Hierarchical Feature Fusion With Inception V3 for Multiclass Plant Disease Classification

Muhammad Esham Qureshi [1], Muhammad Hassaan Ashraf [1,2*], Muhammad Waqar Arshad [1], Ahmed Khan [1], Haider Ali [1] and Zain-ul-Abdeen[1]
[1] Faculty of Computing, Riphah International University, Islamabad, 46000, Pakistan
[2] Department of Computer Science, COMSATS University Islamabad, Islamabad, 44000, Pakistan
E-mail: 36511@students.riphah.edu.pk, hassaan.ashraf@riphah.edu.pk, waqar.arshad@riphah.edu.pk, 37635@students.riphah.edu.pk, 37950@students.riphah.edu.pk, 35515@students.riphah.edu.pk

Artificial Intelligence (AI) and Deep Learning (DL) are revolutionizing plant disease detection, which is crucial for mitigating crop loss, improving food security, and enhancing yield. Traditional manual methods, such as field visits for disease diagnosis, are labor-intensive, time-consuming, and costly, highlighting the demand for automated, real-time applications. However, plant disease classification faces challenges such as class variations, cluttered backgrounds, lesion scale variations, and the need for robust models. Several Convolutional Neural Network (CNN)-based frameworks have been proposed to tackle these challenges, but they often suffer from limitations such as fixed-size kernels and inefficient feature utilization across deeper layers. These limitations lead to feature loss during the initial stages of feature extraction, reducing the model's overall effectiveness in multi-class plant disease classification. To address these issues, we introduce HFFIncep-Net, a novel Deep Learning framework that fuses hierarchical features with the power of the Inception V3 architecture. The Hierarchical Feature Fusion (HFF) stem extracts multiscale features, which are fused at multiple levels within the network. This enhances feature representation and mitigates information loss in the initial layers. Incorporating two Inception V3 blocks, the model captures a diverse set of features at different scales, further improving classification performance. Additionally, the inclusion of a Global Average Pooling (GAP) layer reduces computational complexity while maintaining high accuracy. To enhance training stability and performance, we employ the Swish activation function, which facilitates smoother gradient propagation during backpropagation. When evaluated on the PlantVillage and Cassava Leaf Disease (CLD) datasets, HFFIncep-Net achieves 94.61% precision, 94.31% recall, 94.31% F1 score, and 97.80% accuracy, outperforming existing CNN-based methods. These results demonstrate that HFFIncep-Net is an effective solution for early-stage, multi-class plant disease classification, with significant potential for practical deployment in agricultural applications.

Povzetek: Članek predstavi model HFFIncep-Net, ki z večnivojsko fuzijo značilk in izboljšano arhitekturo Inception V3 natančno klasificira več vrst bolezni rastlin kljub vizualnim in merskim variacijam.

1 Introduction

Agriculture serves as the foundation of the food distribution network[1], and the financial systems of many developing nations rely heavily on it [2]. The prevalence of plant diseases significantly impacts crop health and diminishes yields, leading to substantial losses [3]. Such yield reductions can have a ripple effect on the food supply network and the nation's economy. Plant diseases are typically caused by various agents, including viruses, bacteria, parasites, fungi, or environmental factors. Many cultivators may be unaware of certain types of plant diseases[4]. Recognizing plant diseases in their early stages is crucial for controlling outbreaks and preventing extensive damage to crops [5]. To effectively handle these disease occurrences, regular discussions with specialists are crucial. Nevertheless, regular visual evaluations by specialists in isolated regions of developing countries tend to be costly, less accurate, and time-consuming. Therefore, automatically identifying crop disease symptoms present a valuable, quick, and cost-effective solution [6].

Typically, the symptoms of crop diseases manifest on the leaves, creating digital representations of leaves makes them ideal contenders for identifying plant diseases. In recent years, various intelligent systems have been developed for crop disease diagnosis, which can be categorized into two main types: (i) Traditional computer vision-based detection [6] and (ii) Deep learning-based detection [7]. Standard systems depend on feature extraction and classification approaches. Over the past few decades, numerous feature extraction methodologies, such as Local Binary Patterns (LBP) [8], Scale-Invariant Feature Transform (SIFT) [9], Speeded Up Robust Features (SURF) [10] [11], Histogram of Oriented Gradients (HOG) [12], and Gabor Transform (GT) [13],

have been employed. At the same time, machine learning classifiers, including Support Vector Machine (SVM) [11], Naive Bayes (NB) [14], Random Forest (RF) [14], and Fisher Linear Discriminant (FLD) [15], have been utilized for classifying plant diseases. The precision of traditional diagnostic systems is significantly dependent on the manually crafted feature extraction techniques and classification algorithms employed. Importantly, choosing optimal and resilient handcrafted features for disease identification is especially difficult as each feature comes with its own shortcomings, such as region of interest, scale variation, illumination variation, inter and intra class variation, and dead neurons.

Lately, with the emergence of domain-specific architectures[16] such as GPUs, TPUs, and similar technologies, Deep Learning (DL)-based classifiers, specifically CNNs, have gained significant popularity. A variety of CNN-based models, such as PDD-Net, VGG-16, AlexNet, Inception V3, and ResNet-50 are adept at automatically extracting and learning the best low-level features for classification tasks [17]. These CNN architectures have demonstrated strong efficiency in image classification tasks. However, despite their effectiveness, real-time plant disease categorization faces several obstacles, including early-stage diagnosis, variations in disease classes, limited sample sizes in benchmark datasets, and class imbalance issues[18].

To overcome these issues, this study introduces a blended CNN architecture that leverages recent advancements in the field. The proposed model includes three pooling blocks to boost the retrieval of both local and global multiscale features, enhancing accuracy in identifying small-scale lesions. Furthermore, a dense connection block is associated to promote the flow of information across network layers, thereby enhancing multilevel feature extraction. To reduce model complexity, a GAP layer replaces fully connected layers. The Swish activation function is applied to address potential issues such as gradient vanishing, often linked with traditional activation functions during backpropagation. Swish enhances convergence during training, leading to better performance. HFFIncep-Net is evaluated using the publicly available PlantVillage dataset containing five different plants and nineteen disease classes and CLD dataset containing 5 classes of Cassava leaves demonstrating significant enhancements in classification accuracy when contrasted with the state-of-the-art CNN-based architecture. This study highlights the following significant advances targeted at solving the inadequacies of current computer vision systems for plant disease detection:

- The HFFIncep-Net CNN-based framework employs bicubic interpolation to handle constrained data, reduce information loss, and enhance training efficiency.

