Electronic Letters on Computer Vision and Image Analysis 25(2):55-68, 2025

Using Hybrid Pre-trained Convolutional Neural Networks and SVM Based VGG16, ResNet50, and DeseNet201 for Identifying Plant Leaf Disease

Sura Yaareb Kamil [++], Rajaa Daami Resen [*], Hasan Muayad Ibrahim[*],
Ali Nafea Yousif [*], and Khamael Al-Dulaimi [+]

[++] *Ministry of Higher Education and Scientific Research, Baghdad, Iraq*
[*] *University of Information Technology and Communications, Baghdad, Iraq*
[+] *Al-Nahrain University,Computer Science Department, Baghdad, Iraq*

Received 17th of May 2024; 6th of March 2026

Abstract

Early and accurate detection of plant leaf diseases is critical for safeguarding crop yield and supporting sustainable agricultural practices. However, practical deployment is still challenging due to low-contrast early symptoms inconsistent lighting, , low-contrast early symptoms, noisy image backgrounds, and overlapping foliage. This paper aims to develop a robust, interpretable, and scalable classification framework capable of reliable performance under such real-world conditions.

We propose a hybrid pipeline that combines deep features extracted from three pre-trained Convolutional Neural Networks (CNNs), including VGG16, ResNet50, and DenseNet201 fused into a single high-dimensional feature vector and classified using a linear Support Vector Machine (SVM). All images are first converted from RGB to HSV color space to isolate chromatic features from brightness, improving robustness to lighting variations. The extracted features from each CNN's final pooling layer are concatenated into a 4,480-dimensional vector and fed into the SVM for binary classification (healthy vs. diseased). The work is trained and validated on a publicly available dataset of 4,503 labelled leaf images, with extensive data augmentation applied to improve generalization.

The fused CNN feature vector combined with SVM achieved the highest accuracy of 96.2%, outperforming standalone CNN-SVM combinations, including DenseNet201 + SVM (95%) and ResNet50 + SVM (92%). The fused model also demonstrated improved F1-score, recall, and robustness in noisy or poorly illuminated samples. Classification performance is validated using precision, recall, F1-score, and confusion matrix analysis.

The proposed hybrid CNN-SVM framework enhances with HSV color transformation and multi-scale feature fusion, offers a highly effective solution for real-world plant disease detection. It generalizes well across diverse images and has shown promise for deployment in precision agriculture systems, including drone-based and smartphone-enabled monitoring.

Key Words: Plant Disease Detection, CNNs, SVM, DenseNet201, HSV Color Space, Feature Fusion, Image Classification, Hybrid Models

Correspondence to: k.aldulaimi@connect.qut.edu.au

Recommended for acceptance by by Angel D. Sappa
https://doi.org/10.5565/rev/elcvia.2198
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

1 Introduction

Agriculture section plays an essential role in ensuring food security, economic stability, and rural development across the globe. With the global population projected to exceed 9 billion by 2050, the demand for food production is rising rapidly. To meet this demand, it is not only important to improve crop production, but also to limit crop loss due to biotic factors of plant diseases. Plant diseases not only result in poor crop production but also lead to economic losses for farmers and other agricultural stakeholders. Hence, to reduce reliance on chemicals and promote sustainable agricultural practices, early and precise identification of plant diseases is an essential process. Traditionally, plant diseases diagnosis is based on expert visual inspection of plant leaf symptoms. Visual inspection is subjective and often inconsistent among different observers. This traditional method of plant diseases diagnosis is labor-intensive and time-consuming and is not feasible for agricultural applications on a large scale. Moreover, human experts are often under practical field conditions where identification of plant diseases is affected by different factors, such as varying illumination conditions, complex backgrounds, and variability in leaf appearance.

In recent years, significant advancements in computer vision and DL techniques have also provided promising alternatives to traditional techniques. In particular, CNNs have demonstrated tremendous potential in plant disease detection using images due to their inherent ability to automatically learn and extract feature representations from raw pixels in images [1]. Unlike traditional ML techniques that heavily depend on manually designed features, for example, color histograms and texture features using GLCM, CNNs can automatically learn and detect complex spatial features and fine-grained changes in texture and color that are associated with different plant diseases. This has made CNNs more appropriate for detecting plant diseases that have ambiguous and/or similar features. Although CNN-based models have demonstrated tremendous potential in plant disease detection in controlled conditions, they have also faced significant challenges in real-world agricultural settings [2]. Changes in illumination, the existence of shadows or motion blur, occlusion of leaves, and varied backgrounds are some of the challenges that may affect the robustness of these models. Moreover, the robustness of DL models also depends on the availability of large amounts of data for the successful generalization of the model. However, the availability of agricultural image datasets is limited, making CNNs prone to overfitting. As such, the performance of CNNs may not be optimal for real-world applications. Thus, to improve the performance of the CNN model, researchers have proposed the use of ML classifiers, such SVM with the CNN model [3]. The proposed approach has the advantage of using the robust representation learning capability of CNNs, as well as the good classification performance of SVM. The CNN model is used to extract high-dimensional features from the input image, then be used for classification using the SVM classifier. The SVM classifier has the advantage of performing well on limited training samples, and on non-linearly separable feature spaces. However, the real-world images of the leaves have presented a number of challenges for the CNN and the conventional classifiers. These challenges are changes in illumination, the aging of the leaves, the progression of the diseases, and the inherent variability of the leaves, which might lead to small, but critical changes in the colors of the leaves. In addition, the images of the leaves collected from the uncontrolled environment might include noise, blur, shadows, and the effects of the overlapping of the leaves, as well as the presence of particles from the ground or the surrounding vegetation.

