IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 1, February 2026, pp. 642~654
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i1.pp642-654

Deformable spatial pyramid pooling-enhanced EfficientNet with weighted feature fusion for pomegranate fruit disease diagnosis

Harish Bommenahalli Mallikarjunaiah[1], Balaji Prabhu Baluvaneralu Veeranna[2]
[1]Department of Information Science and Engineering, Malnad College of Engineering-Hassan, Affiliated to Visvesvaraya Technological University, Belagavi, India
[2]Department of Computer Science and Engineering, Malnad College of Engineering-Hassan, Affiliated to Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jul 9, 2025
Revised Jan 1, 2026
Accepted Jan 22, 2026

Keywords:

CLAHE
Deep features
Deformable spatial pyramid pooling
Handcrafted features
Pomegranate fruit disease
Support vector machine

ABSTRACT

Pomegranate is a fruit of high nutritional and economic importance. Still, it is highly susceptible to different diseases during its growing stages, leading to significant yield losses and financial setbacks for farmers. This article proposes a novel disease detection model that integrates handcrafted features with deep features extracted using a developed deformable spatial pyramid pooling (DSPP)-EfficientNet architecture. Handcrafted features such as color (RGB and HSV histograms), texture features from gray level co-occurrence matrix (GLCM), and shape attributes extracted from contour descriptors and Hu moments are captured and fused with deep features by weighted fusion strategy, resulted in the most discriminative information. The fused features are categorized using a support vector machine (SVM) in a classification phase, which effectively classifies different classes of pomegranate fruit diseases. The combined deep and handcrafted features obtained 96.66% accuracy, 96.26% precision, 96.50% recall, 96.37% F1-score, and 95.64% specificity on the pomegranate fruit disease dataset which compared to existing techniques.

Corresponding Author:

Harish Bommenahalli Mallikarjunaiah
Department of Information Science and Engineering, Malnad College of Engineering-Hassan
Affiliated to Visvesvaraya Technological University
Belagavi, India
Email: harishbm137@gmail.com

1. INTRODUCTION

Pomegranate, known for their high nutrient context and unique blend of sweet and tart flavors, is a major fruit crop primarily cultivated in arid and semi-arid areas [1], [2]. This has different qualities and is native on Afghanistan and Iran. In recent years, pomegranate cultivation has extended to areas over Africa, Australia, America, Europe, and Middle East [3]–[5]. Consumers enjoy pomegranates in different formats, involves fresh fruit, juice, oil, jam, and infusion [6]. Fruits are predisposed to diseases and particularly pomegranates are greatly prone to infections in different development phases [7]. The vulnerability of pomegranates to diseases causes effective economic challenges to farmers, highlighting the requirement to identify development phases for suitable plant monitoring and fulfilment on critical requirements [8]. Obtaining knowledge about variant segments on pomegranate fruit growing ensures a cultivators make better choices about fertilization, controlling pests and optimum timing of harvest, through optimizing the fruit quality and yield [9]. This is significant for acknowledging the timing and attributes integrated with every development phase that vary based on parameters like the variety of pomegranate, the conditions of the environment and the criteria for cultivation [10]. The cultivation of pomegranate has essential significance in the agricultural field because of the growing demand for their nutritious fruits and their different applications in beverages, food and cosmetics [11]. Though, the industry struggles with several challenges, including disease management, aligning quality evaluation with development stages, fluctuation of climate and managing soil nutrient balance [12]–[14]. Farmers often struggle to precisely identify the pomegranate development phases, leading to inefficient resource allocation and affecting decisions related to harvest timing and entire crop managing [15]–[18]. Imprecision in identification resulted in source wastage, losses in finance and drawbacks on opportunities to advanced effectiveness and sustainability on farming practices [19]. For addressing these challenges, advanced techniques like transfer and deep learning (DL) are assigned for mitigating issues faced through farmers and exploring potential avenues on simple and improved agricultural practices.

Wang *et al.* [20] developed tomato leaf disease detection algorithm-based attention mechanisms and multi-scale feature fusion. Initially, convolutional block attention module (CBAM) was implemented in a backbone feature extraction network for improving a capability to capture affected features. Next, shallow feature maps were developed to re-parameterized generalized feature pyramid network (RepGPFN) to enhance localization capability for small lesion features. At last, RepGFPN replaced path aggregation feature pyramid network (PAFPN) in a you only look once version 6 (YOLOv6) method for obtaining efficient integration of deep semantic and shallow spatial data. Model suffered from noise and compression artifacts, which minimizes quality of disease detection. Khan *et al.* [21] presented a method which employed robust feature extraction, including gray level co-occurrence matrix (GLCM) and scale invariant feature transform (SIFT), integrated with support vector machine (SVM) for effective classification. The extended dataset of 2,700 tomato leaf images with a minimum of 300 images for every nine different disease classes. This data facilitated the training and testing of various machine learning (ML) and DL-based algorithms. Though uneven lighting and low contrast make disease symptoms complex for detection, employing contrast limited adaptive histogram equalization (CLAHE) improves local contrast and enhances visibility of these symptoms. Oad *et al.* [22] suggested the artificial intelligence (AI) method, which detected and explained plant diseases by image analysis. The suggested method identified several diseases in fruits and vegetables through assigning an ensemble learning classifier with four DL algorithms visual geometry group 16 (VGG16), VGG19, ResNet101v2, and Inception-V3. Additionally, provided explanations for predictions by local interpretable model-agnostic explanations (LIME) employed for interpreting predictions of DL algorithms. Visualizations produced from several algorithms for pixels influence on precise and incorrect predictions. The suggested method missed subtle color, texture, and shape features; in contrast, this article captured detailed handcrafted features for precisely extracting visual variations. Naseer *et al.* [23] introduced transfer learning-enabled CRNet algorithm to capturing spatial features from pomegranate images in 5 phases of pomegranate growth. The captured spatial features were fed into random forest (RF) algorithm, resulting in a development of new probabilistic feature set. These features assisted in precisely identifying pomegranate developmental phases. For evaluating performance, existing classification algorithms considered were convolutional neural network (CNN), k-neighbors classifier (KNC), logistic regression (LR), and Gaussian naïve Bayes (GNB). Traditional cell-level learning (CLL) algorithms have difficulty handling multi-scale and irregular disease patterns. Combining deformable spatial pyramid pooling (DSPP) with EfficientNet enables adaptive, multi-scale pooling concentrated on disease-relevant regions.