- The architecture includes multilevel and multiscale features, a reliable Swish activation function, and GAP layer to improve overall performance and

mitigate errors such as overfitting and gradient vanishing.

- HFFIncep-Net outperforms baseline models such as PDD-Net, DenseNet-201, ResNet-50, AlexNet, MobileNet, and Inception V3 as evidenced by assessment criteria such as accuracy, precision, recall, and F1-score.

This work makes an important contribution to the development of a strong and effective Plant Disease Diagnosis system that effectively detects plant disease and manages variations in disease classifications. The paper is structured as follows: Section 2 provides a literature review on the current state-of-the-art frameworks for plant disease diagnosis. Section 3 describes the materials and techniques utilized in the study, including an in-depth discussion of the proposed HFFIncep-Net architecture. Section 4 explains a full review of the experiments and results. Section 5 summarizes the basic conclusions of the article and suggests possibilities.

2   Literature review and analysis

Over the last decade, several CNN-based frameworks for plant disease detection have been proposed. Alghamdi et al. [19] utilized VGG16 as the baseline architecture and proposed PDD-Net, which integrates multilevel and multiscale CNN features within the VGG16 model. Their approach introduced dense connection blocks to extract multilevel features and a spatial pyramid block to capture multiscale features. These modifications demonstrated that multilevel features are essential for addressing class variations, while multiscale features effectively handle Region of Interest (RoI) scale variations. To further enhance performance, PDD-Net employed the flattened Threshold Swish (fTS) activation function to mitigate gradient vanishing and dead neuron issues. Additionally, the final layer of PDD-Net replaced the fully connected layer with a GAP layer, which reduced the overall number of parameters and minimized overfitting. Although the proposed framework achieved robust results on the PlantVillage and CASSAVA benchmark datasets, PDD-Net utilized the simple feed-forward stem of VGG16, which led to information loss in the deeper layers of the network [20].

Keceli et al. [21] focused on the development of automated methods to enhance the recognition of plant species and the detection of plant diseases. A novel approach is presented that utilizes a multi-input, multi-output CNN integrating raw image data with pre-trained features derived from the AlexNet model. The proposed architecture incorporates convolutional, pooling, and fully connected layers to support accurate classification. The study employed the ReLU (Rectified Linear Unit) activation function within the convolutional layers, though this approach can lead to issues such as gradient vanishing, while classification tasks are handled by fully connected layers culminating in a SoftMax layer. The model employs a categorical cross-entropy loss function to address both species classification and disease detection tasks,

combining the two loss values through a weighted approach to optimize learning. Two datasets were utilized for validation: the PlantVillage dataset, and the FISB dataset, which contains real-world agricultural images. By using AlexNet as a feature extractor, the model benefits from reduced computational complexity and shorter training times due to its relatively low parameter count. The findings demonstrate the model's effectiveness, achieving high accuracy rates: 98% for plant type and health status classification, and 89% for disease classification on the PlantVillage dataset. Similarly, on the FISB dataset, the model achieved 97% accuracy for plant type classification and 94% for disease detection. This study emphasizes the advantages of employing a multi-task learning approach, outperforming single-task frameworks in various metrics such as accuracy, precision, recall, and F1 score.

Chen et al. [22] introduced MS-DNet, a novel lightweight neural network designed for crop disease recognition. Building upon DenseNet, MS-DNet incorporates in depth wise separable convolutions (DSC) and Squeeze-and-Excitation (SE) blocks to enhance feature extraction while reducing model complexity. Preprocessing steps involved converting images to RGB color space, resizing them to 224×224 pixels, and augmenting the dataset using traditional methods (rotation, flipping) and generative adversarial network-based augmentation. The network employed ReLU activation functions in dense blocks and a SoftMax classifier for multi-class classification, utilizing a modified Focal Loss function to handle class imbalance.

Sutaji & Yıldız. [23] proposed LEMOXINET, an ensemble model combining MobileNetV2 and Xception architectures to enhance plant disease detection performance. By concatenating features extracted from both networks, the model aimed to improve accuracy while maintaining computational efficiency. Preprocessing included resizing images to 224×224 for MobileNetV2 and 299×299 for Xception, normalizing pixel intensity values, and augmenting training data with various transformations to simulate different image conditions. The model utilized ReLU activation functions in initial layers and SoftMax in the final classification layer. The model was evaluated on datasets such as iBean, citrus, rice, and Turk-Plants.

Sanida et al. [24] proposed an architecture to improve tomato disease identification by employing a two-stage transfer learning approach to reduce training time and enhance accuracy. Their model integrated the first ten layers of VGG16 with two inception blocks, resulting in a significantly smaller network than the standard VGG16. Images from the PlantVillage dataset were resized to 224×224 pixels and split into training, validation, and testing sets (70%, 20%, and 10% respectively). An enhanced categorical cross-entropy loss function addressed class imbalance during training.

Paul et al. [25] presented a lightweight custom CNN and leveraged transfer learning models VGG-16 and VGG-19 to classify tomato leaf diseases across 11 classes. Preprocessing involved converting images to JPEG format and resizing them to 224×224 pixels. The custom CNN

used the SoftMax activation function in the final layer and employed categorical cross-entropy as the loss function.

Zhou et al. [26] presented a restructured residual dense network (RRDN) to improve tomato leaf disease identification. By integrating the strengths of deep residual and dense networks, RRDN reduced training parameters and enhanced the flow of information and gradients. Preprocessing involved resizing images to 196×196 pixels and normalizing pixel values. The network utilized a 3-layer Residual Dense Block (RDB) with ReLU activation after each convolution and Leaky ReLU post-normalization to mitigate the dead neuron phenomenon. A cross-entropy loss function was employed alongside a SoftMax activation in the output layer for multi-class classification. Evaluated on the AI CHALLENGER dataset containing 13,185 images across 9 classes, RRDN achieved 95% accuracy on both validation and test sets, demonstrating its efficacy in disease identification.

Paymode & Malode. [27] focused on detecting and classifying diseases in crops, particularly tomatoes and grapes, using a Convolutional Neural Network (CNN) based on the VGG16 architecture. Image preprocessing included filtering, grayscale transformation, sharpening, and scaling. Data augmentation techniques such as rotation, translation, and random transformations were applied to enhance the dataset's diversity. The model employed ReLU activation functions and a SoftMax activation in the final layer for classification. Utilizing the PlantVillage dataset, which comprises 54,303 high-quality images across 38 crop classes, the proposed CNN achieved remarkable accuracies of 98.40% for grapes and 95.71% for tomatoes.

