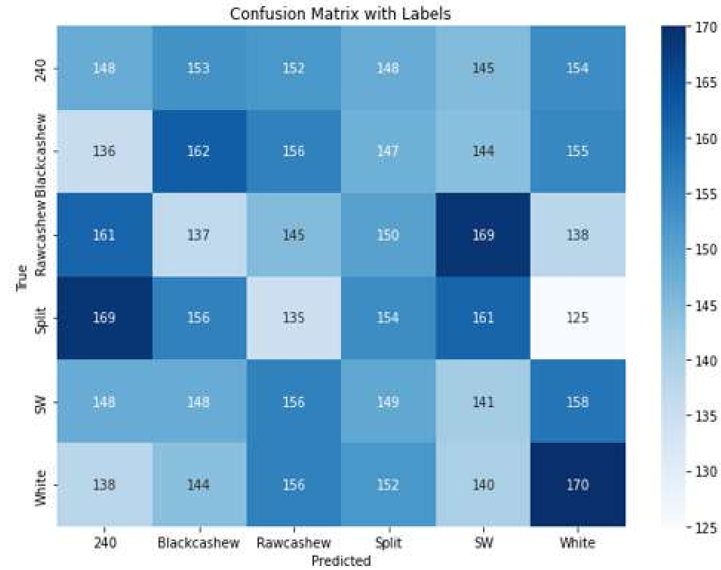

Figure 11. Confusion Matrix for Thresholding Operations Dataset Containing 7200 images

the same as those of recall, and the values varied between 0.159 and 0.175, indicating equal good performance on prediction. The negative predictive value was still high (0.831-0.835), which means that it is effective in classifying negative cases. False positive rate was also low and reliable (0.164-0.169) which ensured that there was less misclassification of the non-target classes. Nonetheless, the rate of false negative was relatively high (0.824-0.846) indicating that not all positive cashew groups were identified as such, which could be explained by high intra-class closeness and visual overlap of the types of cashew kernel. The false discovery rate had a common trend that supported the problem of differentiating minor morphological differences between classes.

Computationally, the training time was significantly different with the preprocessing method and dataset representation. The training time of most preprocessing methods was about 2 hours, whereas the maximum time of computation of DWT vertical component processing was the highest (about 5 hours and 19 minutes), which implied more complicated computations involved in the feature extraction. Other methods like contrast enhancement, DCT as well as thresholding needed relatively lower training time (around 2 hours), which proved that the HCER-ResNet50 model converged efficiently. Generally, the experimental findings validate that the proposed HCER-ResNet50 architecture can be stable and reliable when conditions of the preprocessing strategies of large-scale cashew nut classification are varied. The best overall performance was obtained with the use of contrast-enhanced images whereas DWT and threshold-based datasets also demonstrated favorable results. The high specificity and low false-positive rates demonstrate the suitability of the HCER-ResNet50 model in differentiating non-target classes. The average recall values indicate that sensitivity can be further improved through enhanced feature extraction and data balancing.

Table 3 shows the performance analysis of the proposed HCER-ResNet50 model on the 7200-image dataset with optimizer utilizing various image preprocessing models. The method was trained using the Adam optimizer, which ensures effective parameter optimization and rapid convergence during the training process. Therefore, all pre-processing methods gained maximum classification result with accuracies over 0.95. Particularly, the PNG Dataset given the greatest accuracy of 0.9720 and the minimum error rate of 0.0280. The additional performance measures, such as precision, recall, specificity, and negative predictive value, also specify