Hu *et al.* [24] implemented the lightweight detection technique that depends on enhanced YOLOv5. Initially, Faster-C3 module was developed to replace the actual cross stage partial (CSP) module in YOLOv5 for effectively minimizing number of parameters in the feature extraction process. Then, CoordConv and enhanced content-aware reassembly of features (CARAFE) were implemented in the neck network for enhancing the refinement of position data in feature fusion and refining semantic data in the down-sampling process. At last, channel-wise knowledge distillation algorithm was utilized in model training for enhancing detection accuracy without maximizing the number of model parameters. Considering handcrafted and deep features separately causes incomplete data representation; a weighted fusion strategy was developed to effectively integrate both feature types. Jiang *et al.* [25] developed the method based on efficient feature segmentation transformer (EFS-Former). The extended local detail (ELD) module extended the receptive field of the model through extending the convolution, good handling of fine spots and efficiently minimizing data loss. H-attention minimized computational redundancy through imposing multi-layer convolutions, enhancing feature filtering. Parallel fusion framework efficiently utilized various intervals of feature extraction of CNN and transformers encoders, obtaining intrinsic feature extraction and integrating semantic data in channel and spatial dimensions in feature fusion module (FFM). The initial classifiers failed to differentiate between closely resembling disease symptoms; incorporating an SVM classifier improved boundary-enabled separation, improving overall classification accuracy. Traditional disease detection and classification algorithms are manual, time-consuming and lead to human error, making timely and precise diagnosis challenging. Visual symptoms like discoloration, texture changes and shape deformation are subtle and vary in intensity, causing challenges for traditional methods. Additionally, inconsistent lighting and image noise reduced reliable feature extraction. The primary objective of this paper is to develop hybrid image-based disease detection algorithm for pomegranate fruits through combining handcrafted with DL-based features. The handcrafted features like color, texture and shape are integrated with deep features captured by DSPP-EfficientNet. Then, the weighted feature fusion strategy is employed for integrating these features and classification is performed by SVM.

In this manuscript, proposed DSPP-enhanced EfficientNet with weighted feature fusion is developed in pomegranate disease classification for challenges like varying lesion shapes, subtle texture difference and complex backgrounds. EfficientNet act as lightweight model for deep feature extraction, provides optimal balance between accuracy and computational cost. The enhancement of model's capability to extract multi-scale contextual data, the DSPP is incorporated, allows the model to capture features in different receptive fields. To address the challenges on deep features, proposed model incorporated weighted fusion strategy which integrated deep features with handcrafter features. This fusion process ensures the feature space through semantic data with domain-relevant low-level patterns, enhancing class separability. The usage of learning weights in fusion ensures feature type proportionally for final decision, improving robustness to image variability. The significant contributions of the research are described as follows.

i) Hand crafter features like color histograms, GLCM-based texture descriptors and Hu moments and deep features from developed DSPP-EfficientNet are captured to differentiate the low-level and high-level characteristics of diseases.
ii) Integration of DSPP with EfficientNet for effectively focusing on disease relevance regions, by utilizing adjustable pooling grids and enhancing multi-scale feature representation.
iii) The weighted feature fusion strategy integrates handcrafted and deep features, which balances for much precise and robust disease classification.

This research paper is organized as follows: section 2 provides details of a proposed DSPP-EfficientNet algorithm. Section 3 validates a performance of the DSPP-EfficientNet algorithm. Finally, the conclusion is given in section 4.

2. PROPOSED SECTION

This paper proposes a novel disease detection model that integrates handcrafted features with deep features extracted using a DSPP enhanced EfficientNet architecture. The pomegranate fruit disease dataset is used in this article and the images are pre-processed by using a median filter and CLAHE. Next, the handcrafted and deep features are extracted by DSPP-EfficientNet and these features are fused using weighted feature fusion. In the classification phase, we used the SVM for multi-class classification. Figure 1 represents the process of pomegranate fruit disease classification.

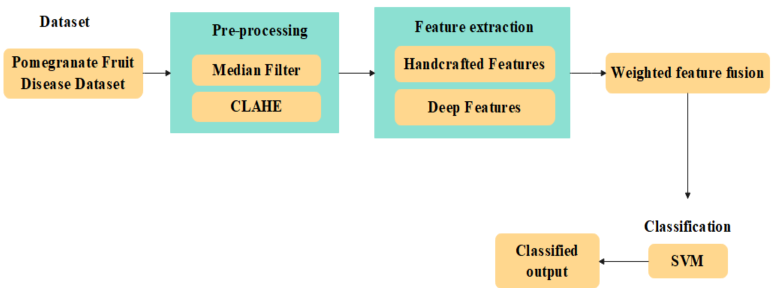

Figure 1. Process of pomegranate fruit disease classification

2.1. Dataset

This paper uses the publicly available pomegranate fruit disease dataset, a standardized resource for categorizing pomegranate fruit diseases [26]. The dataset contains 5,099 labelled and classified images of healthy and diseased pomegranate fruit, divided into five various pomegranate fruit diseases classes such as healthy, bacterial blight, anthracnose, cercospora fruit spot, and alternaria fruit spot. The raw images are given as input to pre-processing phase to improve the quality of images.

2.2. Pre-processing

2.2.1. Median filter

It is a non-linear smoothing method utilized for eliminating the impulse noise when preserving the edges. The window size 5×5 slides over the image. For every pixel, the window captures neighboring pixel values. The values are sorted and median is chosen. Then, the center pixel is replaced with the median values and its mathematical formula is given as (1). It removes background noise caused by compression artifacts. Unlike mean filter, median filter preserves edges, that are essential for identifying disease boundaries and retains fine patterns like spots or diseased regions on a leaf.

$$Output(x, y) = Median\{I(i, j) | (i, j) \in Window\ around\ (x, y)\} \tag{1}$$

2.2.2. CLAHE

CLAHE is a strong method for contrast enhancement, particularly in non-uniform lighting conditions. This is variant of adaptive histogram equalization (AHE) which limits the amplification of noise. The image is separated to small regions called tiles. For every tile, execute a histogram and redistribute the pixel intensities to flatten the histogram. Clip the histogram to the threshold to avoid noise over-amplification. Assign bilinear interpolation for smoothing transitions among tiles, and its mathematical formula is given as (2).

$$CLAHE(x, y) = Interpolated\ output\ from\ tiles\ including\ (x, y) \tag{2}$$

This improves the local contrast, used to highlight the local disease features like discoloration or patches. Also deals with uneven lighting, compensating to non-uniform illumination in captured images. Enhances the visibility of both subtle and strong disease-relevant textures and color variations.

