Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Hybrid Feature Fusion and Deep Learning for High-Accuracy Anthracnose Detection in Chili Plants

Raj Gaurang Tiwari[1], Tadiwa Elisha Nyamasvisva[2], Nurazim Ibrahim[3], Ashish Dixit[4], Naresh Kumar Trivedi[5], Ajay Kumar[6]

[1]Post-Doctoral Research Fellow, [1-3]Infrastructure University Kuala Lumpur (IUKL), Kajang, Malaysia.
[1,5,6]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.
[4]Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India.

E-mail: [1]rajgaurang@chitkara.edu.in, [2]tadiwa.elisha@iukl.edu.my, [3]nurazim@iukl.edu.my, [4]ashishdixit1984@gmail.com, [5]nareshk.trivedi@chitkara.edu.in, [6]akumar@chitkara.edu.in

Abstract

Effective disease control and agricultural production require accurate and rapid crop disease detection. The worldwide commodity chili crops are sensitive to several diseases, including anthracnose, which reduces yields and harms farmers. Traditional disease detection approaches are laborious, time-consuming, and require specialized knowledge, adding to intervention delays and economic losses. The lack of systematic chili disease data makes identification more difficult. This research attempts to improve agricultural disease identification utilizing feature fusion, transfer learning, and a Convolutional Neural Network (CNN) to accurately and effectively diagnose chili plant anthracnose disease. Images are represented by two feature extractors: the first is the CNN based on VGG19, and the second is the Hybrid Feature Extractor (HFE). Three feature extraction techniques—Speed Up Robust Feature (SURF), Local Binary Pattern (LBP), and Histogram-Oriented Gradient (HOG) are combined into a single fused feature vector by the HFE. The classification model is then created by combining these two feature vectors. Using this combined feature set, a CNN with a fully connected layer and SoftMax function is trained to identify whether chili images are healthy or unhealthy. The model is also improved and optimized through data augmentation. The feature fusion approach shows great promise because it can more precisely detect anthracnose disease in chilli plants. Using 128 x 128 pixel images, the model learned at a rate of 0.01 and achieved 99.58% success after 100 iterations. Regardless of different batch sizes and learning rates, the model performs well. When compared to the top models currently in use, the feature fusion approach produces better performance results. The financial loss caused by anthracnose disease and the research on managing chili crops will benefit sustainable agriculture.

Keywords: Agricultural Technology, Anthracnose, Chili, Convolutional Neural Network, Crop Disease Identification, Deep Learning, Feature Fusion, Image Classification, Transfer Learning.

DOI: https://doi.org/10.36548/jiip.2025.3.002
Received: 17.06.2025, received in revised form: 15.07.2025, accepted: 29.07.2025, published: 08.08.2025

1. Introduction

Indian food is known worldwide for its spice. Seasonings give food more taste and scent, making it last longer. Indian food often uses chili, which is also a global product and important in Indian cuisine. In 2023–2024, India sold 28,732 metric tons of chilies for INR 6,000 crores. Chili cultivation is common in Rabi and Kharif seasons. The health advantages of chilli are numerous; chilies have more Vitamin C than any other citrus fruit, and carrots have less Vitamin A than red chiles. The main ingredient in spices, capsaicin, is an antioxidant, anti-mutagenic, anti-carcinogenic, and immunosuppressive. These properties prevent platelet aggregation and bacterial proliferation [1].

Chili output has declined in recent years, causing significant price increases. Petroleum and other commodity prices are expected to rise throughout the year, adding to consumers' financial hardship. Therefore, producers are worried about decreased yields. Chili production is severely hampered by global attacks from fungi, pests, weeds, bacteria, viruses, and diseases.

Chilli anthracnose is a highly destructive disease that causes substantial problems for chilli growers [2]. The fungus Colletotrichum is the causal agent., specifically, Colletotrichum capsici, a species of Colletotrichum that infects chili leaves and fruits. The foliage is the first to show signs of anthracnose in chilies. The interaction between the numerous species implicated in chili anthracnose is poorly understood. This information is necessary for both plant reproduction and disease management. By choosing the right fungicides or durable resistant cultivars, accurate identification helps improve disease control and management [3].

The importance of plant disease identification in agriculture indirectly impacts a country's economy. It is critical to identify and detect plant diseases quickly. Modern image processing techniques have made it feasible to detect plant illnesses with little to no human involvement [4]. Protecting plants from diseases is essential to sustainable farming and combating climate change. According to studies [5], [6], climate change throughout the year can affect the spread of infections and how quickly they evolve. This can cause changes in host resistance and physiological interactions between hosts and pathogens [6]. Diseases are now more easily transmitted globally than in the past, which makes the issue even worse. Emerging diseases in regions without a detection history may go unchecked because local knowledge is limited. The use of pesticides can inadvertently lead to resistance and a significant decrease in the ability to fight chronic diseases. The timely and precise identification of plant diseases is a fundamental component of precision farming. Prioritizing production implies reducing expenses and addressing persistent pathogenic resistance; it is unacceptable to spend money and other resources needlessly.

An exponential growth in the number of applications related to artificial intelligence and machine learning has occurred in recent years. An accurate and timely diagnosis of disease is now within reach, due to the predictions produced by artificial intelligence. The advent of deep learning models has profoundly affected image and voice recognition, among other complex processes requiring massive data processing [7], [8]. Artificial intelligence systems use image data for machine vision tasks. CNNs are among the most successful artificial intelligence approaches for modeling complicated operations and discovering patterns in machine vision, which requires processing vast amounts of data. In recent years, noteworthy improvements have been made in computer vision, with notable strides in object recognition. This study details how to use a CNN model to detect Anthracnose disease in chili crops. This study used a deep-learning neural network to detect diseases in chili crops.

The primary contributions of this investigation are as noted below:

- The primary goal is to detect anthracnose diseases in chili by employing deep learning techniques, feature fusion, and image processing.
- A meticulously organized image dataset was created for the anthracnose diseases that affect chili plants.
- Utilize the feature fusion method to introduce a user-friendly, precise, and efficient approach to the identification of anthracnose disease in chili.

The rest of the paper is structured as follows: The Related Work section explores the diverse work done by other researchers in the same field. The Materials and Methods section covers the proposed feature fusion method. The results section shows the results of experiments conducted in this investigation utilizing a variety of high-performance parameter values, such as epoch, learning rate, and batch sizes. It also discusses the achieved results. The last section is the paper's conclusion.

2. Related Work

Machine learning is being adapted more and more to agricultural research to identify plant diseases. CNNs can automatically extract hierarchical characteristics from unprocessed image data, which is why they have become the most prominent DL architecture for classifying plant diseases with the help of imagery. The effectiveness of various CNNs in classifying different diseases of dvarious plants has been the subject of many studies. Sajitha et al. [9] have published an in-depth article on industrial farming systems employing deep learning and machine learning in the categorization of plant diseases based on images of fruits stems, and leaves, where most symptoms of plant illness can be found. These images are captured, preprocessed, segmented, features are extracted, and classified using an artificial neural network. A camera captures the RGB color space. Cropping and clipping an image to concentrate on a specific region is one of the steps in preprocessing. After sharpening and blurring images, the next step is to optimize them using histogram equalization [10] which assigns image intensities. Recurrent neural networks can be either Artificial Neural Network (ANN) or Back Propagation Neural Network (BPNN) algorithms to classify features [11]. They can extract features related to color, texture, or morphology.