Table 2. Performance of HCER-ResNet50 for 7200 Images

	PNG Dataset	Contrast	DCT	DFT	DWT	DWT Approx	DWT Diagonal	DWT Vertical	DWT Horizon	Threshold
Time	2h:09m:19s:22ms	2h:6m:22s:67ms	2h:57m:37s:20ms	2h:5m:33s:54s	1h:58m:55s	5h:19m:30s:61ms	2h:00m:18s	2h:19m:27s:01ms	2h:11m:09s:4ms	1h:59m:01s:99ms
Accuracy	0.725	0.721	0.719	0.720	0.722	0.722	0.721	0.221	0.718	0.723
Error Rate	0.275	0.280	0.281	0.280	0.278	0.278	0.279	0.279	0.282	0.277
Recall	0.175	0.165	0.159	0.163	0.167	0.166	0.162	0.164	0.154	0.170
Specificity	0.835	0.833	0.831	0.832	0.833	0.833	0.832	0.833	0.831	0.834
Precision	0.175	0.165	0.159	0.163	0.167	0.166	0.162	0.164	0.154	0.170
Negative Predicted Value	0.835	0.833	0.831	0.832	0.833	0.833	0.832	0.833	0.831	0.834
False Positive Rate	0.164	0.166	0.168	0.167	0.166	0.166	0.168	0.167	0.169	0.166
False Negative Rate	0.824	0.834	0.840	0.837	0.832	0.834	0.838	0.836	0.846	0.830
False Discovery Rate	0.824	0.834	0.840	0.836	0.832	0.834	0.838	0.836	0.846	0.830

the overall effectiveness of the proposed method for automated cashew nut classification. Tables 1 and 2 present the performance of the model without the optimizer, whereas the results in Table 3 were achieved using the optimizer, resulting in improved classification performance.

Table 4 denotes the comparative analysis of the proposed HCER-ResNet50 methodology. The proposed model attains a highest performance, accuracy of 0.9720, Precision of 0.0280, and specificity 0.9040, correspondingly. In accordance to accuracy, the Naïve Bayes, GBM, CNN, L-MTCNN, Inception-V3, EfficientNet-B0, and ResNet-50 techniques achieves lesser performance, accuracy of 0.6650, 0.6023, 0.6599, 0.6430, 0.5891, 0.6036, and 0.7250, respectively. All baseline models have been trained and evaluated under identical experimental setup, comprising same dataset, preprocessing, data partitioning, and training configuration, when the presented HCER-ResNet50 was evaluated under the same conditions with the selected optimizer configuration.

The present research deploys a HCER-ResNet50 pipeline to extract features on images of cashew nut products. The input images are processed by the system by converting them to grayscale, and a three-layer convolutional architecture that is written in NumPy is used. Firstly, the RGB images are turned into grayscale in order to minimize the complexity of computations without loss of structural features. The edges and patterns are then identified by a sequence of convolution operations that use manually designed filters. The next step is the convolution step, and the second step is the Rectified Linear Unit (ReLU) activation to observe non-linearity and optimize learning of features. The convolutions are followed with max-pooling to decrease the spatial dimensions and preserve major characteristics. The network is made up of three convolutional layers with the Layer 1 extracting low-level features like edges as shown in Figure 13.

Figure 12. Grey Scale image of 240 Cashewnuts

Figure 12 shows the grayscale image of 240 Cashewnuts.

Figure 13 shows how Layer 1 extracts low-level features such as edges.

The intermediate textures and shapes as shown in Figure 14 are captured in the second layer. Lastly the third Layer 3 acquires high-level structural patterns as shown in Figure 15. The feature maps at every stage are visualized in order to analyze the transformation of information contained in the image across layers. This strategy provides interpretability by demonstrating how the HCER-ResNet50 architecture leverages deep residual feature learning to extract hierarchical representations for accurate image classification. The implementation is used as a multi-level basis of cashew nut image classifi-

Table 3. Performance of HCER-ResNet50 for 7200 Images with Optimizer

	PNG Dataset	Contrast	DCT	DFT	DWT Approx	DWT Diagonal	DWT Vertical	DWT Horizon	Threshold	DWT Approx
Accuracy	0.9720	0.9703	0.9676	0.9685	0.9710	0.9704	0.9718	0.9703	0.9598	0.9702
Error Rate	0.0280	0.0277	0.0324	0.0315	0.0290	0.0296	0.0282	0.0297	0.0402	0.0238
Recall	0.8178	0.8011	0.8108	0.7988	0.7921	0.7924	0.8073	0.7913	0.7709	0.8008
Specificity	0.9040	0.8420	0.8540	0.8960	0.8630	0.7790	0.8730	0.8190	0.7800	0.8130
Precision	0.7979	0.8026	0.7080	0.8303	0.8426	0.8217	0.8124	0.7789	0.7914	0.8674
Negative Predicted Value	0.8370	0.8940	0.9020	0.9210	0.8400	0.7500	0.8750	0.8850	0.8510	0.8550
False Positive Rate	0.1060	0.2580	0.2460	0.1840	0.1370	0.2210	0.1270	0.1810	0.2200	0.1870
False Negative Rate	0.1822	0.1989	0.1892	0.2012	0.2079	0.2076	0.1927	0.2087	0.2291	0.1992
False Discovery Rate	0.2021	0.1974	0.2920	0.1697	0.1574	0.1783	0.1876	0.2211	0.2086	0.1326