Pan et al. [28] introduced TFANet, a two-stage feature aggregation network designed for multi-category soybean leaf disease identification. The model aimed to capture intricate features by aggregating information at different stages. Preprocessing steps involved resizing images to 224×224 pixels and applying extensive image augmentation techniques, including rotation, flipping, color dithering, brightness enhancement, Gaussian noise addition, and adaptive histogram equalization. The model utilized the Sigmoid activation function and employed cross-entropy as the loss function for training. Evaluated on the Auburn Soybean Disease Image Database (ASDID) containing 9,648 images, TFANet demonstrated high efficacy with an accuracy of 98.18%, precision of 98.21%, recall of 98.60%, and an F1-score of 98.39%.

Zhang et al. [29] proposed SDINet, a Siamese Dilated Inception Network tailored for apple leaf disease detection using minimal training samples. The model leverages dilated convolutions within Inception modules to capture multiscale features. Extensive data augmentation was performed, generating 10 augmented images per original through cropping, rotation, rescaling, flipping, brightness adjustment, contrast changes, noise addition, saturation, light enhancement, and random erasing. Images were resized to 256×256 pixels. ReLU activation functions were applied after each convolutional layer and dilated Inception module to enhance nonlinearity. Evaluated on the AppleDisease5 dataset comprising 2,000 images

across five disease classes, SDINet achieved an accuracy of 94.11%, precision of 93.58%, recall of 92.80%, and an F1-score of 93.19%. The model contains fewer than 10 million parameters, significantly less than traditional models like AlexNet.

Dai et al. [30] proposed the DFN-PSAN model, a multi-level deep information feature fusion extraction network designed for plant disease detection in natural field environments. The model integrates a position-sensitive attention mechanism (PSA) to enhance feature representation while reducing parameters by 26% compared to optimal attention mechanisms. Preprocessing steps included converting images to RGB format, resizing to 256×256 pixels, and applying Gaussian filtering and non-local means noise reduction. Weather data augmentation simulated various environmental conditions such as fog, illumination changes, and raindrops to improve robustness. The network utilized ReLU activation in convolutional layers and Sigmoid in attention layers, with a SoftMax classifier at the output. Using cross-entropy loss with label smoothing, the model was trained on datasets like Katra-Twelve, BARI-Sunflower, and FGVC8. DFN-PSAN achieved accuracy of 98.37%, 94.23%, and 93.24% on these datasets respectively, with overall performance exceeding 95.27%.

In another research Dai et al. [31] presented PPLC-Net, a neural network-based model for plant disease detection enhanced by Weather-based data augmentation and a multi-level attention mechanism. The model incorporates dilated convolutions and GAP layers to improve feature extraction. Images were cropped to 256×256 pixels with adaptive scaling. Cross-entropy loss function was employed during training. Evaluated on a dataset of 4,503 images covering 22 leaf classes from 12 plant types, PPLC-Net achieved an impressive accuracy of 99.702%, precision of 98.545%, recall of 98.340%, and an F1-score of 98.442%. With 15.486 million parameters, the model balances complexity and performance in plant disease classification.

Thakur et al. [32] presented VGG-ICNN, a lightweight Convolutional Neural Network designed for crop disease identification using plant leaf images. The model combines the initial four convolutional layers of a pre-trained VGG16 with three blocks of GoogleNet Inception v7, resulting in approximately 6 million parameters significantly fewer than many high-performing models. Preprocessing involved resizing images to 224×224 pixels and normalizing pixel values. The dataset was split into training, validation, and test sets. ReLU activation functions were applied within the Inception blocks, and a SoftMax activation function was used in the fully connected classification layer. The model was evaluated on five public datasets: PlantVillage, Embrapa, Apple, Maize, and Rice, achieving accuracies ranging from 90% to 99%. Specifically, it achieved 99.16% accuracy on PlantVillage and 93.66% on Embrapa.

2.1. Research gap identification and analysis

After conducting a comprehensive and detailed review of the literature discussed above, the following key challenges have been identified and highlighted for exploration and consideration.

2.1.1   Class variation

In agricultural image analysis, class variation is a significant challenge. Intra-class variation refers to differences within the same category or class, such as plants of the same species exhibiting different features like the shape and color of the RoI. Inter-class variation highlights the difficulty of distinguishing between different plant species or diseases that may exhibit similar visual features [19]. These variations make a challenge for machine learning models to classify or detect specific agricultural traits accurately. Addressing class variation require robust feature extraction and classification techniques. To illustrate both these challenges a visual representation is provided in Figure 1 and Figure 2.

Figure 1: Inter-Class Variation a) Grape Black Measles b) Grape Black Rot.

Figure 2: Intra-Class Variation a) Grape Leaf Blight b) Grape Leaf Blight.

2.1.2   Region of interest scale variation

Scale variation refers to changes in the size of RoI in the image caused by different angles and distances of camera while capturing the image [33]. In agriculture, this is particularly problematic for machine learning algorithms as diseases might appear small and distant in one image and large and close-up in another [33]. A visual

representation of the ROI challenge is illustrated in Figure 3.

Figure 3: Region of Interest Scale Variation a) Potato Early Blight at early-stage b) Potato Early Blight at advance stage.

2.2   Research gap

Despite significant progress in using CNNs for plant disease detection, several gaps persist:

1. **Gradients Vanishing in Simple Feed-Forward Models:** Models like VGG16 and AlexNet often struggle with gradients vanishing, causing training difficulties and reducing their ability to learn complex patterns [20]. This limits their effectiveness in handling subtle disease features.

2. **Neglect of Multilevel Features in Inception-Net:** While Inception-based models capture features at multiple scales, they often fail to fully integrate features learned at different depths. This can result in missing important information needed for identifying diseases that vary in appearance across different layers [20].

3. **"Dead Neuron" Phenomenon in Traditional Activation Functions:** Common activation functions like ReLU and Sigmoid can lead to "dead neurons," where some neurons stop responding to any input. This reduces the model's capacity to learn new patterns, ultimately affecting performance [34].

3   Proposed framework

In this section, we provide a comprehensive discussion of the proposed approach, outlining the CNN architecture in detail in the following subsections. Before delving into the specifics, we offer a concise overview to provide readers with a brief understanding.

3.1   Data acquisition

For this study, we utilize the publicly available PlantVillage and the CLD benchmark datasets. These datasets, summarized in Table 1, provide a comprehensive collection of images, enabling the training and evaluation of our proposed model across a diverse range of crops and disease types.