2.3. Feature extraction

The pre-processed images are fed as input to a feature extraction phase to capture handcrafted and deep features to differentiating the various classes of disease regions in pomegranate fruits. The captured, handcrafted and deep features to differentiate diseased regions are described in this section. The extracted handcrafted and deep features are explained in this sub section.

2.3.1. Color features

In this section, the handcrafted features such as color features, texture features and shape features are captured and explained as follows. Numerous plants have color changes like yellowing, brown spots, pale patches, capturing the color distribution helps to identify these symptoms. Initially converts the image to RGB and HSV color spaces. Then divide every channel into bins and count the number of pixels in every bin to form a histogram, at lastly normalize the histogram to develop a feature vector.

2.3.2. Types of color features

Different color feature representations are often utilized in image analysis to capture differences in color information.

i) RGB histogram: captures distribution of red, green and blur intensities. It helps to identify raw color changes
ii) HSV histogram: it effectively separates color (hue) from illumination (value) and saturation, making it invariant to lighting conditions. This is robust for lighting conditions and highlights the subtle disease-specific color shifts.

2.3.3. Texture features

To capture the texture features, in this article the GLCM is used. It employs a concept of pixel intensity distribution, that includes black, white, and various gray shades. In image for each pixel, homogeneity value is measured and changes are acquired then there has greatest chance to get abnormal region. In a pre-processed image, disease areas offer irregular surface patterns like roughness, lesions, powdery textures, these are captures by texture descriptors. GLCM captures spatial relationships among pixel intensities, representing how pairs of gray levels i and j occur at certain angles and distances. Convert the image into grayscale and execute the GLCMs in multiple orientations. Finally, statistical features are captures from every orientation. Texture features extracted from GLCM include:

i) Contrast: it captures local intensity variations
ii) Correlation: it captures pixel correlation
iii) Energy: it captures textural uniformity
iv) Homogeneity: it captures the closeness of diagonal distribution

2.3.4. Shape features

The diseases cause leaf deformation, curling or holes that alters normal shape geometry. Captured the leaf contour by edge detection like canny and identify external boundaries. Execute the aspect ratio, extends the leaf area or bounding box area and solidity. These capture the geometric distortions because of the disease. Hu moments are the shape descriptors driven from image moments, which are rotation, scale and translation invariant. It encodes the spatial distribution of pixels. $Hu[i]$=function of center moments up to 3rd order. Generally, seven Hu moments are captured per image. Shape descriptors capture the global deformations like structural damage. Hu moments provide a compact and invariant shape representation. Mathematical formula for computing hand crafted features is given in (3). In (3), the F_H represents handcrafted features, the F_C represents captured colored features, the F_T represents captured texture features and F_S represents captured shape features.

$$F_H = [F_C, F_T, F_S] \tag{3}$$

2.3.5. Capturing deep features using EfficientNet

EfficientNet is CNN architecture created through Google Brain Team. This article analyzes network scaling and identifies which optimizing the depth, width, and resolution of the network enhances its performance. For developing the new model, it scales the neural network for creating DL-based algorithms that provide much high effectiveness and accuracy while comparing with previously utilized CNNs. The EfficientNet processed high-scale visual recognition with consistency and accuracy. While comparing to other algorithms like VGGNet, GoogleNet, Xception, ResNet, and InceptionNet, the EfficientNet outperformed these algorithms. EfficientNet utilizes composite scaling algorithm that develops various methods in CNN. Number of layers in the network with respect to network depth. Convolutional layer width is proportional to amount of filters included. Height and width of input image define a resolution of the image. Figure 2 represents an architecture of EfficientNet for feature extraction.

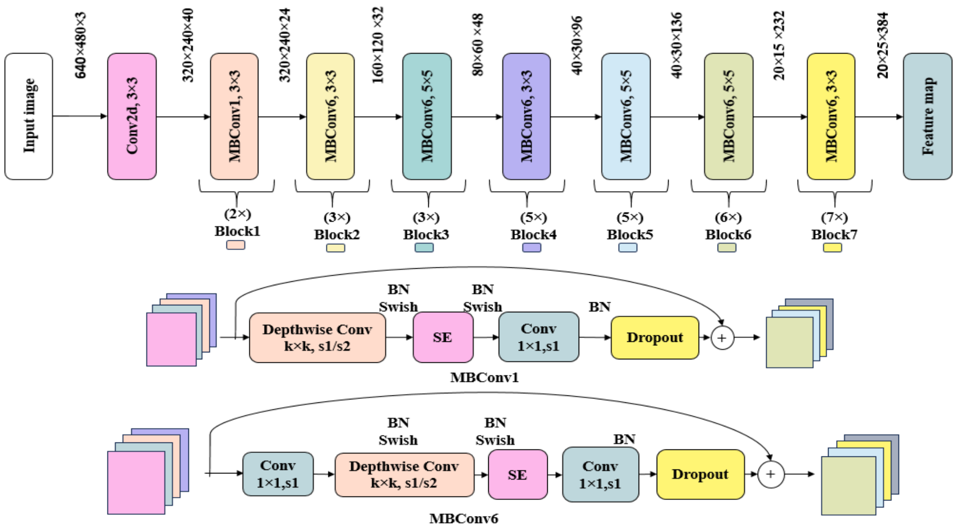

Figure 2. Architecture of EfficientNet for feature extraction

The algorithm captures the characteristics over layers by several convolutional layers with receptive field of 3×3 and a mobile inverted bottleneck Conv. Mathematical from (4) to (8), represents to scale depth, width, and resolution with respect to ϕ. In (4) to (8), the d, w, and r represent depth, width and resolution of the network and α, β, and γ represents constant terms are defined by a grid search hyper parameters tuning method. Coefficient is the user-defined variable that handles whole-scale sources of the method.