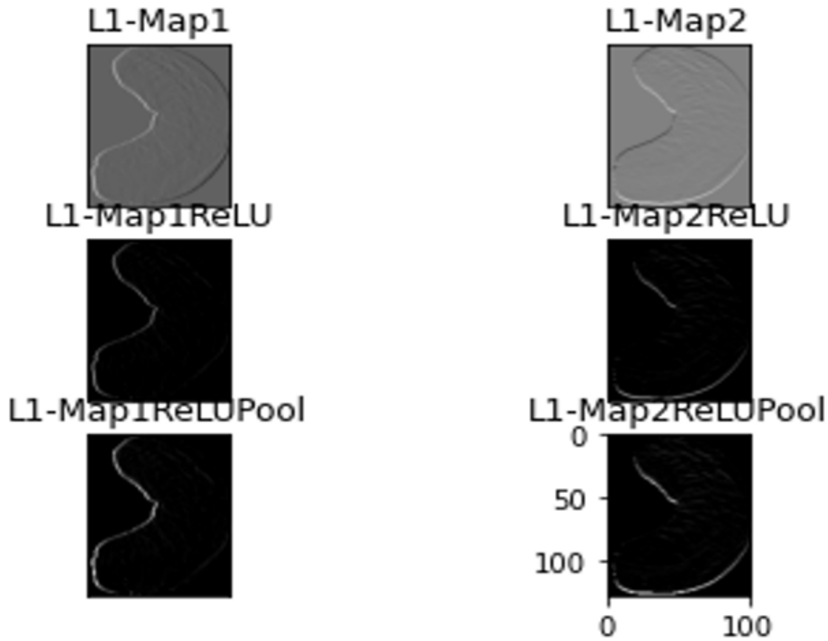

Figure 13. Layer 1 extracts low-level features such as edges

Table 4. Comparative Analysis of the Proposed Model

Algorithm	Accuracy	Error Rate	Specificity
Naïve Bayes	0.6650	0.3350	0.5383
GBM	0.6023	0.3977	0.6951
CNN	0.6599	0.3401	0.6472
L-MTCNN	0.6430	0.3570	0.6703
Inception-V3	0.5891	0.4109	0.5598
EfficientNet-B0	0.6036	0.3964	0.6940
ResNet50	0.7250	0.2750	0.8350
HCER-ResNet50	**0.9720**	**0.0280**	**0.9040**

cation and has the potential to be applied to automated detection of defects and quality grading.

Figure 14 shows Layer 2 captures intermediate textures and shapes.

Figure 15 shows Layer 3 learns high-level structural patterns.

6 Conclusion

The current research proposed Hybrid Compression and Enhancement based ResNet-50 (HCER-ResNet50) architecture of automated cashew nuts grading with image processing and deep learning. The effectiveness of compressing methods based on transform techniques (DWT, DFT and DCT) and enhancement techniques such as contrast stretching and thresholding was examined using a complete dataset gathered in the real-field conditions. A HCER-ResNet50 model was constructed with image processed datasets to classify cashew nuts into six grades showed that it was achievable to integrate image acquisition, image pre-processing, and deep learning into a grading algorithm. The results of the experiment demonstrated that the proposed model demonstrated similar performance in classifica-