3.1.1   PlantVillage dataset

The PlantVillage dataset is a widely recognized resource for plant disease detection, comprising 18,908 images from various plant species. It offers high-quality, annotated images across multiple classes, facilitating accurate model training and evaluation.

This research focuses on the classification of multiple plants and their associated diseases as shown in Figure 4.
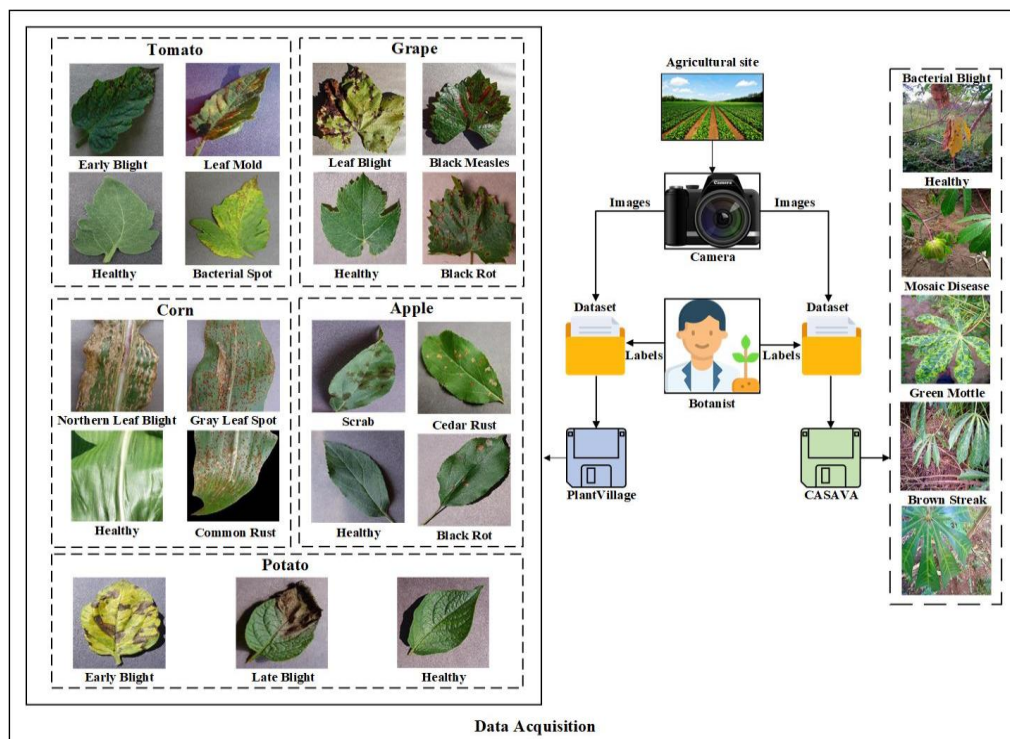

Figure 4: Data preparation.

Table 1: Dataset's description.

| Plants | Classes | Image Frequency | | | |
		Train (70%)	Validate (10%)	Test (20%)	Total
Corn	Northern Leaf Blight	689	98	198	985
	Gray Leaf Spot	359	51	103	513
	Healthy	813	116	233	1162
	Common Rust	834	119	239	1192
Apple	Apple Scab	441	63	126	630
	Healthy	1151	164	330	1645
	Apple Rust	192	27	56	275
	Black Rot	434	62	125	621
Grape	Leaf blight	753	107	216	1076
	Black Measles	968	138	277	1387
	Healthy	296	42	85	423
	Black Rot	826	118	236	1180
Potato	Early Blight	700	100	200	1000
	Late Blight	700	100	200	1000
	Healthy	106	15	31	152
Tomato	Healthy	1113	159	319	1588
	Early Blight	700	100	200	1000
	Bacterial Spot	1488	212	427	2127
	Leaf Mold	666	95	191	952
Cassava	Healthy Cassava Leaves	350	50	100	500
	Cassava Bacterial Blight	350	50	100	500
	Cassava Brown Streak	350	50	100	500
	Cassava Mosaic Disease	350	50	100	500
	Healthy	350	50	100	500

For apples, the identified classes comprise of Apple Rust, Black Rot, Apple Scab and Healthy. Corn is classified into Northern Leaf Blight, Gray Leaf Spot, Common Rust, and Healthy samples. Black Rot, Black Measles, Leaf Blight, and Healthy leaves belong to the grape class. Potato diseases are grouped into Late Blight, Early Blight, and Healthy states. Lastly, tomato classes comprise Leaf Mold, Early Blight, Bacterial spot, and Healthy samples. This classification forms the basis of the study, highlighting a diverse range of plants and associated diseases. The PlantVillage dataset provides a different variety of plants, both healthy and diseased, making it compatible for training a multi-class plant disease detection model.

3.1.2   Cassava leaf disease

The CLD dataset focuses on cassava, one of the most crucial staple crops in many developing countries. Cassava leaves are prone to various diseases, significantly impacting crop yield and food security. To evaluate our model on real-world images, we selected 500 images per class from the CLD dataset, total of 2,500 images. This dataset contains five distinct classes, including both healthy and diseased leaf images as shown in Figure 4.

3.2   Preprocessing

Figure 5 explains the process of standardizing image sizes, which ensures consistency in the dataset, reduces computational demands, and improves the model's performance by providing uniform input images. This study employs bicubic interpolation, a method that calculates new pixel values using the weighted sum of 16 neighboring pixels. This technique provides smoother transitions and improved image quality compared to other methods like nearest-neighbor or bilinear interpolation. Additionally, it excels at retaining smooth gradients, further supporting more accurate feature extraction during training.
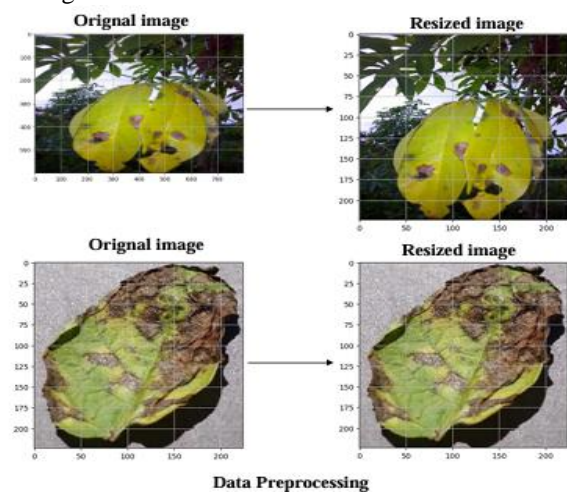

Figure 5: Image resizing.