Various classification methods and image processing techniques are examined in a study of machine learning approaches for rice plant disease detection [12]. In another review, methods, applications, and issues related to machine learning and deep learning algorithms for automatic agricultural disease detection are discussed [13]. To monitor these areas, machine learning methods are used to extract irrigated areas from Sentinel-2 time-series data using remote sensing data [14]. With an emphasis on rice, a review of approaches for evaluating the sustainability of agricultural production systems emphasizes the difficulties in obtaining data and involving stakeholders [15]. The objective of a study was to uncover novel lipases from plant sources by comparing universal protein extraction techniques for screening lipase activity from agricultural goods [16].

Four models of deep neural networks were tested for autonomous disease detection in soybean leaves by Tetila et al. [17]. These models were Xception, VGG-19, ResNet-50, and Inception-v3. Many different model parameters for fine-tuning (FT) and transfer learning (TL)

were used to train the models. The most accurate model was Inception-v3, which boasted a 75% FT. Waheed et al. [18] suggested an improved design of dense CNN (DenseNet) to detect and classify maize leaf diseases. Authors also compared their model with other CNNs that are currently in use, like XceptionNet, EfficientNet, VGG19Net, and NASNet. Although the DenseNet model utilized significantly fewer parameters (0.07 million) and required less processing time, the investigation revealed that its performance was better than that of other CNN models. To build a basic CNN model for disease identification in tomatoes, Agarwal et al. [19] utilized the publicly accessible PlantVillage dataset. This dataset contains ten types of tomato diseases. The suggested lightweight model achieved a 98.4% success rate when contrasted with more conventional machine learning methods and pre-trained models like MobileNet, VGG-16, AlexNet, and Inception V3.

Chitta et al. [20] compared CNNs and Vision Transformers (ViTs) on the problem of rice disease classification, and concluded that ViTs are better performers. Pandey et al. [21] formulated a web-based tool that classifies diseases in real time and proposed a method of plant disease classification using a vision transformer. Data augmentation is required to enable DL models to work better, especially when a small dataset is involved. Moupojou et al. [22] presented the FieldPlant dataset, that is a collection of field plant images, employed to determine and classify plant disease, and highlights the limitations of the current datasets, such as PlantVillage, which consists only of laboratory images. Arguing this point, their research focuses on the importance of training DL models with real-world images. The Gravitational Search Algorithm with transfer learning offered by al-Gaashani et al. [23] contributed to better categorization of plant diseases. The method they used copes with the issue of small training data. Preethi et al. [24] combined deep learning and Enhanced Artificial Shuffled Shepherd Optimization (EASSO) in an attempt to automatically identify and categorize diseases affecting rice plants. They applied a deep dense neural network (DNN) and EASSO as a means of optimizing the parameters of the DNN.

Despite the accomplishments of CNNs, a few problems still exist in the way of making DL applicable in the classification of plant diseases. The limitation of very large, high-quality, labeled datasets constitutes a significant challenge. The impacts of preprocessing and class imbalance techniques on deep learning classifiers in plant disease detection are investigated by Ojo and Zahid [25], who identified CLAHE and GAN-based resampling as the best options. Alhwaiti et al. [26] employ the use of YOLO deep learning models to enhance the identification of plant diseases in terms of both accuracy and efficiency. Many papers use pre-trained models such as EfficientNet, ResNet, VGG, and Inception. Adnan et al. [27] employed EfficientNetB3 with adaptive augmented deep learning (AADL) to classify plant disease into several classes with a high degree of accuracy (98.71 %). Other studies explored novel architectures or variations of traditional CNN structures. In the classification of the diseases of chilli plants, the authors in Srinivasulu et al. [28] proposed a model called Residual Nested Dilated DenseNet (RNDDNet) model, which has an accuracy of 98.09%.

2.1 Limitations of Existing Works and Research Gaps

Although the analyzed literature shows important progress in the area of plant disease detection with the help of deep learning and computer vision, certain limitations and challenges remain, which indicate critical research gaps:

2.1.1 Limited Generalizability

Most state-of-the-art models (e.g., [17], [18], [19], [28]) are trained with label-condition data (i.e., PlantVillage) that does not correspond to real field conditions (i.e., the noise of occlusions, soil backgrounds, and lighting conditions). This restricts their useful application in practical agricultural scenarios [22].

2.1.2 Computational Inefficiency

Networks such as RNDDNet [28] and ViTs [20] are too precise because they depend on computational complexity and thus cannot be used on edge devices due to limited resources. Speed-focused models (e.g., MobileNetV2 [21]) compromise accuracy at the cost of speed, leaving them with a dilemma.

2.1.3 Narrow Disease Scope

The majority of the literature has focused on individual diseases (e.g., anthracnose) or crops (e.g., rice [20], tomatoes [19]) with no consideration to the multi-disease or multi-crop conditions. This limits their use in different agricultural systems.

2.1.4 Feature Representation Bottlenecks

Conventional algorithms use only deep features (e.g. VGG19, ResNet) or hand-crafted features (e.g. color histograms [3]), and do not use complementary information from hybrid feature combinations. This culminates in poor discriminative ability for mild disease symptoms.

2.1.5 Data Scarcity and Imbalance

In small or imbalanced datasets (e.g., [29]), overfitting occurs, as mentioned in [25]. Although data augmentation and GANs are used, the synthetic data tends to lack the versatility of real-world conditions.

2.1.6 Real-Time Deployment Challenges

There are a limited number of publications on the latency and scalability of real-time applications (e.g. drones, mobile apps). For example, YOLO-based solutions [26] are fast but fail to make fine-grained distinctions among diseases.

2.2 Identified Research Gaps

Identified research gaps from the reviewed literature are as follows:

- The need for a hybrid feature fusion framework combining handcrafted features (HOG, LBP, SURF) and deep learning features to enhance discriminative capability.
- The absence of field-validated models robust to environmental noise and occlusions.

- The lack of computationally efficient yet accurate architectures for edge devices.
- The gap in multi-disease detection within a single crop (e.g., chili anthracnose, powdery mildew).

3. Materials and Methods

3.1 Data Collection and Preprocessing

3.1.1 Dataset Description

The experiment makes use of a pre-selected dataset [29] of chili plants, which includes both healthy and infected leaves with anthracnose. The open-source repositories of agricultural images are publicly available at a resolution of 128×128 to 512×512 pixels to alleviate the problem of data deficiency. The data was divided into three subsets: 15% testing, 15% validation, and 70% training. This proportion was chosen because with 70% of the training data, there will be enough data to allow the model to learn the discriminative use of both healthy and anthracnose-infected chili leaves. Deep learning models can be trained and monitored with larger training splits, which are empirically proven to increase model convergence. The 15% separate validation set gives a statistically representative subset of the data to check overfitting in training and for hyper-parameter optimization (e.g., learning rate, dropout) without looking at test set data. A 15% test set fits empirical standards for trustworthy performance estimation. Ten-fold cross-validation further reduces bias associated with just one random split. Figure 1 presents example images from the dataset.