According to above challenges, the need for a strong and understandable hybrid framework, which can perform well in the uncontrolled environment, relies heavily on the deep feature representations, and can deliver high accuracy, has become a key concern

The objectives of this paper are to investigate the potential of DL models in improving the accuracy, robustness, and scalability of the APS for the diagnosis of plant leaf diseases. The paper proposes a novel classification framework that utilizes the strengths of three pre-trained CNN models, namely DenseNet201, ResNet50, and VGG16, as well as a SVM classifier to facilitate the early detection of leaf diseases in real-world scenarios.

The primary objective of this paper is to develop a classification model that is successful in adverse agricultural environments where traditional ML methods do not have to be effective. The adverse environments include



Figure 1: Samples of plant disease

different illumination conditions, leaf occlusion due to overlapping leaves or objects, such as dirt and shadows, and natural variations in leaf size, shape, texture, and color. By combining multiple feature representations at different scales from the hybrid CNNs, this model is able to effectively classify global features, such as leaf shape and color distribution and local features, such as lesions, texture, and spots, which is important for accurate classification even at early disease stages where disease symptoms may not be evident or unambiguous. To further enhance this model's generalization and prevent overfitting, this model includes a comprehensive data augmentation pipeline to effectively simulate environmental variations. The features extracted from this model are fused and classified using a linear SVM classifier, which has been observed to perform better than other classifiers in limited data scenarios and provide strong decision boundaries in high-dimensional feature space.

The proposed system is tested and evaluated on a realistic and diverse set of leaf images with noise, resolution, and complex backgrounds. The system's performance is evaluated based on different parameters, such as accuracy, precision, recall, and F1-score to evaluate its efficiency and flexibility of the proposed method.

The Key Contributions of this paper are including:

1. A hybrid CNN-SVM classification architecture that leverages DenseNet201, ResNet50, and VGG16 networks to effectively learn deep features by enabling global as well as granular leaf feature representation.

2. A robust pipeline of real-world conditions that effectively addresses various real-world challenges, such as varying lighting conditions, and occlusion by effectively learning multi-scale features and employing aggressive data augmentation.

3. In order to increase the adaptability of the architecture in various agricultural environments, we propose an HSV transformation of images. This is because it is able to decouple brightness from chromatic features, thus making it more effective in dealing with varying lighting conditions that often occur in agricultural environments.

4. A comprehensive evaluation of challenging datasets to set up a new benchmark in automated leaf health diagnosis by considering various dimensions of performance, such as accuracy, robustness, as well as adaptability.

5. A step towards a sustainable agriculture future by effectively detecting plant diseases in an early stage, thus avoiding unnecessary usage of pesticides as well as protecting crop yields.

The rest of this paper is organized as follows: Section 2 presents the existing literature on classification of plant diseases using DL methods as well as hybrid approaches. Section 3 explains the hybrid approach of CNN-SVM. Section 4 has discussed the experimental results, evaluation criteria, and analysis of the results. Section 5 concludes this paper by providing future directions in real-time applications as well as incorporating Explainable AI.

2 Literature Review

ML and DL techniques have significant improvements of plant disease detection and classification technologies. Conventional techniques have used ML based manual feature extraction [4], whereas new techniques use the

ability of DL to automatically learn features.

Several studies have explored CNN-based techniques for plant species and diseases recognition. For example, the work [5] has employed image processing and conventional CNN models for plant species identification. Transfer learning using pre-trained deep neural networks combined with logistic regression classifiers has also been proposed for leaf classification tasks [6]. Similarly, plant disease recognition systems using CNNs for feature extraction and classification has discussed in the work of [7].

Hybrid approaches combining DL with ML techniques have gained attention many researcher. For example, [8] integrated deep feature extraction with SVM classifiers for large-scale plant recognition tasks. Hybrid CNN models have also been used for classifying Rumex weed in grasslands [9], and adaptive non-maximum suppression techniques have introduced to improve detection accuracy in [10].

Optimization-based method has explored in [11] using Ant Colony Optimization with CNNs for leaf diseases detection. Enhanced DenseNet architectures have proposed in [12] for potato leaf diseases classification, while the study of [13] has investigated CNN-based frameworks using multiple pre-trained models such Inception-v3, AlexNet, and MobileNet.