3.3   HFFIncep-Net for feature extraction and classification

In Figure 6, the proposed HFFIncep-Net architecture, a novel solution designed to overcome the limitations of existing CNNs for multi-class plant disease classification, particularly at early stages. This model integrates hierarchical feature fusion [20] and multiscale feature extraction [35], enhancing its ability to analyze complex

plant disease symptoms with minimal information loss. By combining a robust feature extraction mechanism with advanced architectural enhancements, HFFIncep-Net ensures superior classification performance.

At the core of the architecture is the HFF-Net Stem, which processes 224×224-pixel colored input images to extract hierarchical multiscale features. The stem consists of four parallel convolutional layers, each configured to operate at different scales, enabling the model to capture a diverse range of spatial features. The first convolutional layer (conv.1) extracts feature maps (FMs) of dimensions $224 \times 224 \times 64$ using a 3×3 filter with a stride (S) of 1, preserving fine details in the image. The second layer (conv.2) reduces the input dimensions, producing $112 \times 112 \times 64$ FMs with a 3×3 filter and S = 2, while the third (conv.3) and fourth (conv.4) layers generate $56 \times 56 \times 32$ FMs using 5×5 and 7×7 filters, respectively, both with S = 4. These multiscale feature maps are fused through a carefully designed hierarchical process to retain critical details while minimizing computational complexity.

The hierarchical fusion mechanism employed in the HFF-Net Stem ensures that features from different scales are effectively combined. The FMs from conv.1 are down sampled using a Max Pooling (MP) layer and concatenated with those from conv.2, producing feature maps of dimensions $112 \times 112 \times 128$. These concatenated FMs are further reduced to $112 \times 112 \times 64$ using a Channel Pooling Layer (CPL), preserving essential feature information while optimizing processing efficiency. The process continues as features from conv.3 and conv.4 are integrated, producing a final stem output of $56 \times 56 \times 128$ feature maps (FMs). This robust feature extraction process mitigates information loss and ensures a rich representation of the input image, optimized for subsequent multiscale analysis [20].

To further enhance the model's capacity for multiscale feature extraction, two Inception V3 blocks are incorporated after the HFF-Net Stem. These blocks process the stem output through multiple convolutional filters of varying dimensions 1×1, 3×3, and 5×5 alongside a Max Pooling operation. The outputs from these filters are concatenated, creating a unified representation that captures patterns at various scales. This approach is particularly effective in addressing the diverse manifestations of plant diseases, which may appear as small spots or large lesions. By leveraging these multiscale feature representations, the Inception blocks significantly enhance the model's ability to distinguish between different disease classes.

The performance of the HFFIncep-Net is further improved through several architectural enhancements. The Swish activation function is employed across the network, offering a smoother gradient flow and improved learning stability compared to traditional activation functions. Batch Normalization is applied in the convolutional layers to standardize feature distributions,
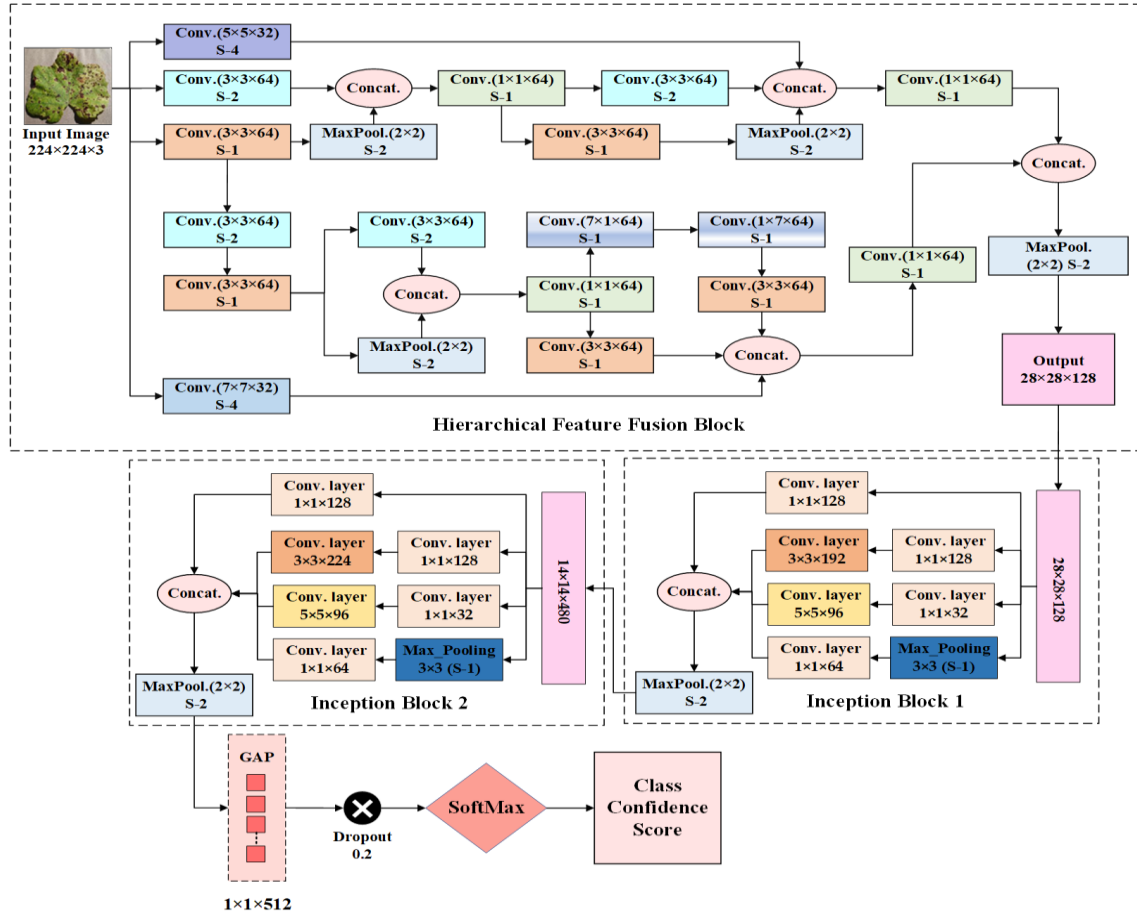

Figure 6: The proposed HFFIncep-Net CNN architecture.

reduce internal covariate shifts, and prevent overfitting. Additionally, a GAP layer is incorporated at the higher levels of the network, aggregating spatial information without the need for fully connected layers. This model not only improves classification accuracy but also reduces the risk of overfitting, ensuring the model's ability to perform effectively on unseen data.