$$d = \alpha^{\phi} \tag{4}$$

$$w = \beta^{\phi} \tag{5}$$

$$r = \gamma^{\phi} \tag{6}$$

$$s.t\ \alpha.\beta^2.\gamma^2 \approx 2, \tag{7}$$

$$\alpha \geq 1, \beta \geq 1, \gamma\gamma \geq 1. \tag{8}$$

This method adjusted the depth, width and resolution of the network for optimizing the network's accuracy and memory consumption using available resources. EfficientNet adjusts every dimension by pre-defined group of scaling coefficients, outperforming other DL-based algorithms. The method is released with scaling levels ranging from 0 to 7, where every level represents an increase in accuracy and model parameter size. With recent advancements, EfficientNet provides enhanced ubiquitous connectivity and brings capabilities of DL to various platforms, effectively meeting different application requirements.

i) Spatial pyramid pooling module (SPP): the SPP module is driven from pyramid scene parsing network (PSPNet), the pyramid pooling mitigates the drawbacks of fixed-size requirement for CNN input image. This is employed for extracting features at multiple scales and synthesizes the global data. Figure 3 represents the architecture of DSPP. Pyramid pooling module includes 4 stages such as pyramid pooling, convolution, up-sampling and concatenation process. By pyramid pooling, spatial features on 4 various spatial scales are detected. For improving the capability of non-linear learning of multiple scale features, 1×1 convolution is incorporated for handling feature size and to minimize number of feature channel through $N-th$ amount of channels of an input feature map, where N represents amount of pyramid pooling scales. Convoluted feature maps are inserted by a bilinear filter for matching input size of a feature map. Input feature maps are fused with 4 up-sampled feature maps, so that the global context features are maintained to multiple scale features. Four levels with sizes of 1×1, 2×2, 3×3, and 6×6 are utilized in SPP module.

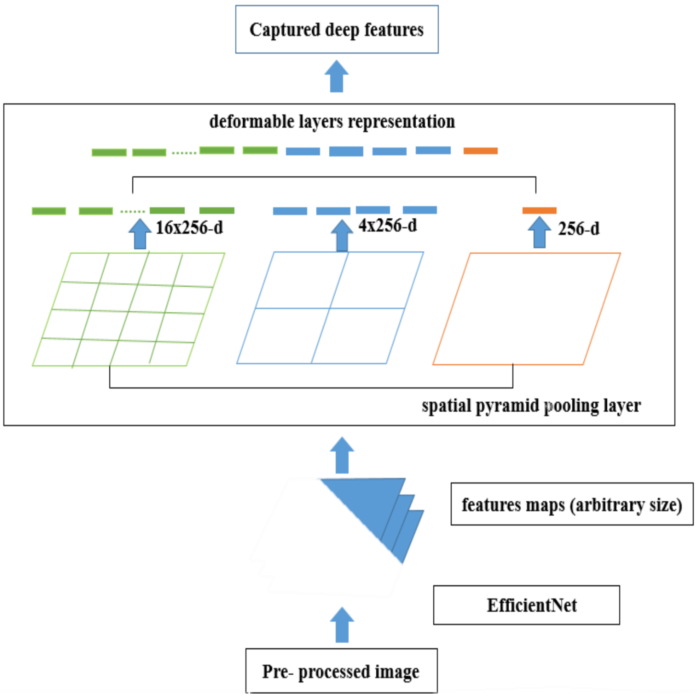

Figure 3. Architecture of DSPP

ii) DSPP modifies the SPP through making the position of pooling layers learnable. the pooling areas deform spatially based on input features, so that they align with semantic or diseases areas. Every pooling layer's sampling position is adjusted through offset, learns dynamically through a lightweight sub-network. The mathematical formula for this process is given in (9).

$$p' = p + \Delta p \quad (9)$$

In (9), the p represents actual pooling position, the Δp represents learned feature and the p' represents final sample position utilized in pooling. These offsets Δp are acquired from an individual convolution layer trained and integrated with the method. DSPP modifies where the features are pooled based on input context. This enables the model to focus effectively on disease-affected areas, even when these regions are distorted and irregular. The extracted handcrafted and deep features are fused together by weighted feature fusion and it is explained.

3.4. Weighted feature fusion

Weight of each feature is considered as evaluation of its feature significance. Sensitivity of evaluated features is defined through employing values within range of [0, 1]. The extracted hand-crafted and deep features are concatenated before being fed into the SVM classifier for classification. The W_i represents weight, which represents capability to classify the different classes of pomegranates and its mathematical expression is, as given in (10), where M represents number of features.

$$W_i = \frac{R_i}{\sum_{i=1}^{M} R_i}, i = 1, 2, \dots, M \tag{10}$$

Before performing weighted feature fusion, that is essential for standardize features to prevent toeing with smaller data values from features with higher data values. This ensures that the calculation outcomes are not distorted because of various dimensions in features. Mathematical formula for normalization of feature value q_i of $ith\ f_{bee}$ is given in (11).

$$q_i' = \frac{q_i - min(q_i)}{max(q_i) - min(q_i)}, i = 1, 2, \dots, M \tag{11}$$

The feature that defines a data to highest extent is acquired by multiplying handcrafted features and deep features through their corresponding weight W_i, then summing the results, with total value normalized to 1. Handcrafted features like color histograms, texture patterns and shape descriptors captured by different traditional algorithms and deep features are automatically learned and encode high-level semantic and spatial data like complex textures and contextual patters captured by DSPP-EfficientNet are fused. The vector of handcrafted features is represented as $F_H \in R^m$ and vector of deep features are represented as $F_D \in R^n$. Here, these two features are combined and represented as F_{fused}. The mathematical formula for weighted feature fusion is given in (12). In (12), the $w_H, w_D \in [0,1]$ represents scalar weights and by adding this value is 1, the $\phi(.)$ represents normalized function.

$$F_{fused} = w_H \cdot \phi(F_H), w_D \cdot \phi(F_D) \tag{12}$$

3.5. Support vector machine

SVM is classification model depended on the principle of structural risk reduction. Main aim of SVM is to classify different samples and increase the spacing among optimally separated hyperplane and the entire training sample. Consider original dataset and it is given as (13).

$$\{(x_i, y_i) | x_i \in R^d, y_i \in \{-1, +1\}, i = 1,2, \dots, n\} \tag{13}$$

In (13), the n represents number of training data samples, the x_i represents input of model, the d is dimension of training sample, the y_i represents sample class, -1 and 1 represent category labels. For linear separable, equation for separation plane is presented as $w \cdot x + b = 0$. A mathematical expression for sample (x_i, y_i) that needs to satisfy is given in (14).