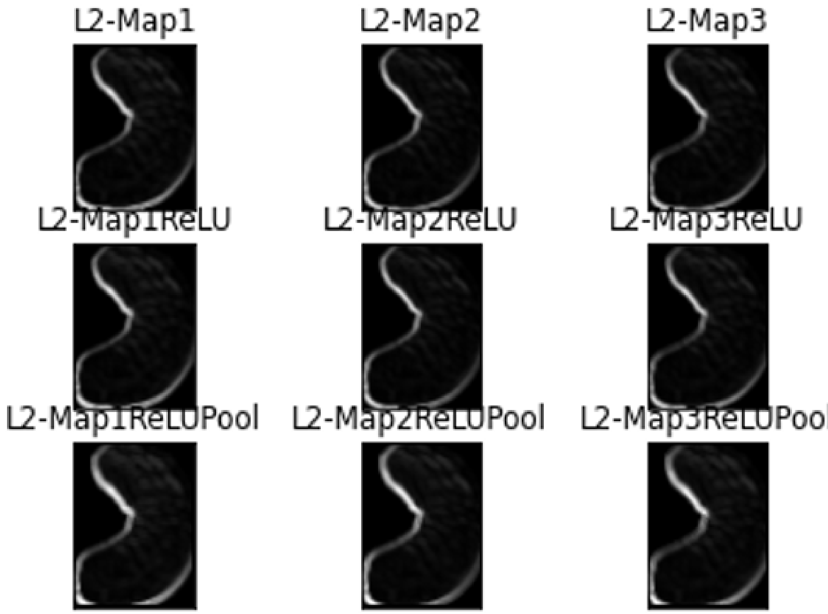

Figure 14. Layer 2 captures intermediate textures and shapes

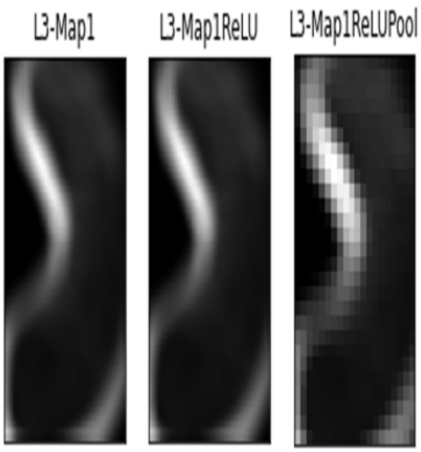

Figure 15. Layer 3 learns high-level structural patterns

tion when using any pre-processing strategy with the accuracy of 0.9598-0.9720. The best results were gained with DWT horizontal component and DCT and DWT approximation made similar results which means that transform based feature extraction is effective in the case of cashew grading. This model had low error rates and high specificity and constant negative predictive values which have ensured that the model was reliable in terms of distinguishing defective and non-defective samples. The computational analysis showed that the majority of the image processing methods needed 18-22 minutes to converge to the model, and the processing time of the DWT vertical component was much-higher, though the gain in accuracy is insignificant, which demonstrates the trade-off between the computational complexity and the increase in the processing speed.

Moreover, the usage of the Adam optimizer allowed effective parameter optimization values and fast convergence during training time, furthering to enhanced accuracy and reliability of the proposed HCER-ResNet50 method. All in all, the developed HCER-ResNet50 framework is a scalable, consistent, and automated way of grading cashew nuts that would decrease the reliance on manual checking of the cashew and enhance the efficiency of the grading process. The results also demonstrate the possibility of additional improvement by the better feature representation, optimization of model parameters, as well as, the integration of the advanced deep learning strategies to enhance sensitivity and classification accuracy in real-time industrial bids.

Although the presented HCER-ResNet50 cashew quality classification architecture attains a satisfactory performance, certain limitations remain. The dataset utilized in this research has been originally acquired and developed by the authors using a camera-based image acquisition setup. As a outcome, variations in illumination conditions, sample orientation, image quality, and background characteristics were present within the dataset. Moreover, some cashew grades show highly similar visual appearances, making the classification operation inherently challenging. These factors might have achieved this classification accuracy. And also, the use of a relatively compact HCER-ResNet50 framework designed to establish an effective baseline framework for automated cashew grading. When the model illustrates its ability to distinguish six various cashew quality classes, more advanced framework might offer enhanced feature representation and classification performance. Future work will focus on increasing the size of dataset with a larger number of samples collected under various environmental conditions to enhance model normalization. Additional improvements may comprise the application of data augmentation techniques, systematic transfer learning approaches, and the investigation of deeper convolutional neural network architectures. These enhancements are expected to strengthen feature extraction capabilities, improve robustness against image variations, and further increase classification accuracy for real-world deployment.