3.4　Activation function

Choosing the right activation function is important for how effectively a model learns and generalizes.

Swish activation function

The Swish activation function is defined as:

$$Sigmoid\ (x) = x * Sigmoid(x) \tag{1}$$

Where $Sigmoid(x)$ is Sigmoid function, and it's given by the equation:

$$Sigmoid\ (x) = \frac{1}{1+e^{-x}} \tag{2}$$

Equation (1) multiplies the input by the sigmoid of the input, creating a smooth curve. This produces negative outputs for negative inputs, which helps avoid the gradient vanishing problem, ensuring gradient flow even for small or negative inputs. These features help the model learn better, keep gradients flowing smoothly, and reduce the chance of neurons becoming "dead" (i.e., stopping learning) [20].

3.5　Loss function

To effectively train the HFFIncep-Net for the plant disease detection, we used the sparse categorical cross entropy loss function. The formula for the loss function is defined in the equation (3):

$$Loss = -\sum_{i=1}^{n} t_i \log\left(p_{i,y_i}\right) \tag{3}$$

Here, L denotes the loss, N indicates the total number of samples, y_i denotes the integer label for the ith sample, and p_{i,y_i} represents the predicted probability that the ith sample belongs to its correct class, represented by y_i.

3.6　Model training and evaluation

In this study, we utilized Google Colab's cloud environment to train and evaluate our HFFIncep-Net for plant disease detection, leveraging the available GPU acceleration. The training setup included a batch size of 32 and was configured for 40 epochs to strike a balance between learning efficiency and model accuracy. We chose the Adam optimizer for its adaptive learning rate capabilities, which helped in effectively facilitating model convergence.

Instead of using model checkpointing, we chose to train the model for the full 40 epochs, ensuring that the final model weights represent comprehensive learning

from the entire dataset. This approach provided a robust framework for detecting various plant diseases and optimized the model's performance in accurately classifying plant health across multiple categories.

3.7　Evaluation metrics

In this study, we assessed the model's ability to detect and categorize different plant diseases using performance indicators, such as accuracy, recall, precision, and the F1-score. Formulas for these metrics are as follows:

$$\boldsymbol{Precision} = \frac{TP}{TP+FP} \tag{4}$$

$$\boldsymbol{Recall} = \frac{TP}{TP+FN} \tag{5}$$

$$\boldsymbol{F1-score} = 2 \times \frac{(Recall \times Precision)}{(Recall+Precision)} \tag{6}$$

$$\boldsymbol{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \tag{7}$$

4　Experiment and results

In this study, we tested the HFFIncep-Net model for plant disease detection and compared its performance with several well-known CNN architectures: ResNet-50, Inception V3, DenseNet-201, VGG16, MobileNet, and AlexNet. Each model has its own strengths. ResNet-50 uses residual connections to solve the vanishing gradient problem, helping it learn better as the network gets deeper. Inception V3 captures features at multiple scales, which is useful for recognizing the varied visual patterns in plant diseases. DenseNet-201 improves feature sharing and reduces computation by connecting each layer to every other layer. VGG16, while simple, is effective in extracting features through its deep structure, and AlexNet provides a solid baseline for comparison. MobileNet, designed for lightweight applications, optimizes computational efficiency while maintaining high accuracy, making it suitable for real-time disease detection in resource-constrained environments.

By comparing HFFIncep-Net against these models, we aim to show its advantages in terms of accuracy, scalability, and handling challenging conditions such as in different lighting, a variety of disease classes, and complex backgrounds. The model demonstrates consistently high precision, recall, and F1-scores, indicating its robustness in classifying both healthy and diseased plants across multiple species. Additionally, the model effectively distinguishes between visually similar diseases, minimizing false positives and false negatives. High classification accuracy observed across different plant species highlights the adaptability of HFFIncep-Net to diverse agricultural conditions. The HFFIncep-Net results on different plants are shown in Table 2.

Table 2: The proposed HFFIncep-Net performance.

Benchmark	Class Label	TP	FP	FN	TN	Pr. %	Rec%	F1. Score %
Apple	Apple Rust	55	0	1	587	100.00	100.00	100.00
	Apple Scab	123	2	3	515	99.21	99.21	99.21
	Black Rot	125	1	6	511	100.00	100.00	100.00
	Healthy	329	8	1	305	99.70	99.70	99.70
Corn	Common Rust	239	0	0	534	100.00	100.00	100.00
	Gray Leaf Spot	87	4	16	666	95.60	84.47	89.69
	Northern Leaf Blight	193	16	5	559	92.34	97.47	94.84
	Healthy	233	1	0	539	99.57	100.00	99.78
Grape	Black Measles	277	2	1	535	99.28	99.64	99.46
	Black Rot	235	0	1	579	100.0	99.58	99.78
	Leaf Blight	213	1	3	598	99.53	98.61	99.06
	Healthy	85	2	0	728	97.70	100.0	98.83
Potato	Early Blight	200	0	0	231	100.00	100.00	100.00
	Late Blight	200	1	0	230	99.50	100.00	99.75
	Healthy	30	0	1	400	100.0	96.77	98.36
Tomato	Bacterial Spot	422	1	5	709	99.76	98.83	99.29
	Early Blight	198	5	2	932	97.54	99.00	98.26
	Leaf Mold	190	2	1	944	98.96	99.48	99.22
	Healthy	318	1	1	817	99.69	99.69	99.69
Cassava	Cassava Green Mottle	75	27	25	373	73.53	75.00	72.16
	Cassava Bacterial Blight	88	19	12	381	82.24	88.00	85.02
	Cassava Brown Streak	70	24	30	376	74.47	70.00	72.16
	Cassava Mosaic Disease	84	28	16	372	75.00	84.00	79.25
	Healthy	74	11	26	389	87.06	74.00	80.00