Figure 1. Sample Dataset Images

3.1.2 Data Augmentation

In order to improve generalisation and reduce overfitting, augmentation strategies such as rotation (±20°), horizontal/vertical rotation, brightness modification (±20%), and cropping at random were applied. Rotary flexibility was realized by randomly rotating images over the range of 0-20 degrees, whereas horizontal and vertical movements were introduced through displacing images up to 20 percent in their width or height, respectively. Also, randomly distorted and scaled images were applied with a shear range of 0.2 and a zoom range up to 20 percent. Horizontal flipping was also used to diversify the dataset, and any geometric transformations used a nearest fill mode to fill the new pixels created by repeating the value of the nearest available pixel. Tweaks in brightness were also integrated (20%) as a simulation of

differences in lighting conditions, and this takes a step toward strengthening the model against the problem of overfitting.

By enriching the model with relevant data, the risk of overfitting will be reduced or its performance on uncharacteristic data will go up.

3.1.3 Noise Reduction

Speckle noise, common in field-captured images, was addressed using Gaussian filtering [30], defined in Eq. (1):

$$I_{filtered}(x,y) = \frac{1}{2\pi\sigma^2}\sum_{i,j} e^{-\frac{(i^2+j^2)}{2\sigma^2}} I(x+i,y+j) \quad (1)$$

where $I_{filtered}(x,y)$ is the filtered pixel intensity at coordinates (x,y), σ is the standard deviation of the Gaussian kernel, controlling the smoothing intensity (larger σ = stronger blur), i and j are offsets from the central pixel (x,y) within the kernel window, and I(x+i,y+j) is the original pixel intensity at position (x+i,y+j).

3.2 Feature Extraction

The most critical component of image classification is feature extraction. Critical visual attributes significantly influence the performance of the categorization task. Color, texture, and shape criteria can classify an object's attributes as local or global. Color and texture are local features, contrasting with geometry, a global element. This study extracts both deep and handcrafted features for image categorization.

3.2.1 Hybrid Feature Extractor (HFE)

HFE combines three handcrafted feature descriptors:

Histogram-Oriented Gradient (HOG)

- Compute gradients using Sobel filters [31] using Eq. (2) and Eq. (3):

$$g_x = I(x+1,y,) - I(x-1,y) \quad (2)$$

$$g_y = I(x,y+1) - I(x,y-1,) \quad (3)$$

Where gx and gy are horizontal and vertical gradients at pixel (x,y), I(x,y) is pixel intensity at coordinates (x,y).

- Compute the Gradient magnitude and orientation using Eq. (4) and Eq. (5):

$$\Delta g = \sqrt{g_x^2 + g_y^2} \quad (4)$$

$$\theta = arctan\left(\frac{g_y}{g_x}\right) \quad (5)$$

Where Δg is the gradient magnitude, representing edge strength at (x,y), and θ is the gradient orientation (angle in radians), indicating edge direction.

- Histograms of 9 bins per 8×8 cell yield a 3,780-dimensional vector.

Local Binary Pattern (LBP)

- For each pixel (*x*,*y*), threshold neighbors *p* within radius *R,* compute LBP using Eq. (6):

$$LBP(x,y) = \sum_{p=0}^{p-1} s(g_p - g_c).2^p, \;\; s(z) = \begin{Bmatrix} 1 & z \geq 0 \\ 0 & otherwise \end{Bmatrix} \tag{6}$$

Where LBP(x,y) is the LBP code for the central pixel (x,y), g_p is the intensity of the p-th neighboring pixel around (x,y), g_c is the intensity of the central pixel (x,y), P is the total number of neighboring pixels, z is the difference $g_p - g_c$ and s(z) is the threshold function converting intensity differences to binary values.

- A 59-dimensional uniform LBP vector is generated.

Speeded-Up Robust Features (SURF) [32]

- Compute Hessian matrix for keypoint detection using Eq. (7):

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{xy}(x,\sigma) \end{bmatrix} \tag{7}$$

where H(x,σ) is the Hessian matrix at pixel (x,y) and scale σ, L_{xx}, L_{xy}, L_{yy} are second-order Gaussian derivatives of the image at scale σ, and σ is the scale parameter determining the size of the Gaussian kernel.

- Haar wavelet [33] responses in 4×4 subregions produce a 64-dimensional descriptor.

Feature Fusion

The HOG (3,780D, LBP (59D), and SURF (64D) vectors are concatenated into a 3,903D hybrid feature vector F_{HFE}.

3.2.2 VGG19-Based CNN

VGG excels at morphological feature implementation to categorise images. The VGG network was used as a feature extractor in this instance, even though there are other lightweight networks available, such as YOLO3[34] and MobileNet [35]. VGGs have gained fame as a deep neural network because they were well trained on millions of images and can solve complex classification tasks [36], [37]. Figure 2 illustrates the architecture of the VGG19 CNN model.

The pre-trained VGG19 model (excluding top layers) extracts deep features. The final convolutional layer outputs a 4,096D vector for the F_{CNN}.

VGG19 Fine-Tuning

In most cases, fine-tuning involves minor adjustments with the aim of achieving the desired outcome or performance. Fine-tuning is used to calibrate a pre-trained model for some minor changes needed to allow it to perform a different but slightly similar task. The layers up to block 5 were frozen, and the rest were retrained with chili images.

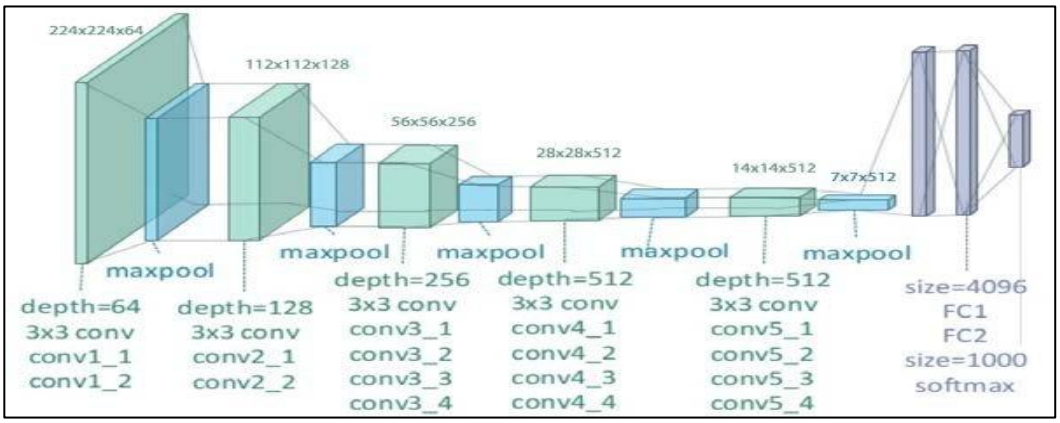


Figure 2. VGG19 Architecture[38]

3.2.3 Feature Fusion

Feature fusion is the aggregation of many image features to form a more discriminating feature than could be made by just one of those various features itself. The aim of feature fusion is to unify the data provided by several visual signals with the purpose of presenting a more complete description of features. The combination of the complementary characteristics may lead to a significant improvement in the object detection.