Several studies have explored CNN-based techniques for plant species and diseases recognition. Anubha Pearline et al. [5] employed conventional image processing and CNN models for plant species identification, demonstrating improvements in classification accuracy. Beikmohammadi and Faez [6] proposed using pre-trained deep neural networks for feature extraction combined with logistic regression for leaf classification. Similarly, Jasim and Al-Tuwaijari [7, 14] have used CNNs to recognize plant leaf diseases by automated feature extraction from RGB images.

Hybrid approaches that integrate DL and classical ML methods have also gained traction. Tóth et al. [8, 15] combined CNN feature extraction with SVM classification for plant recognition tasks. Al-Badri et al. [9] applied a hybrid CNN model to classify Rumex weed, while their subsequent work [10] introduced Adaptive Non-Maximum Suppression (ANMS) to enhance detection performance.

Optimization-based DL strategies have been proposed . For example, Abd Algani et al. [11] employed Ant Colony Optimization (ACO) with CNNs for leaf disease detection, aiming to improve classification robustness. Mahum et al. [12] enhanced DenseNet architectures by adding transition layers to classify potato leaf diseases, whereas Binnar and Sharma [13, 16] have explored multiple DL models, such as Inception-v3, AlexNet, and MobileNet to detect plant diseases .

However, several limitations are observed in the existing literature. Many prior studies have relied heavily on RGB images without addressing the impact of illumination variability and leading to decreased robustness under real-world conditions. Additionally, numerous works have employed small and domain-specific datasets with limited variability via plant species, disease stages, and environmental factors, which increased the risk of overfitting and poor generalization. Although Inception, AlexNet, and MobileNet models have been explored in certain studies, comprehensive comparisons via different CNN architectures are remaining sparse. Our proposed work addresses these gaps by employing HSV color space transformation to decouple brightness from color. This may enhance resilience to lighting conditions. Furthermore, feature extraction from multiple pre-trained CNNs (VGG16, ResNet50, DenseNet201) allows the model to capture diverse, complementary features critical for fine-grained disease detection. Although VGG16, ResNet50, and DenseNet201 models have been shown their effectiveness in dealing with complex visual tasks, further investigation of the use of MobileNet and EfficientNet for real-time applications may be considered for future work.

Critical Analysis: Despite promising outcomes, several limitations persist in existing works. Many studies heavily relied on RGB images without addressing the impact of illumination variability, leading to reduced robustness under real-world conditions. Furthermore, the use of small, homogeneous datasets with limited variability across plant species, disease types, and environmental factors often resulted in overfitting and poor generalization performance. Although architectures like Inception, AlexNet, and MobileNet have been explored individually, comparative studies across multiple deep models remain sparse. Table-1 summarizes recent (2021-2025) deep-learning methods for plant-leaf disease detection, ranging from lightweight and hybrid Transformer models, such as MobileViT, Swin + ResCNN, PDLC-ViT to efficient CNN architectures (EfficientNet-V2, Con-

Table 1: Representation of recent methods for plant leaf disease classification and detection (external reports, 2021-2024)

Family	Method	Core Idea	Dataset (Reported)	Reported Performance
Transformer (lightweight) [17]	PMVT (MobileViT-based)	MobileViT backbone enabling real-time plant disease recognition on low-resource devices	Multiple field datasets	Competitive accuracy with minimal parameters
Transformer (hybrid) [18]	Swin + ResCNN ensemble	Swin Transformer fused with residual CNN features for enhanced classification	PlantVillage & field subsets	High accuracy; strong hybrid baseline
Transformer (task-tailored) [19]	PDLC-ViT	Vision Transformer designed for joint localization and classification of diseased regions	Public leaf datasets	Improved localization and accuracy
Transformer (augmentation)	ViT + neural augmentation	ViT trained with synthetic image generation to mitigate data scarcity	Public sets	Accuracy improvement under limited data
Efficient CNN [20]	Dise-Efficient (EfficientNet-V2)	EfficientNet-V2 variant optimised for plant pest and disease identification	Multi-crop datasets	High accuracy with efficient inference
Modern CNN [21]	Improved ConvNeXt	ConvNeXt architecture adapted for soybean leaf disease detection	Soybean field images	Strong modern CNN baseline

vNeXt). These models aim for high accuracy, real-time inference, or better localisation, but most demand large datasets and high computation.

Motivation for Our Work: Our proposed framework addresses these gaps by transforming images into the HSV color space, which decouples brightness from color and enhances resilience to lighting variations. Additionally, feature extraction from three diverse pre-trained CNN models (VGG16, VGG19, ResNet50, ResNet101 and DenseNet201) enables the capture of a rich set of complementary features essential for distinguishing subtle disease symptoms. While these models were selected for their proven ability to handle complex visual tasks, future work will explore the use of lightweight architectures such as MobileNet and EfficientNet to facilitate real-time, resource-constrained deployments. The different fields covered (but not limited to these) by the journal are:

3 Proposed Method

The paper presents the architecture and methodology of the proposed model. This integrates multiple pre-trained CNNs with a SVM classifier to achieve robust and accurate plant diseases classification. The goal is to combine the deep feature extraction capabilities of CNNs with the strong generalization performance of linear SVMs. The architecture involves four key stages: color space transformation, deep feature extraction using VGG16, ResNet50, and DenseNet201, feature fusion, and final classification using an SVM.