Declarations

Data Availability Statement

The data that support the findings of this study are openly available in: https://www.kaggle.com/datasets/muthumphil11/cashewnuts-datasets
reference number [24].

Code Availability Statement

The code that support the findings of this study are available on request from the corresponding author.

Author Contributions

S. Muthukumaran: Conceptualization, Methodology, Supervision, Software, Writing- Original draft preparation **B. Kamatchi**: Data curation,. **P. Arivazhagan**: Visualization, Investigation. **N. Kalaichelvi**: Writing-Reviewing and Editing **R. Mahalakshmi**:Validation.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript. All authors contributed to the manuscript and approved the final version.

Ethics Approval

This article does not involve any studies with human participants or animals performed by any of the authors.

Consent to Participate

Not applicable.

Informed Consent

Not applicable.

Clinical Trial Registration

Not applicable.

Competing Interests and Funding

The authors did not receive support from any organization for the submitted work.

Funding Statement

The authors received no funding for this study.

Human Participants and/or Animals

Not applicable.

References

[1] M. O. Alaka, J. K. Ogunjobi, O. E. Omoruyi, O. O. Oluwasina, and L. Lajide, “Transforming extracted cashew nut shell into biochar and its application as soil amender for jute mallow (Corchorus olitorius L.) cultivation,” *Sci. Rep.*, vol. 16, no. 1, p. 5101, Jan. 13, 2026. doi: 10.1038/s41598-026-35559-6.

[2] I. Berger, A. Pannure, A. Harris, P. Basu, B. M. Smith, and L. V. Dicks, “Agroecological cashew cultivation increases pollinator abundance, diversity and flower visitation rates, with potential yield benefits,” *Agric. Ecosyst. Environ.*, vol. 396, p. 110006, 2026. doi: 10.1016/j.agee.2025.110006.

[3] Mendes, Sousa, A. de S. Brito Neto, E. da S. Justino, C. J. L. Herbster, R. L. Oliveira, L. R. Bezerra, and E. S. Pereira, “Effect of natural cashew nut shell liquid on lactation performance of dairy goats,” 2026.

[4] J. S. Correia, D. P. Amorim-Neto, E. Lang, M. P. da Silva Lascaris, B. M. D. Pian, L. P. Margalho, et al., “Microbiological quality of plant-based cheese analogues and pathogen behavior in cashew nut-based varieties,” *Int. J. Food Microbiol.*, vol. 451, p. 111670, Apr. 16, 2026. doi: 10.1016/j.ijfoodmicro.2026.111670.

[5] T. Ayedon, R. O. Awode, and A. Y. J. Akossou, "Trading modes of cashew (Anacardium occidentale L.) by small producers in Central Benin," *Int. J. Account. Finance Audit. Manag. Econ.*, vol. 7, no. 1, pp. 456–480, 2026.

[6] M. T. Mallick, D. O. Murty, R. Pal, S. Mandal, H. N. Saha, and A. Chakrabarti, "High-speed system-on-chip-based platform for real-time crop disease and pest detection using deep learning techniques," *Comput. Electr. Eng.*, vol. 123, p. 110182, 2025. doi: 10.1016/j.compeleceng.2025.110182.

[7] M. Farhadi, Y. Abbaspour-Gilandeh, A. Mahmoudi, and J. Mari Maja, "An integrated system of artificial intelligence and signal processing techniques for the sorting and grading of nuts," *Appl. Sci.*, vol. 10, no. 9, p. 3315, 2020.

[8] V. N. Pham, Q. H. Do Ba, D. A. Tran Le, Q. M. Nguyen, D. Do Van, and L. Nguyen, "A low-cost deep-learning-based system for grading cashew nuts," *Computers*, vol. 13, no. 3, p. 71, 2024.

[9] J. Zheng, Z. Lv, D. Li, C. Lu, Z. Shen, H. Chen, et al., "High-efficiency real-time rice leaf disease classification using convolutional neural network accelerator on FPGA for edge computing in precision agriculture," *Measurement*, vol. 256, p. 118122, 2025. doi: 10.1016/j.measurement.2025.118122.