$$y_i[(w \cdot x_i) + b] \geq 1, i = 1,2, \dots, n \tag{14}$$

In (14), the w represents the normal vector plane and b represents constant term. Distance among closest sampling point and separation plane is represented as $1/\|w\|$. Hence, highest spacing of $1/\|w\|$ is equal to highest value of $\|w\|^2$. Separation line defined through w is optimum separation line and samples in separation line $w \cdot x + b = \pm 1$ are known as support vectors. Lagrange optimization algorithm s employed for converting that into a double problem. Mathematical formula for maximisation function is given as (15).

$$max\, W(\alpha) = \sum_{I=1}^{N} \alpha_i - \frac{1}{2} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \tag{15}$$

In (15), the α_i represents the Lagrange multiplier and $\alpha_i \geq 0, i = 1, \dots, n$. This actually identifies optimum solution to quadratic function with constraints and instances respective to non-zero in support vectors. Mathematical formula for obtaining optimal classification function is given in (16).

$$f(x) = sgn\{(w.x) + b^*\} = sgn\{\sum_{i=1}^{n} \alpha_i^* y_i (x_i.x) + b^*\} \tag{16}$$

In (16), the α_i^* represents optimum Lagrange multiplier and the b^* represents classification threshold; these are parameters to determine the optimum hyper-plane partition. Positive or negative function represents class features. For linearly inseparable cases, a slack variable ξ_i is employed, transforming the issue of finding the hyperplane into a quadratic programming problem and its mathematical formula is given as (17).

$$\begin{cases} \phi(w) = min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\varepsilon_i \\ s.t. y_i[(w \cdot x_i) + b] \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \ldots, n \end{cases} \tag{17}$$

In (17), the ξ_i represents positive slack variable which allows misclassification, described deviation number of respective data point x_i from hyperplane. The C represents penalty factor describes degree of punishment to misclassification. This is utilized to control weight among hyperplanes with less spacing in fitness function and guarantees minimum deviation number in data point. Proposed method combined lightweight EfficentNet with DSPP to imprve multi-scle feature extraction. This ensures feature representation, handcrafted features extract color, texture and shape data are integrated with deep features by weighted feature fusion mechanism. This hybrid model uses low-level and high-level descriptors, enables the model to differentiate disease classes. The fused features are then classified by dense neural layer by SoftMax function as in Algorithm 1.

Algorithm 1. DSPP-EfficientNet with weighted feature fusion for pomegranate disease classification
Input: pomegranate fruit disease dataset
Output: predicted disease class for every image, evaluation metrics.
Begin
Pre-processing stage
For every image I in D:
    Resize image to uniform size (224×224)
    Normalize pixel values
    Apply data augmentation
Feature Extraction
For every image I:
    Extract handcrafted features $H = Extract_{Handcrafted}(I)$
    Pass I by EfficientNet to obtain deep features
    Apply DSPP on F to acquire multi-scale ensriched features
Feature Fusion
For every sample
    Fuse features using weighted sum $F_{combined} = \alpha \times H + \beta \times F_{DSPP}$
Classification
    Pass $F_{combined}$ to fully connected layers
    Apply SoftMax activation to obtain class probabilities
    Employ class label with highest probability
Model Training
    Define loss function (CrossEntropyLoss)
    Optimize Using Adam
    Train model over epochs with batch size
Evaluation
    Use test/validation data to compute Accuracy, Precision, Recall, F1-score, Specificity
    Apply K-fold cross-validation for stability.
End

3. EXPERIMENTAL RESULTS

The performance of DSP-EfficientNet with weighted fusion for the SVM method is simulated with Python environment and system configurations are 8 GB RAM, Windows 10 (64-bit) and i5 processor. Accuracy, precision, recall, F1-score, and specificity are metrics considered in this research to validate the performance of developed technique. Error rates like mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute precision error (MAPR), and coefficient of determination (R^2) are considered to evaluate the performance. Mathematical formula for these metrics is given from (18) to (22).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (18)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (19)$$

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (20)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100 \quad (21)$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (22)$$

Figure 4 represents the class-wise results for the pomegranate fruit disease dataset. The five different classes such as alternaria, anthracnose, bacterial_blight, cercospora, and healthy are evaluated to show the performance of the developed technique. This evaluation highlights how each class performs under the proposed DSPP-EfficientNet method with weighted feature fusion.

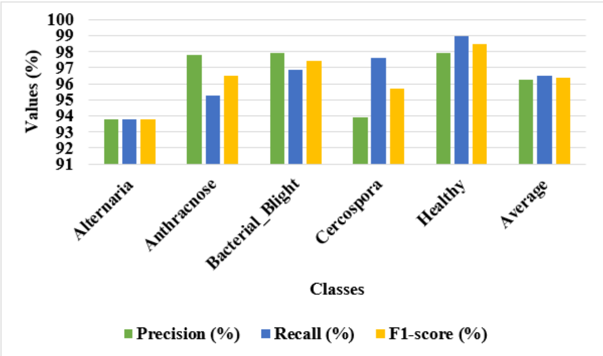

Figure 4. Performance of class wise results for pomegranate fruit disease dataset

In Table 1, the performance of combined handcrafted and deep features with different performance metrics are evaluated. Here, handcrafted features such as color, texture, and shape features are evaluated individually and combined. Then, the deep attributes are evaluated and both handcrafted and deep features are evaluated in combination. The combined deep and handcrafted features obtained 96.66% accuracy, 96.26% precision, 96.50% recall, 96.37% F1-score, and 95.64% specificity on the pomegranate fruit disease dataset.

Table 1. Performance of combined handcrafted and deep features

Features	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
Color features	91.23	90.92	90.51	90.37	90.03
Texture features	91.15	90.79	90.33	90.04	89.84
Shape features	91.34	91.02	90.87	90.66	90.37
Combined handcrafted features	94.78	94.32	94.05	93.76	93.41
Deep features	93.26	93.05	92.85	92.58	92.17
Both handcrafted and deep features	96.66	96.26	96.50	96.37	95.64

Table 2 represents the performance of different feature extraction algorithms with metrics and error rates. The existing feature extraction algorithms like InceptionNet, MobileNet, ResNet, EfficientNet are considered to validate the performance of the developed DSP-EfficientNet with weighted fusion. The DSP-EfficientNet with weighted fusion obtained 96.66% accuracy, 96.26% precision, 96.50% recall, 96.37% F1-score, and 95.64% specificity on the pomegranate fruit disease dataset.