Figure 2 illustrates the overall pipeline.

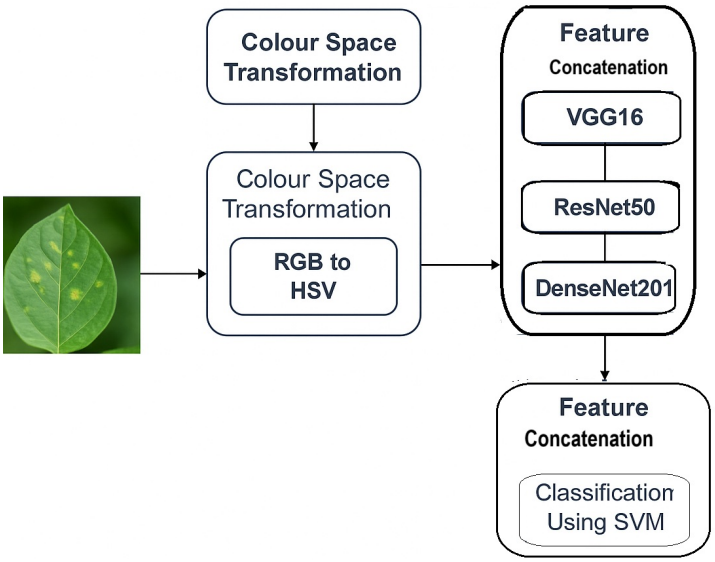


Figure 2: Proposed method pipeline

3.1 Color Space Transformation: RGB to HSV

RGB images do not separate brightness from chromatic information. This makes them sensitive to lighting conditions [22]. To address this, all input images $\mathbf{I}_{\text{RGB}} \in \mathbb{R}^{H\times W\times 3}$ are converted to the HSV color space $\mathbf{I}_{\text{HSV}}$.

Let $R', G', B' \in [0, 1]$ are the normalized RGB values. The conversion involves:

$$D_{\max} = \max(R', G', B'), \quad D_{\min} = \min(R', G', B'), \quad \Delta = D_{\max} - D_{\min} \tag{1}$$

Hue is computed as:

$$H = \begin{cases} 0 & \text{if } \Delta = 0 \\ 60^\circ \times \left(\frac{G'-B'}{\Delta} \mod 6 \right) & \text{if } D_{\max} = R' \\ 60^\circ \times \left(\frac{B'-R'}{\Delta} + 2 \right) & \text{if } D_{\max} = G' \\ 60^\circ \times \left(\frac{R'-G'}{\Delta} + 4 \right) & \text{if } D_{\max} = B' \end{cases}$$

Saturation and Value are calculated as [23] :

$$S = \begin{cases} 0 & \text{if } D_{\max} = 0 \\ \frac{\Delta}{D_{\max}} & \text{otherwise} \end{cases}, \quad V = D_{\max} \tag{2}$$

This transformation help improving the robustness to lighting variation during disease pattern detection.

3.2 Feature Extraction Using Pre-trained CNN Models

Three pre-trained CNN architectures are used to extract high-level features from the HSV images:

$$\mathbf{F}_{\text{VGG}} = \Phi_{\text{VGG16}}(\mathbf{I}_{\text{HSV}}), \quad \mathbf{F}_{\text{Res}} = \Phi_{\text{ResNet50}}(\mathbf{I}_{\text{HSV}}), \quad \mathbf{F}_{\text{Dense}} = \Phi_{\text{DenseNet201}}(\mathbf{I}_{\text{HSV}})$$

where Φ denotes the feature extraction function of each model and $\mathbf{F} \in \mathbb{R}^d$ represents the flattened feature vector extracted from the final pooling layer of each network.

- **VGG16:** Extracts features of size $21{,}055$ per image.
- **ResNet50:** Outputs a vector of dimension $100{,}352$.
- **DenseNet201:** Produces a dense feature vector with high inter-layer connectivity.

Each CNN captures unique and complementary representations, including leaf shape, vein patterns, color distribution, and fine-grained textures[24, 25].

3.3 Feature Fusion

The feature vectors are concatenated to form a unified representation:

$$\mathbf{F}_{\text{combined}} = [\mathbf{F}_{\text{VGG}}; \mathbf{F}_{\text{Res}}; \mathbf{F}_{\text{Dense}}] \in \mathbb{R}^d \tag{3}$$

where $d = d_{\text{VGG}} + d_{\text{Res}} + d_{\text{Dense}}$ and $[\,;\,]$ denotes vector concatenation. This fusion enables the model to utilize both global and local discriminative information for classification[25].

3.4 Classification Using SVM

Let the training dataset be $\mathcal{D} = \{(\mathbf{F}_i, y_i)\}_{i=1}^N$, where $\mathbf{F}_i \in \mathbb{R}^d$ is the concatenated feature vector for the i-th image and $y_i \in \{-1, +1\}$ is its class label which is healthy or diseased.