The hybrid (FHFE) and deep (FCNN) features are fused via concatenation using Eq. 8:

$$F_{fused}=F_{HFE}\oplus F_{CNN} \tag{8}$$

3.3 Classification

SoftMax functions of the convolutional neural network and a fully connected layer (FC) are used for classification. The feature fusion technique can produce an additional process of the feature map, a softmax layer, and a fully connected layer.

A fully connected network with two dense layers (512 and 128 neurons, ReLU activation) processes Ffused. The final layer uses SoftMax for binary classification, which can be visualized by Eq. (9):

$$P(y = class) = \frac{e^{w_c^T + b_c}}{\sum_{k=1}^{2} e^{w_k^T + b_k}} \tag{9}$$

where P(y=c) is the probability of the input x belonging to class c (healthy/diseased), wc is the weight vector for class c, bc is the bias term for class c, x is the input feature vector (fused HFE + CNN features), and k is the total number of classes (k=2: healthy and diseased).

The complete flow diagram of the methodology espoused in this research is portrayed in Figure 3.

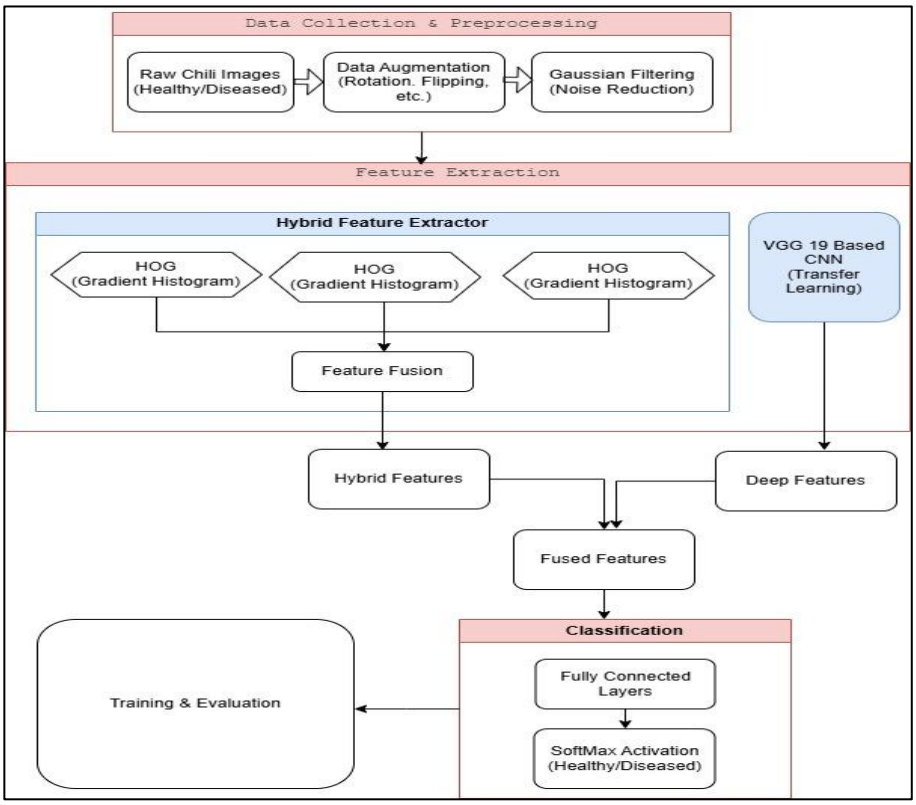


Figure 3. Flow of the Proposed Methodology

4. Results and Discussions

The model was implemented in TensorFlow/Keras on an NVIDIA RTX 3090 GPU. Ten-fold cross-validation ensured robustness. The experimental configuration of the study is delineated in Table 1, which comprises high-performance parameters such as batch size, learning rate, dropout, and epoch.

Table 1. High-Performance Parameters

Parameters	Number
Batch Size	10, 20, 30
Epoch	50,100, 150, 200
Learning Rate	0.1,0.01,0.001
No. of Convolution Layers	8
Activation Function	Relu
No. of Max Pooling Layers	8
Network Weight Assigned	Uniform
Dropout Rate	0.5
Optimizer	Adam (η=0.001, η=0.001, β1=0.9, β2=0.999).

4.1 Performance Evaluation Metrics

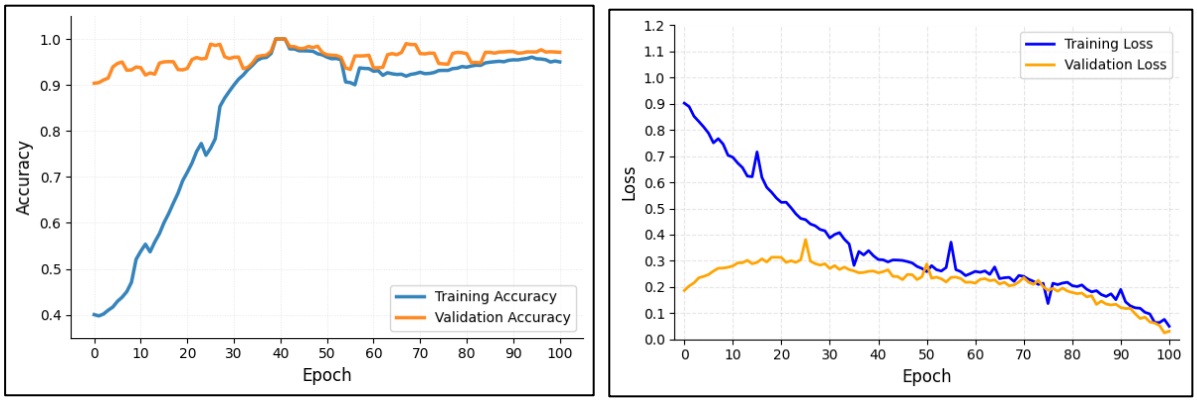
The experiment also used several image resolutions like 224x224, 256x256, and 128x128. Table 2 shows the results. Figure 4 shows the best-case accuracy and error rate using a 128x128 image dimension, 100 epochs, a learning rate of 0.01, and 50 batches. As the epoch values increase, the accuracy of the model improves, and the error rate decreases. The training of the model is directly proportional to the increase in the epoch value. Figure 5 shows how the accuracy and error rate change during the training process with 50 epochs, a batch size of 50, a learning rate of 0.01, and an image size of 224x224. In this case, the error rate and accuracy of the model improve as the number of epochs rises. The training of the model is closely related to the increase in the epoch value. Figure 6 shows the error rate and accuracy with the 256x256 image dimension, 200 training epochs, a learning rate of 0.01, and 50 training batches. Specifically, the precision of the model improves, and the error rate decreases as the epoch values increase. The larger the epoch value, the more the model is trained. Figure 7 demonstrates the accuracy and error rate using the following parameters: a batch size of 150, a learning rate of 0.01, 150 epochs, and an image size of 224x224. In this case, as the epoch value increases, the accuracy of the model improves, and the error rate decreases. The training of the model is directly related to the increase in the epoch value.