Table 3 represents the performance different classifiers with metrics and error rates. The existing classifiers like extreme gradient boosting (XGBoost), decision tree (DT), RF, and traditional SVM are considered to validate the performance of the developed DSP-EfficientNet with weighted fusion. The DSP-EfficientNet with weighted fusion obtained 96.66% accuracy, 96.26% precision, 96.50% recall, 96.37% F1-score, and 95.64% specificity on the pomegranate fruit disease dataset. Figures 5 and 6 represents a confusion matrix and ROC Curve for proposed method using Pomegranate fruit disease dataset.

Table 2. Performance of different feature extraction algorithms

Methods	Classification performance					Error rate				
	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)	MSE	RMSE	MAE	MAPE	R^2
InceptionNet	94.52	94.31	94.05	94.25	93.75	0.51	0.49	0.64	0.59	0.51
MobileNet	94.78	94.53	94.17	94.42	94.28	0.44	0.42	0.59	0.53	0.56
ResNet	95.02	94.83	94.42	94.55	94.63	0.39	0.34	0.52	0.46	0.60
EfficientNet	95.43	95.21	95.07	95.13	95.05	0.32	0.28	0.45	0.37	0.68
DSP-EfficientNet with weighted fusion	96.66	96.26	96.50	96.37	95.64	0.25	0.22	0.37	0.31	0.76

Table 3. Performance of different classifiers with metrics and errors rates

Methods	Classification performance					Error rate				
	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)	MSE	RMSE	MAE	MAPE	R^2
XGBoost	94.52	94.31	94.05	94.25	93.75	0.51	0.49	0.64	0.59	0.51
DT	94.78	94.53	94.17	94.42	94.28	0.44	0.42	0.59	0.53	0.56
RF	95.02	94.83	94.42	94.55	94.63	0.39	0.34	0.52	0.46	0.60
SVM	95.43	95.21	95.07	95.13	95.05	0.32	0.28	0.45	0.37	0.68
DSP-EfficientNet with weighted fusion for SVM	96.66	96.26	96.50	96.37	95.64	0.25	0.22	0.37	0.31	0.76

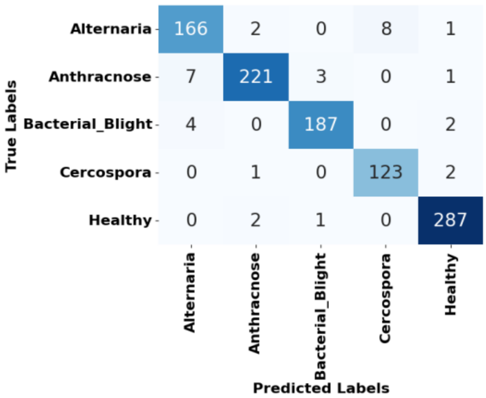

Figure 5. Confusion matrix

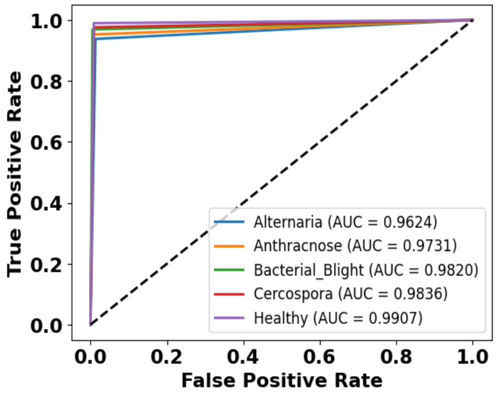

Figure 6. ROC curve

Table 4 presents the results of 5-fold cross-validation to evaluate generalization ability of proposed DSPP-EfficientNet with weighted feature fusion method for pomegranate disease classification. The proposed mode is trained and test across five different data splits, ensures robust model performance. The results from Table 4, shows that the proposed model performed well on individual splits and generalized efficiently across various subsets of data.

Table 4. K-fold cross-validation results for proposed DSPP-EfficientNet with weighted feature fusion model across different metrics

K-Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
K=1	96.30	96.10	96.30	96.20	95.50
K=2	96.80	96.50	96.60	96.55	95.80
K=3	96.90	96.60	96.70	96.65	95.90
K=4	96.50	96.30	96.40	96.35	95.60
K=5	96.60	96.30	96.50	96.40	95.40
Mean	96.66	96.36	96.50	96.43	95.64
Standard deviation	0.20	0.19	0.14	0.17	0.20

Table 5 presents the comprehensive statistical and computation evaluation for DSPP-EfficientNet with weighted feature fusion mode for pomegranate disease classification. Metrics like mean accuracy, confidence interval, standard deviation, p-value from t-test, training time, inference time per image, and memory usage are considered to evaluate the performance of proposed model. The results from the Table 5 shows the model effectiveness, generalization ability and computational efficiency.

Table 5. Statistical significance and computational efficiency of proposed DSPP-EfficientNet with weighted feature fusion model

Metrics	Values
Mean accuracy (%)	96.66
Confidence interval	[96.47, 96.85]
Standard deviation	±0.20
p-value from t-test	0.002
Training time	120 s
Inference time per image	38 s
Memory usage (MB)	235 MB

4.1. Comparative analysis

The performance of the proposed DSP–EfficientNet with weighted fusion for the SVM algorithm is compared with YOLOv6+CBAM [20], GLCM, SIFT+SVM [21], ensemble learning [22], and novel CRNet [23] using the pomegranate fruit disease dataset. The developed DSP-EfficientNet with weighted fusion for the SVM algorithm obtained 96.66% accuracy and 96.26% precision on the pomegranate fruit disease dataset. Table 6 represents a comparative analysis of a proposed technique.