4.1   Proposed model results

4.1.1. Model HFFIncep-Net for Corn

Figure 7 and Figure 8 display the training and validation loss, as well as training and validation accuracy patterns for corn disease classification, providing further insights into the model's performance. The training loss decreases slowly over the epochs and stabilized after the 8th epoch, indicating the model's capacity to acquire features from the training data. The training accuracy respectively increases and stabilizes at a high level.
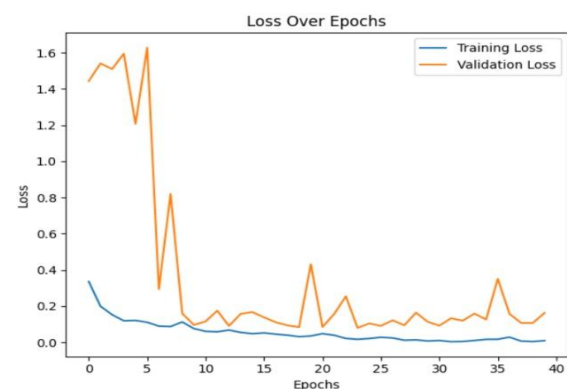

Figure 7: Loss curves of HFFIncep-Net for corn diseases classification.
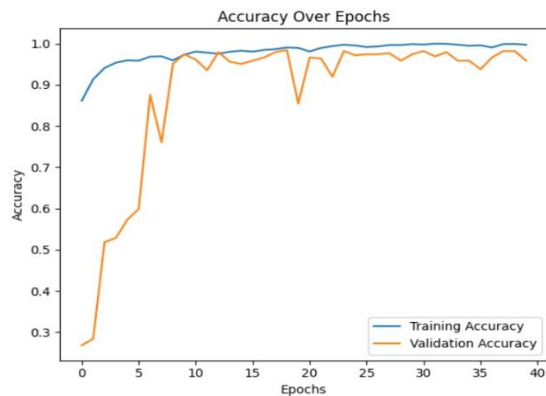
Figure 8: Accuracy curves of HFFIncep-Net for corn diseases classification.

Figure 9 presents the confusion matrix for the HFFIncep-Net in corn disease detection, illustrating the complexity of the classification task, particularly due to significant class variations such as inter- and intra-class differences between Gray Leaf Spot and Northern Leaf Blight. Despite these challenges, the HFFIncep-Net shows impressive performance in identifying various corn diseases.

For Common Rust, the model correctly identified all the 239 samples, proving its strong ability to detect this disease. In the case of Gray Leaf Spot, the HFFIncep-Net accurately classified 87 out of 103 samples, demonstrating its effectiveness in identifying this disease.

The model performed exceptionally well with the more difficult Northern Leaf Blight disease of corn, correctly identifying 193 out of 198 samples, showing it can handle tougher cases. Lastly, for Healthy Corn, the model achieved perfect accuracy by correctly diagnosing all 233 Healthy samples.

Figure 9: Confusion matrix of HFFIncep-Net for Corn diseases.

4.1.2   Model HFFIncep-Net for Tomato

In Figure 10 and Figure 11, training and validation loss as well as training and validation accuracy plots for tomato disease are presented subsequently that provides further understandings into the model's performance. The validation loss and accuracy stabilized after the 5th epoch, indicating that the model effectively learned features by this point.
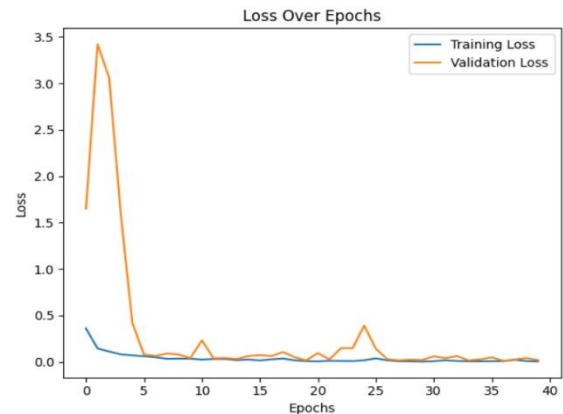

Figure 10: Loss curves of HFFIncep-Net for tomato diseases classification.
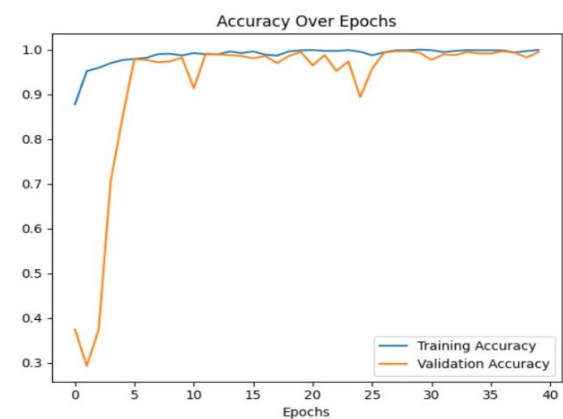

Figure 11: Accuracy curves of HFFIncep-Net for tomato diseases classification.

Figure 12 presents the confusion matrix for the HFFIncep-Net in tomato disease detection, highlights that there is a class variation challenge in bacterial spot and early blight. Although these challenges, the HFFIncep-Net shows impressive performance in identifying various Tomato diseases.

For bacterial spot, the model correctly identified 422 out of 427 samples, proving its strong ability to detect Bacterial Spot disease. In the case of Early Blight, the HFFIncep-Net correctly classified 198 out of 200 samples, establishing its effectiveness in identifying this disease.

The model also performed well with the Leaf Mold disease of Tomato, correctly identifying 190 out of 191

samples just misclassifying one sample, which shows that model is performing accurately in different and complex. Lastly, for Healthy Tomato, the HFFIncep-Net achieved high accuracy, only misclassified 1 sample out of 319 Healthy samples.
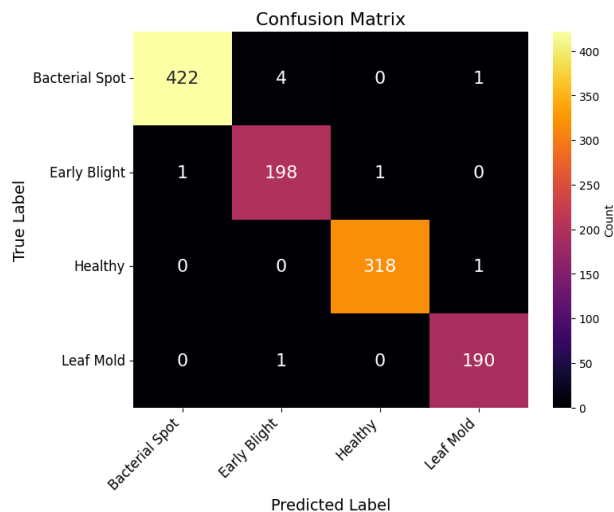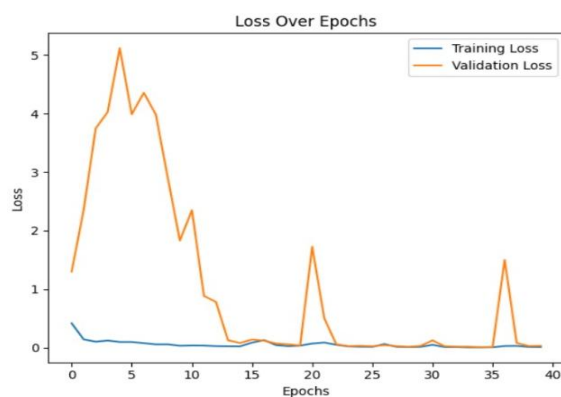

Figure 12: Confusion matrix of HFFIncep-Net for tomato diseases.