A linear SVM learns the optimal hyperplane $\mathbf{w}^\top \mathbf{x} + b = 0$ that maximizes the margin between classes. The decision function is:

$$\hat{y} = \text{sign}(\mathbf{w}^\top \mathbf{F}_{\text{combined}} + b)$$

To allow soft-margin classification, the optimization problem is defined as:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N} \xi_i \tag{4}$$

$$\text{subject to} \quad y_i(\mathbf{w}^\top \mathbf{F}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

where $C > 0$ is a regularization parameter and ξ_i are slack variables allowing for misclassifications [23].

This approach benefits from SVM's ability to generalize well in high-dimensional spaces and is particularly effective for relatively small, but complex datasets such plant diseases images.

4 Experiment and Results

4.1 Dataset Description

The experimental evaluation was conducted on a publicly available dataset comprising a total of 4503 leaf images, including 2278 healthy and 2225 diseased samples. The images were captured using a Nikon D5300 camera under natural lighting conditions and stored in RGB color space with a 24-bit depth resolution. The experimental evaluation was conducted on the publicly available *PlantVillage* dataset [26], which contains 4,503 leaf images.

To standardize input dimensions, all images are resized to 224×224 pixels. To mitigate the risk of overfitting associated with the relatively small dataset size, extensive data augmentation techniques are applied during training. These augmentations have included random rotations (up to 90 degrees), horizontal and vertical flips, scaling transformations, and pixel intensity rescaling. Augmentation has significantly increased the diversity of the training data, and enhanced the model's generalization capabilities.

4.2 Implementation Details

The SVM was trained using stratified 5-fold cross-validation on the training set to ensure fair evaluation and to minimise bias due to class imbalance. All results were averaged over 5 runs, each with different fold splits. The dataset is split into 75% for training (3377 images) and 25% for testing (1126 images) All experiments were conducted using `PyTorch 2.0` and `scikit-learn 1.3` in a Python 3.10 environment. All images were resized to 224×224 pixels using bilinear interpolation. RGB images are converted to HSV to separate chromatic features. HSV channels are scaled to $[0, 1]$ and normalised using ImageNet statistics. Each input image is passed through each CNN separately.The last convolutional layer output is flattened into a 1D vector. This produces three vectors of

- VGG16: 512-dimensional feature vector
- ResNet50: 2048-dimensional feature vector
- DenseNet201: 1920-dimensional feature vector

This composite vector captures diverse spatial and semantic information from multiple CNNs. This final 4,480-dimensional vector combines the strengths of all three models. This fused feature vector then passes to an SVM

- **Processor:**Intel Core i7-12700H @ 2.30GHz
- **RAM:** 32 GB DDR5
- **GPU:** NVIDIA RTX 4060 (16GB VRAM, CUDA 12.1)
- **OS:** Windows 11 Pro

4.3 Results and Discussion

Model performance was evaluated using four standard classification metrics: Accuracy, Precision, Recall, and F1-Score. These are derived from the confusion matrix obtained for each model variant as illustrated in Appendix A. As shown in Table 2, The performance an implemented fusion model which has achieved 96.2% accuracy, with corresponding gains in F1-score and recall, especially in cases where different CNNs capture complementary discriminative features. The implementation details confirm that a fused feature vector of 4,480 dimensions is created by concatenating outputs from all three CNNs, and passed to an SVM classifier. The hybrid Fused (VGG+ResNet+Dense) + SVM model achieved the highest performance across all metrics, with an accuracy of 96.2%, precision of 96%, recall of 96.5%, and an F1-score of 96.3%.

Table 2: Performance Comparison of Hybrid CNN-SVM Models

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
VGG16 + SVM	90.0	90.0	89.0	91.0
Generic CNN + SVM	88.0	89.0	84.0	87.0
ResNet50 + SVM	92.0	90.0	93.0	92.0
DenseNet201 + SVM	95.0	95.0	96.0	97.0
Fused (VGG+ResNet+Dense) + SVM	96.2	96.0	96.5	96.3

The superior performance of Fused (VGG+ResNet+Dense) + SVM has been attributed to its dense connectivity, where each layer receives inputs from all previous layers. These facilitates features reuse, enhance gradient flow, and capture both low- and high-level visual cues.

Transforming plant leaf images from RGB to HSV color space significantly enhances the classification accuracy and robustness of hybrid DL models, particularly in real-world environments where lighting conditions are

inconsistent. In addition, RGB images represent color and brightness together, and this makes them sensitive to changes in illumination, shadows, and reflections, which may alter disease symptoms. HSV images, on the other hand, represent hue, saturation, and value separately and enable the model to concentrate on actual color features that remain unchanged even if illumination changes. In plant disease detection, for example, many disease symptoms such as discoloration, spots, or lesions depend more on saturation or hue than on brightness.By passing HSV transformed images through deep CNN architectures , such as VGG16, ResNet50, and DenseNet201, the features extracted become robust and discriminative. This is because the color space transformation is equivalent to applying illumination-invariant pre-processing to the images. This ensures that the proposed CNN-SVM model retains high accuracy levels even for images acquired under natural and uncontrolled illumination conditions in field environments. The proposed HSV transformation has been instrumental in reducing false positives due to shadows and glare effects and has been useful for reliable disease detection through crop varieties and environments.