Table 6. Comparative analysis of the proposed technique

Methods	Type of plant	Accuracy (%)	Precision (%)
YOLOv6+CBAM [20]	Tomato disease	NA	92.9
GLCM, SIFT+SVM [21]	Tomato disease	92.3	NA
Ensemble learning [22]	Plant village dataset	92.3	NA
Novel CRNet [23]	Pomegranate growing stage	98	NA
Proposed DSP-EfficientNet with weighted fusion for SVM	Pomegranate fruit disease dataset	96.66	96.26

4.2. Discussion

The experimental results determine the effectiveness of proposed DSPP-EfficientNet with weighted feature fusion model that accurately classifies pomegranate diseases. The incorporation of DSPP ensures model capability to capture multi-scale features, address variations in diseases patterns, lesion size, and background textures identified in fruit images. When integrated with handcrafted features, deep features are extracted by EfficientNet more semantically meaningful, leads to enhanced classification performance across all metrics. By including weighted feature fusion, the proposed model integrated handcrafted features (texture, color) with deep features, efficiently ensures representation space and enhancing class separability. EfficientNet ensures effective and compact feature learning, when DSPP extracts contextual data at multiple receptive fields. The metrics of precision, recall, and F1-score shows the model's capability to differentiate disease classes when maintaining low false positive and false negative rate. The specificity of proposed model shows the ability to correctly identify healthy samples that is essential for real-world agricultural applications. Statistical analysis and k-fold cross validation show the model robustness and stability, with less standard deviation and confidence interval, ensures consistent performance across various data splits.

4. CONCLUSION

This article presents a robust and efficient approach for pomegranate fruit disease detection by combining handcrafted and deep features within a novel hybrid framework. In preprocessing stage uses a median filter and CLAHE, which helps to effectively remove noise and enhances contrast, enabling clear visualization of disease symptoms. Handcrafted features such as color histograms, GLCM-based texture descriptors, and shape features, including Hu moments, are captured as domain-specific visual cues. Deep features are acquired by EfficientNet model integrated with DSPP that adaptively pools feature from disease-relevant regions. These features are fused by a weighted feature fusion strategy that assigns importance to each feature type before being classified using an SVM. Experimental outcomes evaluating on pomegranate fruit disease dataset show that the proposed method significantly outperforms traditional feature extraction and classification approaches and obtained 96.66% accuracy. The proposed hybrid method demonstrates a strong generalization ability, effectively integrating handcrafted and deep features, enabling it to extract different visual patterns over various disease types. The use of a median filter and CLAHE ensures that variations in lighting and noise conditions don't degrade feature quality. Additionally, incorporating DSPP enables the method to adaptively focus on irregular and multi-scale disease areas, enhancing its robustness to spatial variations. Moreover, the weighted feature fusion strategy balances the handcrafted data with automatically learned deep representations, enhancing adaptability for unseen data. In future research, the proposed model can be extended to detect multiple diseases across various plant species. Integration with real-time mobile or drone-based applications could enable in-field disease monitoring. Incorporating temporal data or hyperspectral imaging may enhance detection accuracy further. Additionally, explainable artificial intelligence (XAI) techniques can be explored to improve transparency in model decisions.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Harish Bommenahalli Mallikarjunaiah	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Balaji Prabhu Baluvaneralu Veeranna		✓				✓	✓	✓	✓	✓	✓	✓		

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Kaggle dataset at https://www.kaggle.com/datasets/sujaykapadnis/pomegranate-fruit-diseases-dataset, reference number [26].

REFERENCES

[1] S. Munera, A. R.-Ortega, S. Cubero, N. Aleixos, and J. Blasco, "Automatic detection of pomegranate fruit affected by blackheart disease using X-ray imaging," *LWT - Food Science and Technology*, vol. 215, 2025, doi: 10.1016/j.lwt.2025.117248.

[2] M. Prajwala, P. P. Kumar, S. M. Gopinath, S. Palaiahnakote, M. Basavanna, and D. P. Lopresti, "Domain-independent adaptive histogram-based features for pomegranate fruit and leaf diseases classification," *CAAI Transactions on Intelligent Technology*, vol. 10, pp. 317–336, Nov. 2024, doi: 10.1049/cit2.12390.

[3] S. D. Deb, R. K. Jha, and S. Kumar, "ConvPlant-Net: a convolutional neural network based architecture for leaf disease detection in smart agriculture," in *2023 National Conference on Communications (NCC)*, Feb. 2023, pp. 1–6, doi: 10.1109/NCC56989.2023.10067920.

[4] K. D. Pushya, C. Girish, D. Prakash, and R. Manikantan, "A novel approach for detecting pomegranate leaf pathologies using deep learning," *International Journal of Science and Research Archive*, vol. 13, no. 1, pp. 1206–1218, Sep. 2024, doi: 10.30574/ijsra.2024.13.1.1735.

[5] S. D. Lilda and R. Jayaparvathy, "Enhancing cardiovascular disease classification in ECG spectrograms by using multi-branch CNN," *Computers in Biology and Medicine*, vol. 186, Mar. 2025, doi: 10.1016/j.compbiomed.2025.109737.

[6] M. R. Tonmoy *et al.*, "Attention mechanism-based ultralightweight deep learning method for automated multi-fruit disease recognition system," *Agronomy Journal*, vol. 117, no. 2, 2025, doi: 10.1002/agj2.70035.

[7] R. Karthickmanoj *et al.*, "Revolutionizing agricultural productivity with automated early leaf disease detection system for smart agriculture applications using IoT platform," *Environment, Development and Sustainability*, vol. 26, no. 7, Jul. 2024, doi: 10.1007/s10668-024-05228-7.

[8] A. Maity, J. Sharma, A. Sarkar, and B. B. Basak, "Zinc nutrition improves fruit yield, quality, and reduces bacterial blight disease severity in pomegranate (*Punica granatum* L.)," *Journal of Plant Nutrition*, vol. 46, no. 9, pp. 2060–2076, Sep. 2023, doi: 10.1080/01904167.2022.2118610.

[9] R. K. Lakshmi and N. Savarimuthu, "DPD-DS for plant disease detection based on instance segmentation," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 3145–3155, Aug. 2023, doi: 10.1007/s12652-021-03440-1.

[10] T. V. Kafare, N. Sharma, and A. L. Wanare, "Improved convolutional network with transfer learning and texture feature extractor for plant disease detection," *Journal of Phytopathology*, vol. 173, no. 1, Feb. 2025, doi: 10.1111/jph.70028.