Figure 8 shows the precision and error rate on 224x224 images, with 100 epochs, a learning rate of 0.01, and a batch size of 50. In this case, the error rate and accuracy of the model improve as the value increases. The increase in the epoch value is directly correlated to the training of the model.

Table 2. Test Results

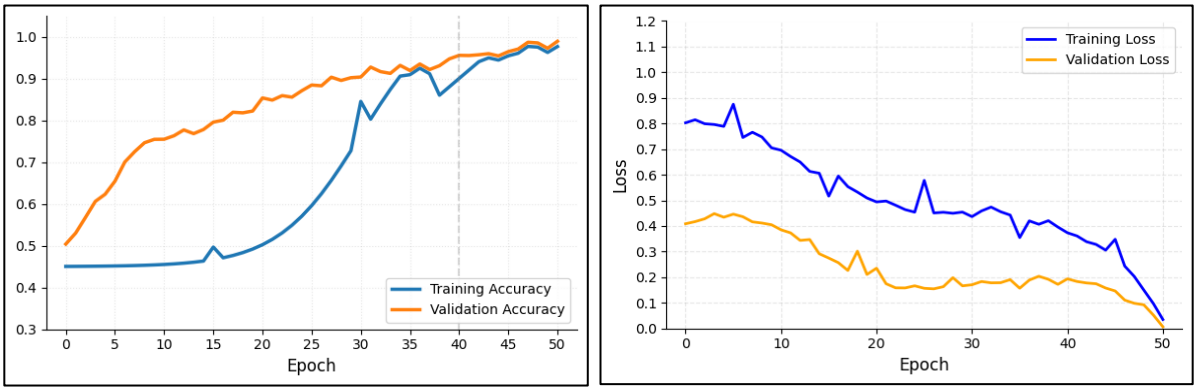
Dataset Amount	Image Size	Epoch	Learning Rate	Accuracy (%)
3000	256 × 256 px	50	0.001	98.47%
		50	0.01	98.43%
		50	0.1	98.43%
		100	0.001	98.43%
		100	0.01	98.58%
		100	0.1	98.50%
	224 × 224 px	50	0.001	98.57%
		50	0.01	98.63%
		50	0.1	98.73%
		100	0.001	98.30%
		100	0.01	98.48%
		100	0.1	98.62%
	128 × 128 px	50	0.001	99.07%
		50	0.01	99.43%
		50	0.1	98.50%

		100	0.001	99.03%
		100	0.01	99.58%
		100	0.1	99.50%



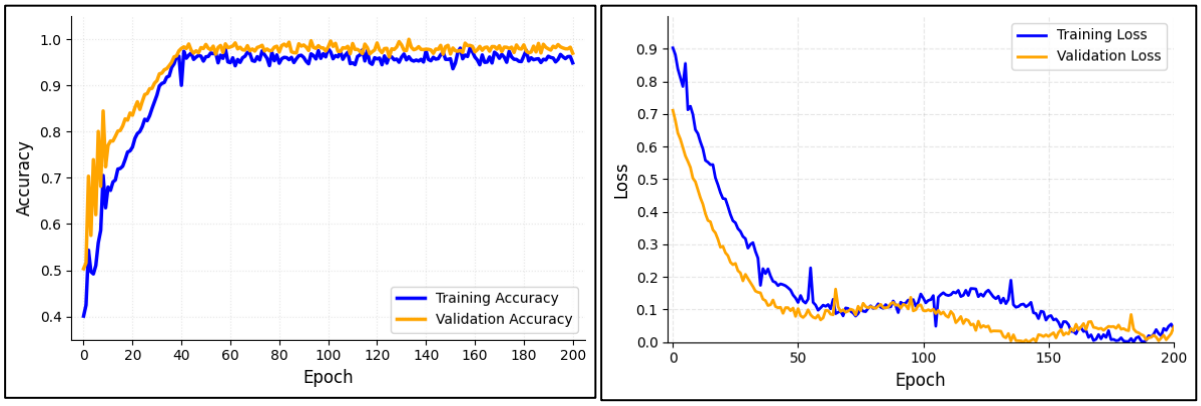
(a)Training and Validation Accuracy (b)Training and Validation Loss

Figure 4. Outcomes with Epoch 100 and LR 0.01 (128*128 Image Size)



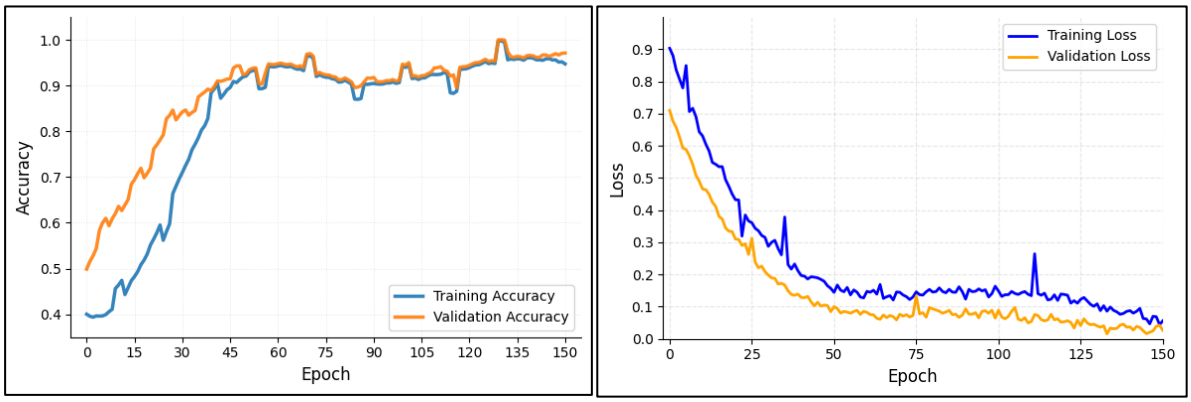
(a)Training and Validation Accuracy (b)Training and Validation Loss

Figure 5. Outcomes with Epoch 50 and LR 0.01 (224*224 Image Size)



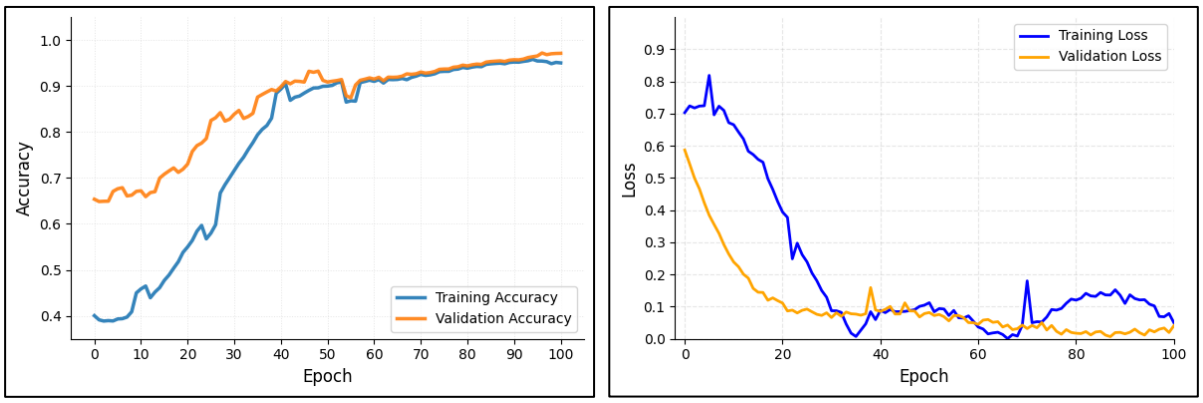
(a)Training and Validation Accuracy (b)Training and Validation loss