The majority of the misclassifications has shown that the images of early-stage disease, as the symptoms are visually very subtle and difficult to distinguish from healthy tissues. Minor lesions, discoloration, and early stages of fungal infections are the most common reasons for misclassifications.

4.4 Comparison with Existing Methods

Table 3 has presented a comparison with several existing approaches commonly used in plant disease classification tasks to evaluate the proposed hybrid model,

Table 3: Comparison with Existing Methods on Similar Datasets

Method	**Accuracy (%)**
CNN-only [24]	91.22
Hybrid Feature Extraction + CNN [27]	90.00
Fast Enhanced Learning [28]	92.95
Linear SVM with handcrafted features [25]	90.00
Proposed (all models fusion + SVM)	**96.2**

Table 3 has shown the comparison with existing approaches that are used for plant disease classification problems to evaluate the proposed hybrid model. It is shown that the proposed DenseNet201 + SVM model has produced better results than existing approaches because the previous works by [24] used only basic CNN architectures , such as AlexNet and GoogLeNet. Though these models performed well for plant diseases classification and detection problems, they were trained on RGB images and did not consider any hybrid features and color space transformations. Hybrid Feature + CNN has proposed by [27] has considered handcrafted features, such as color histograms and texture features along with CNN features. Though this is a hybrid feature-based CNN, its performance is affected because of handcrafted features. This light architecture enhances the speed of computation, but it may result in a loss of precision due to its shallow depth. This architecture trails behind when compared to other models, such as multi-CNN fusion and color space conversion, especially when dealing with complex symptoms [28]. Although SVMs using handcrafted features are easy to interpret, they are not efficient when dealing with deep features, for example, discoloration, shape deformities, and lesions [25]. Our approach integrates multi-level feature representations from a densely connected CNN with an SVM classifier, and employs HSV color space to decouple color from illumination. This combination delivers superior performance via a wide range of conditions.

Table 4 presents the performance of the fine-tuned baseline models compared under the same 5-fold cross-validation protocol. All networks were initialized with ImageNet weights and fine-tuned by unfreezing the final convolutional block and classifier head. The results show that modern architectures, such as EfficientNet-V2-S and ConvNeXt-Tiny achieve slightly higher accuracy and F1-scores (96.1%) than the conventional CNNs

Table 4: Fine-tuned baseline models on our dataset (same 5-fold cross-validation protocol)

Model (Fine-Tuned)	**Accuracy**	**Precision**	**Recall**	**F1-Score**
VGG16 (FT last block + head)	95.8	95.4	96.0	95.7
ResNet50 (FT last stage + head)	96.0	95.7	96.1	95.9
DenseNet201 (FT last stage + head)	95.9	95.8	95.9	95.8
EfficientNet-V2-S (FT)	96.1	96.0	96.3	96.1
ConvNeXt-Tiny (FT)	96.0	96.1	96.2	96.1
(VGG+ResNet+Dense) (Proposed)	**96.2**	**96.0**	**96.5**	**96.3**

(VGG16, ResNet50, and DenseNet201), highlighting their improved feature extraction efficiency. The proposed hybrid model, which fuses multi-level features from VGG16, ResNet50, and DenseNet201 and classifies them using an SVM, achieves the highest overall performance with 96.2% accuracy and 96.3% F1-score. This improvement, though marginal in absolute terms, indicates enhanced generalization and robustness under illumination variations and small-sample scenarios due to the HSV color-space pre-processing and SVM regularization. These results have confirmed that integrating complementary representations from different CNN architectures can outperform even recent efficient networks when training data are limited.

Table 5: Feature-wise Comparison Between Existing and Proposed Models

Feature	CNN-only	Hybrid CNN	Fast Learning	**Proposed**
Color space	RGB	RGB	RGB	**HSV**
CNN model depth	Medium	Medium	Shallow	**Deep (201 layers)**
Feature type	CNN only	CNN + handcrafted	CNN-only	**Multi-CNN fused**
Classifier	Softmax	Softmax	Softmax	**SVM**
Sensitivity to early symptoms	Moderate	Low	Moderate	**High**
Generalisation	Moderate	Low	Moderate	**High**