[11] J. Wang *et al.*, "PG-YOLO: An efficient detection algorithm for pomegranate before fruit thinning," *Engineering Applications of Artificial Intelligence*, vol. 134, Aug. 2024, doi: 10.1016/j.engappai.2024.108700.

[12] U. Barman *et al.*, "Vit-SmartAgri: vision transformer and smartphone-based plant disease detection for smart agriculture," *Agronomy*, vol. 14, no. 2, Feb. 2024, doi: 10.3390/agronomy14020327.

[13] G. Sambasivam, G. P. Kanna, M. S. Chauhan, P. Raja, and Y. Kumar, "A hybrid deep learning model approach for automated detection and classification of cassava leaf diseases," *Scientific Reports*, vol. 15, no. 1, Feb. 2025, doi: 10.1038/s41598-025-90646-4.

[14] L. M. Abouelmagd, M. Y. Shams, H. S. Marie, and A. E. Hassanien, "An optimized capsule neural networks for tomato leaf disease classification," *EURASIP Journal on Image and Video Processing*, vol. 2024, no. 1, Jan. 2024, doi: 10.1186/s13640-023-00618-9.

[15] Y. Alhwaiti, M. Khan, M. Asim, M. H. Siddiqi, M. Ishaq, and M. Alruwaili, "Leveraging YOLO deep learning models to enhance plant disease identification," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, doi: 10.1038/s41598-025-92143-0.

[16] D. S. Joseph, P. M. Pawar, and K. Chakradeo, "Real-time plant disease dataset development and detection of plant disease using deep learning," *IEEE Access*, vol. 12, pp. 16310–16333, Jan. 2024, doi: 10.1109/ACCESS.2024.3358333.

[17] K. Taji, A. Sohail, T. Shahzad, B. S. Khan, M. A. Khan, and K. Ouahada, "An ensemble hybrid framework: a comparative analysis of metaheuristic algorithms for ensemble hybrid CNN features for plants disease classification," *IEEE Access*, vol. 12, pp. 61886–61906, Apr. 2024, doi: 10.1109/ACCESS.2024.3389648.

[18] W. Shafik, A. Tufail, C. D. S. Liyanage, and R. A. A. H. M. Apong, "Using transfer learning-based plant disease classification and detection for sustainable agriculture," *BMC Plant Biology*, vol. 24, no. 1, Feb. 2024, doi: 10.1186/s12870-024-04825-y.

[19] M. Iftikhar, I. A. Kandhro, N. Kausar, A. Kehar, M. Uddin, and A. Dandoush, "Plant disease management: A fine-tuned enhanced CNN approach with mobile app integration for early detection and classification," *Artificial Intelligence Review*, vol. 57, no. 7, Jun. 2024, doi: 10.1007/s10462-024-10809-z.

[20] Y. Wang, P. Zhang, and S. Tian, "Tomato leaf disease detection based on attention mechanism and multi-scale feature fusion," *Frontiers in Plant Science*, vol. 15, Apr. 2024, doi: 10.3389/fpls.2024.1382802.

[21] R. Khan, N. U. Din, A. Zaman, and B. Huang, "Automated tomato leaf disease detection using image processing: an SVM-based approach with GLCM and SIFT features," *Journal of Engineering*, vol. 2024, no. 1, Aug. 2024, doi: 10.1155/2024/9918296.

[22] A. Oad *et al.*, "Plant leaf disease detection using ensemble learning and explainable AI," *IEEE Access*, vol. 12, pp. 156038–156049, Oct. 2024, doi: 10.1109/ACCESS.2024.3484574.

[23] A. Naseer, M. Amjad, A. Raza, K. Munir, N. A. Samee, and M. A. Alohali, "A novel transfer learning approach for detection of pomegranates growth stages," *IEEE Access*, vol. 12, pp. 27073–27087, Feb. 2024, doi: 10.1109/ACCESS.2024.3365356.

[24] Y. Hu, G. Liu, Z. Chen, J. Liu, and J. Guo, "Lightweight one-stage maize leaf disease detection model with knowledge distillation," *Agriculture*, vol. 13, no. 9, Aug. 2023, doi: 10.3390/agriculture13091664.

[25] D. Jiang, M. Sun, S. Li, Z. Yang, and L. Cao, "EFS-Former: an efficient network for fruit tree leaf disease segmentation and severity assessment," *Agronomy*, vol. 14, no. 9, Sep. 2024, doi: 10.3390/agronomy14091992.

[26] S. Kapadnis, "Pomegranate fruit disease dataset," *Kaggle.* Accessed: Jan 25, 2025. [Online]. Available: https://www.kaggle.com/datasets/sujaykapadnis/pomegranate-fruit-diseases-dataset

BIOGRAPHIES OF AUTHORS

Harish Bommenahalli Mallikarjunaiah is currently pursuing a doctor of philosophy (Ph.D.) in Artificial Intelligence at Malnad College of Engineering, Hassan, where my research focuses on "Design and development of machine learning-based efficient models for pomegranate fruit cultivation." He holds a master of technology (M.Tech.) in Computer Science and Engineering from Davangere University, Karnataka. His academic interests include machine learning, image processing, and data analytics. His passionate about contributing to innovative and impactful research in my domain. He has presented papers at national and international conferences and have contributed to peer-reviewed journals. His also actively involved in academic and departmental activities, including seminars, workshops, and collaborative projects. His goal is to advance knowledge in my field through rigorous research and to eventually contribute as a faculty member and mentor. Beyond academics, he enjoy painting and creative arts, which helps me maintain a well-rounded and balanced approach to life and work. He can be contacted at email: harishbm137@gmail.com.

Balaji Prabhu Baluvaneralu Veeranna is currently serving as the head of the department and associate professor in the Department of Computer Science and Engineering (AI and ML) at Malnad College of Engineering, Hassan. He holds a Ph.D. in Computer Science from Visvesvaraya Technological University and completed his postdoctoral research at the Indian Institute of Science (IISc), Bangalore. With over 10 years of teaching and research experience, his areas of expertise include artificial intelligence, machine learning, big data analytics, remote sensing, and cloud computing. He has authored more than 18 research papers in SCI and Scopus-indexed journals, published patents, and co-authored a book on IoT applications. He has also received multiple awards, including the best young researcher award and excellence in research award. His current research interests focus on the application of AI in remote sensing, smart agriculture, and IoT-enabled systems. He can be contacted at email: bvb@mcehassan.ac.in.