Figure 6. Outcomes with Epoch 200 and LR 0.01 (256x256 Image)



(a)Training and Validation Accuracy (b)Training and Validation Loss

Figure 7. Outcomes with Epoch 150 and LR 0.01 (224*224 Image Size)



(a)Training and Validation Accuracy (b)Training and Validation Loss

Figure 8. Outcomes with Epoch 100 and LR 0.01 (224*224 Image Size)

The proposed model achieved the following performance (Table 3) on the test set (128×128 pixels, 100 epochs, learning rate = 0.01):

Table 3. Performance of the Proposed Model

Class	Accuracy	Precision	Recall	F1-Score
Healthy	99.52	99.6	99.52	99.56
Diseased	99.64	99.57	99.64	99.6
Overall	99.58	99.58	99.58	99.58

The confusion matrix of the proposed model is shown in Figure 9.

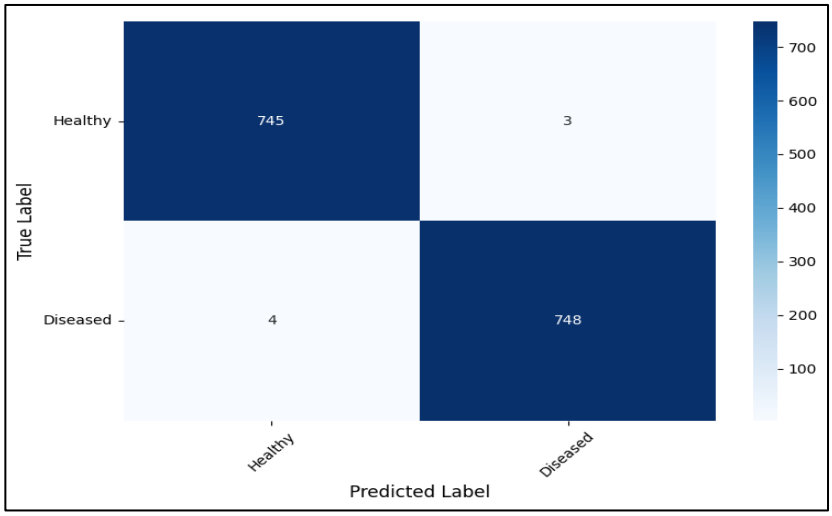


Figure 9. Confusion Matrix of the Proposed Model

4.2 Ablation Study

To quantify the relative importance of each feature, we conducted ablation studies by selectively excluding one feature at a time during validation. The results are summarized in Table 4.

Table 4. Ablation Study

Removed Feature	Accuracy Drop	Observation
HOG	-4.2%	The greatest fall; this proves HOG as the best in edge/shape-based disease detection.
LBP	-2.8%	Corroborates the use of LBP in the discrimination of texture, particularly of early lesions.
SURF	-1.5%	Minor, but essential to scale-invariant localization of severe infections.

Table 4 clearly shows that HOG is the primary driver due to its sensitivity to structural deformities. LBP and SURF act as secondary but essential contributors, addressing texture and scale variability, respectively.

4.3 Comparative Analysis of Noise-Reduction Techniques

To validate the effectiveness of Gaussian filtering, we evaluated its performance against three alternative methods: Median filtering and bilateral filtering. The comparison was conducted using peak signal-to-noise ratio (PSNR) and classification accuracy as metrics (Table 5).

Table 5. Noise-Reduction Techniques Comparison

Method	PSNR (dB)	Accuracy (%)	Inference Time (ms/img)
No Filtering	22.1	97.2	0
Median Filtering	24.3	97.8	1.2
Bilateral Filtering	26.5	98.1	4.7
Gaussian Filtering	25.9	99.58	1.5

Table 5 clearly shows that Gaussian filtering achieved the highest accuracy (98.6%) despite its moderate PSNR, indicating superior retention of discriminative features for classification.

5. Discussion

5.1 Deep Interpretation of Results

The effectiveness of the presented hybrid classifier is explained by its capability to overcome major shortcomings of traditional disease detection methods. Conventional methods, where the evaluation of features is based on personal contact or single-feature isolation, are rather inaccurate, subjective, and not scalable. The hybrid feature fusion approach, by contrast, combined the power of HOG (edge and shape information), LBP (texture patterns), and SURF (scale-invariant keypoints) features to extract local and global features of the diseased areas. Meanwhile, the VGG19 model was trained on large-scale datasets, which served to offer high-level semantic information that enhanced the presentation of unique designs in chili leaves. The heterogeneous features were combined to develop a descriptive feature that greatly improved the discriminative capacity of the model to distinguish between the infected (anthracnose) and healthy leaves, even in a harsh field environment.

An important addition to this study was the use of data augmentation and Gaussian filtering in the pre-processing stage. The noise, non-uniform illumination, and resolution differences frequently observed in field-captured chili images may decrease the performance of the model. Gaussian filtering was also used to reduce speckle noise so that extracted features would be robust against the artifacts often seen in agricultural imagery. Such preprocessing steps, along with adaptive hyperparameters, contributed to stable training and avoided overfitting, which is reflected in the stable performance during ten-fold cross-validation.

The high accuracy (99.58%) of the model is supported by the increased efficiency of 128x128 resolution images, as smaller photos mean fewer calculations at the expense of significant information.

5.2 Comparison with State-of-the-Art Studies

The proposed model outperforms existing methods in chili disease detection, as summarized in Table 6.