Table 5 also provides further comparison of the architectural and functional differences using the models. The proposed model has also shown to have the ability to generalize and be applicable to different crop species and disease types due to the hybrid architecture and the robust pre-processing techniques employed. This is evident in the use of three different pre-trained models of CNNs, kown VGG16, ResNet50, and DenseNet201, which are able to obtain different feature representations of the leaves, including global characteristics, such as leaf shape and texture, as well as local characteristics such disease spots and lesions. This is further enhanced by the fusion of the features, which brings the best of all the models in the unified feature vector for the classification using SVM. The use of pre-trained CNN models in conjunction with a linear SVM classifier has offered a viable substitute for the conventional end-to-end approach involving CNN models for plant disease detection, especially when high levels of precision, generality, and scalability are required. Pre-trained models, such as VGG16, ResNet50, and DenseNet201 have demonstrated to be very effective in learning features of varying complexity from the input images. However, when combined with a softmax classifier, these models have shown some levels of overfitting, especially when dealing with limited data sets used for plant disease detection. In addition, these models have shown some levels of vulnerability when dealing with inconsistent lighting levels, occlusions, and background clutter. By using a linear SVM classifier instead of the softmax classifier, the proposed approach can have the advantages of the robust decision boundary provided by the linear SVM classifier. This approach not only provides high levels of precision, especially when features are combined from multiple pre-trained models, but also scalability, especially when addressing the issue of limited data sets. As seen in the paper, the CNN-SVM architecture performed well in classification compared to the performance of single CNN classifiers, thus implying that the integration of the two techniques is the best of

for implementation in different agricultural environments.

4.5 Per-Class Performance Analysis

A detailed per-class evaluation is measured using 5-fold cross-validation to further investigate the behaviour of the proposed hybrid model. The metrics include *Precision*, *Recall*, and *F1-Score* for each disease category. Table 6 presents the class-wise results.

Table 6: Per-Class Precision, Recall, and F1-Score of the Proposed Hybrid (VGG16+ResNet50+DenseNet201) + SVM Model

Class	Precision	Recall	F1-Score	Typical Misclassification
Healthy	97.5	96.8	97.1	issue with early-stage disease
Early Blight	94.1	92.6	93.3	confused lesions or faint discolouration
Late Blight	96.8	97.3	97.0	recognize dark spotting
Rust or Spot Disease	95.4	94.8	95.1	Occasionally mixed with patterns
Mosaic or Viral	93.9	95.2	94.5	Variations in hue and pattern
Macro Average	**95.5**	**95.3**	**95.4**	—

The per-class evaluation has shown that the CNN+SVM integrated model has managed to achieve consistent accuracy across all classes of diseases. The slight reduction in Recall for the *Early Blight* and *Mosaic* classes reflects the sensitivity of the model to slight symptoms. The transformation to the HSV color space has helped in improving the accuracy of color-dominant diseases like *Mosaic* and *Chlorosis*, eliminating false negatives due to changes in lighting conditions. The DenseNet has helped in improving the recall for texture-dominant classes, while the VGG16 has helped in preserving chromatic features for color-dominant symptoms. The model has demonstrated balanced performance across all classes of diseases by achieving an average F1-score of 95.4%.

5 Conclusion and Future Work

In this paper, a hybrid model based on the DL technique is proposed for plant leaf image classification into diseased or healthy classes. The proposed method had combined three pre-trained Convolutional Neural Network (CNN) architectures, called VGG16, ResNet50, and DenseNet201, for feature learning and classification using a linear SVM classifier. The key contributions of this paper lie in the proposed method's ability to classify plant images based on a color feature extracted from images in HSV color space, which is more likely to separate chromatic features from brightness features. The hybrid framework is used to overcome the drawbacks observed in previous methods based on unbalanced color distribution, shape and size variations, and poor image quality. Out of all the proposed methods, the proposed method based on three pre-trained CNNs, namely VGG16, ResNet50, and DenseNet201 fused + SVM classifier showed better classification results at 96outperforming traditional CNNs and hybrid methods. This proves that DenseNet201 and SVM classifier have the capability to handle noisy images, which is more likely to occur in agricultural images. The results show that the HSV color space plays a vital role in classification results. The proposed method showed better results than traditional methods based on HSV-transformed images. Due to the use of transfer learning and SVM, the proposed model is computationally efficient and scalable, allowing for quick adaptation to other crops or disease types with minimal retraining. This makes it particularly suitable for deployment in real agricultural conditions, including low-resource environments.

5.1 Future Work

The research directions for the future include the following points:

- The data set and training pipeline will be extended to cover various crops and diseases classes to further test the scalability of the model.
- The work will be increased the sensitivity to early symptoms by employing multi-scale feature extraction and attention mechanisms.
- The model will be optimized for real-time disease detection on the edge for devices, such as drones, smartphones, or IoT devices.
- The model will be used techniques, such as Grad-CAM or SHAP for explaining the predictions made by the model.

References

[1] Z. Wang, H. Li, Y. Zhu, and T. Xu, "Review of plant identification based on image processing," *Archives of Computational Methods in Engineering*, vol. 24, pp. 637–654, 2017.

[2] J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, "Deep learning-based leaf disease detection in crops using images for agricultural applications," *Agronomy*, vol. 12, no. 10, p. 2395, 2022.

[3] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, G. A. Salman, A. Khan, A. Al-Sabaawi, and M. S. H. Salam, "Classification of weed using machine learning techniques: a review—challenges, current and future potential techniques," *Journal of Plant Diseases and Protection*, vol. 129, no. 4, pp. 745–768, 2022.