[10] C. Jin, L. Zhou, Y. Zhao, H. Qi, X. Wu, and C. Zhang, "Classification of rice varieties using hyperspectral imaging with multi-dimensional fusion convolutional neural networks," *J. Food Compos. Anal.*, vol. 148, p. 108389, 2025. doi: 10.1016/j.jfca.2025.108389.

[11] L. Zhang, J. Du, J. Li, C. Li, J. Zhang, and L. Wu, "Early detection of tobacco leaf mildew using multi-attention enhanced 3D residual convolutional neural network with hyperspectral imaging," *Infrared Phys. Technol.*, vol. 154, p. 106375, 2026. doi: 10.1016/j.infrared.2026.106375.

[12] S. J. Wei, D. F. Al Riza, and H. Nugroho, "Comparative study on the performance of deep learning implementation in the edge computing: Case study on the plant leaf disease identification," *J. Agric. Food Res.*, vol. 10, p. 100389, 2022. doi: 10.1016/j.jafr.2022.100389.

[13] D. S. M. Zakir, Z. Alwani Shaffiei, M. Ghazali, and V. S. Briane Paul, "Comparative analysis of VGG-16, ResNet50, and EfficientNet-B1 with optimization techniques for crop disease detection," *J. Adv. Res. Des.*, vol. 143, no. 1, pp. 55–64, 2026.

[14] M. Momeny, A. Jahanbakhshi, K. Jafarnezhad, and Y. D. Zhang, "Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach," *Postharvest Biol. Technol.*, vol. 166, p. 111204, 2020.

[15] M. Abbaszadeh, A. Rahimifard, M. Eftekhari, H. G. Zadeh, A. Fayazi, A. Dini, and M. Danaeian, "Deep learning-based classification of the defective pistachios via deep autoencoder neural networks," *arXiv preprint arXiv:1906.11878*, 2019.

[16] S. Ledbin Vini and P. Rathika, "Advanced state-of-the-art approaches in image processing techniques for plant leaf diseases analysis," *Int. J. High Perform. Comput. Appl.*, 2026. doi: 10.1177/10943420251408184.

[17] M. Nouman Noor, M. Masab, F. Haneef, M. Hussain, M. Yaqoob, T. Mazhar, et al., "An effective approach for recognition of crop diseases using advanced image processing and YOLOv8," *Food Sci. Nutr.*, vol. 14, no. 2, p. e71504, Feb. 9, 2026. doi: 10.1002/fsn3.71504.

[18] G. Isik and F. Keskin, "CA-SAUNet: Coordinate-Attention Enhanced Shape Attentive UNet for medical and agricultural image segmentation," *Pattern Anal. Appl.*, vol. 29, no. 1, p. 6, 2026. doi: 10.1007/s10044-025-01590-y.

[19] M. Saha and S. Das, "Optimized plant identification through leaf feature extraction using deep learning and advanced digital image processing techniques," 2026. doi: 10.1108/978-1-83662-996-220251006.

[20] A. S. Rathor, S. Choudhury, A. Sharma, R. K. Pachauri, G. Shah, N. Fatema, et al., "Deep convolutional network for plant monitoring in precision agriculture," *Sustain. Comput.*, vol. 50, p. 101302, 2026. doi: 10.1016/j.suscom.2026.101302.

[21] A. Usha Ruby, J. George Chellin Chandran, B. N. Chaithanya, T. J. Swasthika Jain, and R. Patil, "Wheat leaf disease classification using modified ResNet50 convolutional neural network model," *Multimedia Tools Appl.*, vol. 83, no. 23, pp. 62875–62893, 2024.

[22] M. Hemam, "Evaluation of SE, ECA, and CA modules in ResNet-50 for the classification of grape leaf diseases," *Syst. Comput.*, vol. 2, no. 1, 2026.

[23] M. Winklmair, R. Sekulic, J. Kraus, P. Penava, and R. Buettner, "A deep learning-based approach for classifying the maturity of cashew apples," *PLoS One*, vol. 20, no. 6, p. e0326103, 2025.

[24] S. Muthukumaran, “Cashewnuts Datasets,” Kaggle, 2026. [Online]. Available: https://www.kaggle.com/datasets/muthumphil11/cashewnuts-datasets