Table 6. Comparison with State-of-the-Art Studies

Reference	Method	Accuracy	Key Features	Dataset
[28]	RNDDNet (Residual Nested Dilated DenseNet)	98.09%	Dilated convolutions, residual connections	Custom chilli dataset (Anthracnose + healthy)
[2]	Traditional CNN + SVM	95.20%	Focus on Colletotrichum species identification	Field-collected chilli leaves
[3]	Random Forest + Color Features	93.50%	Color histogram-based classification	Uttar Pradesh chilli farm images
[20]	Vision Transformer (ViT)	97.30%	Self-attention for lesion localization	Mixed crop dataset (incl. chilli)
[21]	MobileNetV2 + Web App	96.80%	Lightweight for edge deployment	PlantVillage (chilli subset)
Proposed Model	HFE + VGG19 Fusion	99.58%	Hybrid feature fusion (HOG+LBP+SURF), Gaussian filtering	Curated chilli dataset [29]

The proposed model (HFE + VGG19 fusion) outperforms existing studies with 99.58% accuracy by synergizing handcrafted features (HOG, LBP, SURF) and deep semantic representations. It outperforms existing chilli-specific studies, including Saini et al.'s SVM-based approach (95.20%) [2] and Singh et al.'s color-feature method (93.50%) [3], by leveraging hybrid feature fusion to capture both texture and deep semantic patterns. While Vision Transformers (ViT) [20] and MobileNetV2 [21] achieve high accuracy on mixed-crop datasets, their performance on chilli Anthracnose remains suboptimal compared to our dedicated fusion framework (99.58%). Srinivasulu et al. [28] proposed RNDDNet, a dilated DenseNet variant, attaining 98.09% accuracy on chili disease detection. Though their residual connections improve feature propagation, the model's reliance on deep architectures increases computational complexity.

5.3 Limitations and Weaknesses

Despite its strengths, the study has notable limitations: The overheads in computations created by feature fusion (HOG + LBP + SURF + VGG19) would adversely impact inference execution, at least under real-time conditions on edge devices with limited resources. Furthermore, while the dataset is limited to a specific anthracnose, it thus limits the model, as its results are not applicable to other chili diseases, such as powdery mildew or bacterial spot. Also, training with top-end GPUs (e.g. NVIDIA RTX 3090) makes the model non-usable in resource-restricted areas where these GPUs are not available. Another drawback is ecological interactions the model's reliability is undetermined in the cases of occlusion or payments with the soil background images that are frequent in the field situations. In-time settings such as variability in lighting, leaf movement due to wind, which creates motion blur, and weather

(rain, humidity) can have serious effects on accurate anthracnose detection in video feeds. The accuracy loss due to lighting variability can be mitigated by combining adaptive histogram equalization and learnable exposure compensation. The integrity of features degraded by motion blur can be preserved with an optical flow-CNN stabilizer. The effects of weather conditions on performance can be diminished through weather-adaptive augmentation during training.

5.4 Implications of the Study

The study has significant practical and theoretical implications, especially when it comes to the transformation of agricultural practices due to the early detection of anthracnose, the reduction of unpleasant losses in production, and the most effective use of fungicides with the help of real-time monitoring tools, such as mobile applications or drones. It promotes sustainability by enhancing precision agriculture through the elimination of unnecessary pesticide use, which aligns with eco-friendly farming activities. In terms of methodological contribution, the proven efficiency of hybrid feature fusion presents a precedent to be observed in other plant disease research; the diagnostic efficacy in plant pathology would increase. The scalability of the model to low-resolution images will make it applicable even when used with IoT devices in rural areas where resources are not well developed, democratizing access to superior agricultural technology. Lastly, the framework can assist policymakers in adopting a data-driven approach to implement AI-based advisory systems that lead to the integration of smart solutions into national agricultural strategies to ensure long-term food security.

6. Conclusion

Effective and speedy identification of crop diseases is critical for food security and sustainable farming. This research addresses the urgent issue of detecting the invasion of anthracnose disease in chili plants by conducting an excellent combination of the handcrafted approach to feature fusion, transfer learning, and convolutional neural networks. The Hybrid Feature Extractor, composed of HOG, LBP, and SURF in combination with the VGG19 model trained on chili leaf images, allowed for the extraction of discriminative and complementary features. This hybrid method generated impressive results in terms of classification accuracy, 99.58%, which surpasses existing state-of-the-art approaches and serves as evidence of the relevance of traditional computer vision methods in collaboration with deep learning in the agricultural domain. The experimental results of the proposed model are further supported by its flexibility to different image resolutions (128×128 to 256×256 pixels) and computational efficiency. Unlike current techniques, which deploy high-resolution inputs or dedicated hardware to generate high levels of accuracy, the proposed approach attained high correct classification accuracy in cases where smaller-sized images are used; hence, the framework can be deployed on resource-limited devices in rural agricultural environments. Although the present study has strong points, it also recognizes limitations. The complexity of computing feature fusion-handcrafted and deep features—poses challenges for real-time implementations on edge devices. As future work, dimensionality reduction strategies (e.g. Principal Component Analysis) to increase inference speed while maintaining the same accuracy, as well as using lightweight architectures (e.g. MobileNet), can be investigated. Additionally, it would also be worth increasing the database to cover other diseases of chili (e.g. powdery mildew, bacterial spot) and environmental differences (e.g. occlusions, soil backgrounds), to facilitate the generalizability of the model.

References

[1] Sahid, Zulfikar Damaralam, Muhamad Syukur, Awang Maharijaya, and Waras Nurcholis. "Quantitative and qualitative diversity of chili (Capsicum spp.) genotypes." Biodiversitas Journal of Biological Diversity 23, no. 2 (2022).

[2] Saini, Tejbhan Jalsingh, Ganapati Bhat, Anshuman Tiwari, Shantikumar Gupta, and Radhamani Anandalakshmi. "Identification, prevalence and pathogenicity of Colletotrichum species associated with chilli anthracnose in India." Journal of Plant Pathology 107, no. 1 (2025): 3-22.

[3] Singh, Vivek, U. K. Tripathi, Abhishek Singh, Ashwani Kumar Patel, and Mukesh Kumar. "Exploring Chilli Anthracnose in Uttar Pradesh's Key Cultivation Regions." Int. J. Plant Soil Sci 35, no. 19 (2023): 2265-2270.

[4] Chin, Ruben, Cagatay Catal, and Ayalew Kassahun. "Plant disease detection using drones in precision agriculture." Precision Agriculture 24, no. 5 (2023): 1663-1682.

[5] Miller, Sally A., Anna L. Testen, Jonathan M. Jacobs, and Melanie L. Lewis Ivey. "Mitigating emerging and reemerging diseases of fruit and vegetable crops in a changing climate." Phytopathology® 114, no. 5 (2024): 917-929.

[6] Yang, Li-Na, Maozhi Ren, and Jiasui Zhan. "Modeling plant diseases under climate change: evolutionary perspectives." Trends in plant science 28, no. 5 (2023): 519-526.

[7] Miao, Ke & Chen, Chenglei & Zheng, Xianqing. (2023). A study on the application of artificial intelligence in the design of intelligent medical robots. Applied Mathematics and Nonlinear Sciences. 9. 10.2478/amns.2023.2.01388.

[8] Mangalika, Udula. "Object Recognition to Content Based Image Retrieval: A Study of the Developments and Applications of Computer Vision." (2024).

[9] Sajitha, P., A. Diana Andrushia, N. Anand, and Mohannad Z. Naser. "A review on machine learning and deep learning image-based plant disease classification for industrial farming systems." Journal of Industrial Information Integration 38 (2024): 100572.

[10] Dyke, Roberto M., and Kai Hormann. "Histogram equalization using a selective filter." The visual computer 39, no. 12 (2023): 6221-6235.