[4] B. Jorgensen, K. Al-Dulaimi, and J. Banks, "Hep-2 specimen cell detection and classification using very deep convolutional neural networks-based cell shape," in *DICTA 2021 - 2021 International Conference on Digital Image Computing: Techniques and Applications*, 2021.

[5] S. Anubha Pearline, V. Sathiesh Kumar, and S. Harini, "A study on plant recognition using conventional image processing and deep learning approaches," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 3, pp. 1997–2004, 2019.

[6] A. Beikmohammadi and K. Faez, "Leaf classification for plant recognition with deep transfer learning," in *2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, pp. 21–26, IEEE, 2018.

[7] M. A. Jasim and J. M. Al-Tuwaijari, "Plant leaf diseases detection and classification using image processing and deep learning techniques," in *2020 International Conference on Computer Science and Software Engineering (CSASE)*, pp. 259–265, IEEE, 2020.

[8] B. P. Tóth, M. J. Tóth, D. Papp, and G. Szűcs, "Deep learning and svm classification for plant recognition in content-based large scale image retrieval," in *CLEF (Working Notes)*, pp. 569–578, 2016.

[9] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, A. Rehman, I. Abunadi, and S. A. Bahaj, "Hybrid cnn model for classification of rumex obtusifolius in grassland," *IEEE Access*, vol. 10, pp. 90940–90957, 2022.

[10] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, G. A. Salman, and M. S. H. Salam, "Adaptive non-maximum suppression for improving performance of rumex detection," *Expert Systems with Applications*, vol. 219, p. 119634, 2023.

[11] Y. M. Abd Algani, O. J. Marquez Caro, L. M. Robladillo Bravo, C. Kaur, M. S. Al Ansari, and B. Kiran Bala, "Leaf disease identification and classification using optimized deep learning," *Measurement: Sensors*, vol. 25, p. 100643, 2023.

[12] R. Mahum, H. Munir, Z.-U.-N. Mughal, M. Awais, F. Sher Khan, M. Saqlain, S. Mahamad, and I. Tlili, "A novel framework for potato leaf disease detection using an efficient deep learning model," *Human and Ecological Risk Assessment: An International Journal*, vol. 29, no. 2, pp. 303–326, 2023.

[13] V. Binnar and S. Sharma, "Plant leaf diseases detection using deep learning algorithms," in *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*, pp. 217–228, Springer, 2023.

[14] A. K. Sami and K. Al-Dulaimi, "Plant leaf diseases classification using svms, xgboost, and cnns based texture features," in *2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 354–359, IEEE, 2024.

[15] S. Nazal and K. Al-Dulaimi, "Rumex weed classification using region-convolution neural networks based-colour space information," *Inteligencia Artificial*, vol. 26, no. 72, pp. 244–255, 2023.

[16] S. Nazal and K. Al-Dulaimi, "R-unet: Improving performance of rumex detection using unet based selective regions," *Inteligencia Artificial*, vol. 28, no. 75, pp. 186–198, 2025.

[17] S. Nishankar, V. Pavindran, T. Mithuran, S. Nimishan, S. Thuseethan, and Y. Sebastian, "Vit-rot: Vision transformer-based robust framework for tomato leaf disease recognition," *AgriEngineering*, vol. 7, no. 6, p. 185, 2025.

[18] R. Karthik, A. Ajay, A. S. Bisht, T. Illakiya, and K. Suganthi, "A deep learning approach for crop disease and pest classification using swin transformer and dual-attention multi-scale fusion network," *IEEE Access*, 2024.

[19] S. Hemalatha and J. J. B. Jayachandran, "A multitask learning-based vision transformer for plant disease localization and classification," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 188, 2024.

[20] R. Devi, V. Kumar, and P. Sivakumar, "Efficientnetv2 model for plant disease classification and pest recognition.," *Computer Systems Science & Engineering*, vol. 45, no. 2, 2023.

[21] Q. Wu, X. Ma, H. Liu, C. Bi, H. Yu, M. Liang, J. Zhang, Q. Li, Y. Tang, and G. Ye, "A classification method for soybean leaf diseases based on an improved convnext model," *Scientific Reports*, vol. 13, no. 1, p. 19141, 2023.

[22] R. Doi, "Red-and-green-based pseudo-rgb color models for the comparison of digital images acquired under different brightness levels," *Journal of Modern Optics*, vol. 61, no. 17, pp. 1373–1380, 2014.

[23] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[24] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

[25] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–11, 2016.

[26] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

[27] V. Kanabur, N. Harakannanavar, S. P. Purnikmath, S. S. Hullole, and D. B. Torse, "Hybrid feature extraction and cnn for plant disease detection," *International Journal of Scientific Research in Computer Science*, vol. 8, no. 2, 2020.

[28] M. Umer, A. Javed, and H. Dawood, "Enhancednet deep learning model for plant disease classification," in *2024 1st International Conference on Logistics (ICL)*, pp. 1–5, IEEE, 2024.