[11] Yoshihara, Akira, Kazuki Fujikawa, Kazuhiro Seki, and Kuniaki Uehara. "Predicting stock market trends by recurrent deep neural networks." In Pacific rim international conference on artificial intelligence, pp. 759-769. Cham: Springer International Publishing, 2014

[12] Awasthi, Rohit Kumar, and Srikant Singh. "An Overview of Machine Learning Methods for the Detection of Diseases in Rice Plants in Agricultural Research." (2023).

[13] Wani, Javaid Ahmad, Sparsh Sharma, Malik Muzamil, Suhaib Ahmed, Surbhi Sharma, and Saurabh Singh. "Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges." Archives of Computational methods in Engineering 29, no. 1 (2022): 641-677.

[14] Yu, Lixiran, Hong Xie, Yan Xu, Qiao Li, Youwei Jiang, Hongfei Tao, and Mahemujiang Aihemaiti. "Identification and Monitoring of Irrigated Areas in Arid Areas Based on Sentinel-2 Time-Series Data and a Machine Learning Algorithm." Agriculture 14, no. 10 (2024): 1693.

[15] Gharsallah, Olfa, Claudio Gandolfi, and Arianna Facchi. "Methodologies for the sustainability assessment of agricultural production systems, with a focus on rice: a review." Sustainability 13, no. 19 (2021): 11123.

[16] Ha, Jisu, Jun-Young Park, Yoonseok Choi, Pahn-Shick Chang, and Kyung-Min Park. "Comparative analysis of universal protein extraction methodologies for screening of lipase activity from agricultural products." Catalysts 11, no. 7 (2021): 816.

[17] Tetila, Everton Castelão, Bruno Brandoli Machado, Gabriel Kirsten Menezes, Adair da Silva Oliveira, Marco Alvarez, Willian Paraguassu Amorim, Nícolas Alessandro De Souza Belete, Gercina Gonçalves Da Silva, and Hemerson Pistori. "Automatic recognition of soybean leaf diseases using UAV images and deep convolutional neural networks." IEEE geoscience and remote sensing letters 17, no. 5 (2019): 903-907.

[18] Waheed, Abdul, Muskan Goyal, Deepak Gupta, Ashish Khanna, Aboul Ella Hassanien, and Hari Mohan Pandey. "An optimized dense convolutional neural network model for disease recognition and classification in corn leaf." Computers and Electronics in Agriculture 175 (2020): 105456.

[19] Agarwal, Mohit, Suneet Kr Gupta, and Kanad Kishore Biswas. "Development of Efficient CNN model for Tomato crop disease identification." Sustainable Computing: Informatics and Systems 28 (2020): 100407.

[20] Chitta, Subrahmanyasarma, Vinay Kumar Yandrapalli, and Shubham Sharma. "Deep Learning for Precision Agriculture: Evaluating CNNs and Vision Transformers in Rice Disease Classification." In 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0, IEEE, 2024, 1-6.

[21] Pandey, Priyanshu, and Rusha Patra. "A Real-time Web-based Application for Automated Plant Disease Classification using Deep Learning." In 2023 IEEE International Symposium on Smart Electronic Systems (iSES), IEEE, 2023, 230-235.

[22] Moupojou, Emmanuel, Appolinaire Tagne, Florent Retraint, Anicet Tadonkemwa, Dongmo Wilfried, Hyppolite Tapamo, and Marcellin Nkenlifack. "FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning." IEEE Access 11 (2023): 35398-35410.

[23] Al-Gaashani, Mehdhar SAM, Nagwan Abdel Samee, Reem Alkanhel, Ghada Atteia, Hanaa A. Abdallah, Asadulla Ashurov, and Mohammed Saleh Ali Muthanna. "Deep transfer learning with gravitational search algorithm for enhanced plant disease classification." Heliyon 10, no. 7 (2024).

[24] [Preethi, P., R. Swathika, S. Kaliraj, R. Premkumar, and J. Yogapriya. "Deep Learning–Based Enhanced Optimization for Automated Rice Plant Disease Detection and Classification." Food and Energy Security 13, no. 5 (2024): e70001.

[25] Ojo, Mike O., and Azlan Zahid. "Improving deep learning classifiers performance via preprocessing and class imbalance approaches in a plant disease detection pipeline." Agronomy 13, no. 3 (2023): 887.

[26] Alhwaiti, Yousef, Muntazir Khan, Muhammad Asim, Muhammad Hameed Siddiqi, Muhammad Ishaq, and Madallah Alruwaili. "Leveraging YOLO deep learning models to enhance plant disease identification." Scientific Reports 15, no. 1 (2025): 7969.

[27] Adnan, Faiqa, Mazhar Javed Awan, Amena Mahmoud, Haitham Nobanee, Awais Yasin, and Azlan Mohd Zain. "EfficientNetB3-adaptive augmented deep learning (AADL) for multi-class plant disease classification." IEEE Access 11 (2023): 85426-85440.

[28] Srinivasulu, Maramreddy, and Sandipan Maiti. "RNDDNet: A residual nested dilated DenseNet based deep-learning model for chilli plant disease classification." Engineering Research Express 6, no. 3 (2024): 035204.

[29] "Anthracnose Disease in Chili." Accessed: Apr. 20, 2025. [Online]. Available: https://www.kaggle.com/datasets/prudhvi143413s/anthracnose-disease-in-chili-palnadu-ap

[30] Agustina, Ina, Fauziah Nasir, and Anggit Setiawan. "The implementation of image smoothing to reduce noise using Gaussian filter." (2017).

[31] Admass, Wasyihun Sema, Yirga Yayeh Munaye, and Girmaw Andualem Bogale. "Convolutional neural networks and histogram-oriented gradients: a hybrid approach for automatic mango disease detection and classification." International Journal of Information Technology 16, no. 2 (2024): 817-829.

[32] Wu, Hao, Lincong Fang, Qian Yu, and Chengzhuan Yang. "Composite descriptor based on contour and appearance for plant species identification." Engineering Applications of Artificial Intelligence 133 (2024): 108291.

[33] Mulcahy, Colm. "Image compression using the Haar wavelet transform." Spelman Science and Mathematics Journal 1, no. 1 (1997): 22-31.

[34] Probierz, Eryka. "On emotion detection and recognition using a context-aware approach by social robots: modification of faster R-CNN and YOLO v3 neural networks." (2023).

[35] Chen, Junde, Defu Zhang, Md Suzauddola, and Adnan Zeb. "Identifying crop diseases using attention embedded MobileNet-V2 model." Applied Soft Computing 113 (2021): 107901.

[36] Likhar, Kiran, and Sonali Ridhorkar. "Enhancing skin cancer detection: A comparative analysis of models with VGG-16, VGG-19, and inception V3." International Journal of Intelligent Systems and Applications in Engineering 12 (2024): 502-514.

[37] Nawar, Abbas Khalifa, Hadi Raheem Ali, Mothefer Majeed Jahefer, and Sabah Abdulazeez Jebur. "Automate facial paralysis detection using vgg architectures." ijciar. v7i1 158.

[38] Zheng, Yufeng, Clifford Yang, and Alex Merkulov. "Breast cancer screening using convolutional neural network and follow-up digital mammography." In Computational Imaging III, vol. 10669, 1066905. SPIE, 2018.