4.1.3   Model HFFIncep-Net for Potato

In Figure 13 and Figure 14, which display the training and validation loss as well as training and validation accuracy plots for potato disease subsequently, gives additional insights into the model's performance.

Figure 13: Loss curves of HFFIncep-Net for potato diseases classification.

Figure 15 presents the confusion matrix for the HFFIncep-Net in potato disease detection. The model performs efficiently on all the classes, which shows its strong generalization capability. It correctly distinguishes between Early Blight, Late Blight, and Healthy leaves with minimal error. The model misclassified only one sample across all categories, demonstrating near-perfect performance on potato disease classification. This high accuracy highlights the model's ability to handle subtle visual differences between similar disease classes. Such performance also indicates the model's robustness and reliability in real-world agricultural applications.
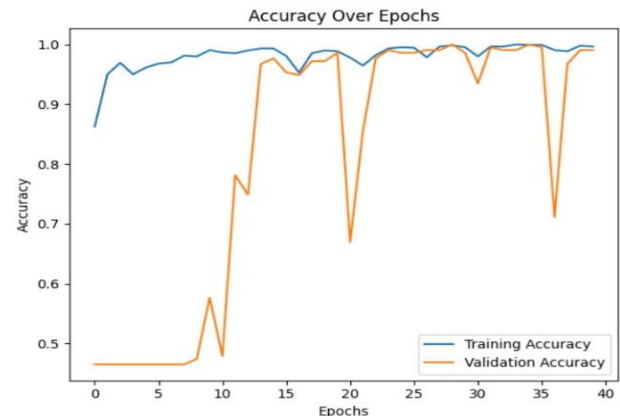

Figure 14: Accuracy curves of HFFIncep-Net for potato diseases classification.

For Early Blight, the model correctly identified all 200 samples. For Healthy, the HFFIncep-Net correctly classified 30 out of 31 samples, and in the case of Late Blight, the model perfectly classified all 200 samples. For Healthy samples, the model misclassified only one sample out of 30. The extremely low misclassification rate demonstrates the model's precision in distinguishing diseased and healthy potato leaves. This further suggests that HFFIncep-Net could be integrated into smart farming systems for automated disease detection and management.
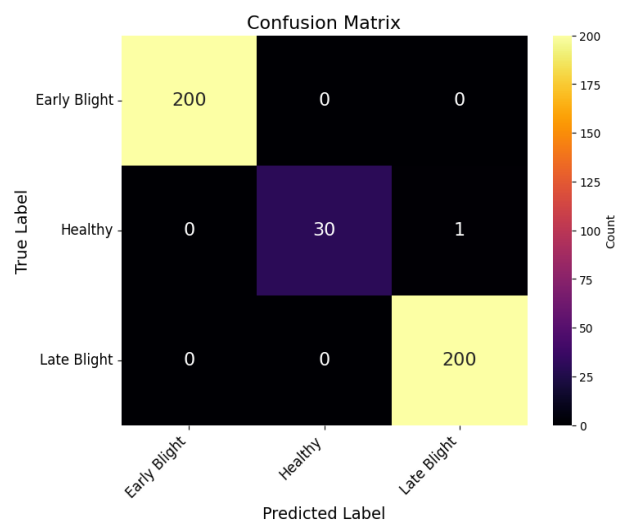

Figure 15: Confusion matrix of HFFIncep-Net for potato diseases.

4.1.4   Model HFFIncep-Net for Grapes

Figure 16 and Figure 17 show the training and validation loss as well as training and validation accuracy plots for grape disease subsequently, which provides further understanding into the model's performance. The validation loss and validation accuracy stabilized after the 7th epoch, which displays efficient model learning ability. This indicates that the model successfully avoids overfitting while maintaining high generalization performance.
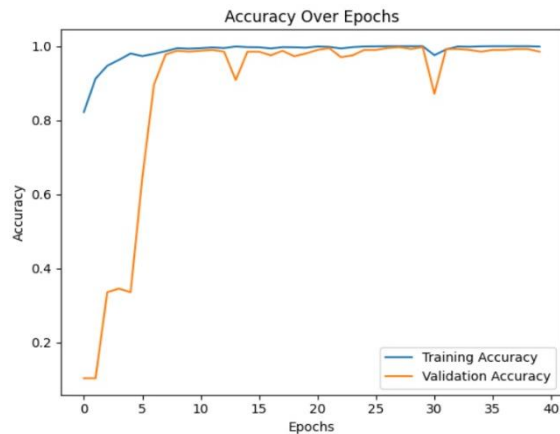
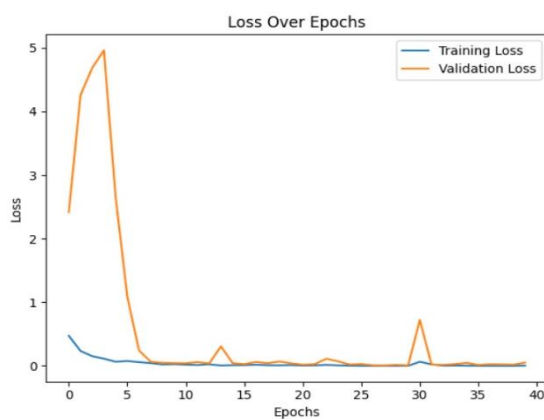Figure 16: Loss curves of HFFIncep-Net for grapes diseases classification.

Figure 17: Accuracy curves of HFFIncep-Net for grapes diseases classification.

Figure 18 presents the confusion matrix for the HFFIncep-Net in grape disease detection, highlights that model is performing effectively on all the classes of grapes, and only misclassifies 4 samples. In the case of Black Measles, the model correctly classifies all the 277 samples. And for the case of Grape Black Rot the model misclassifies only 1 sample out of 236 samples.
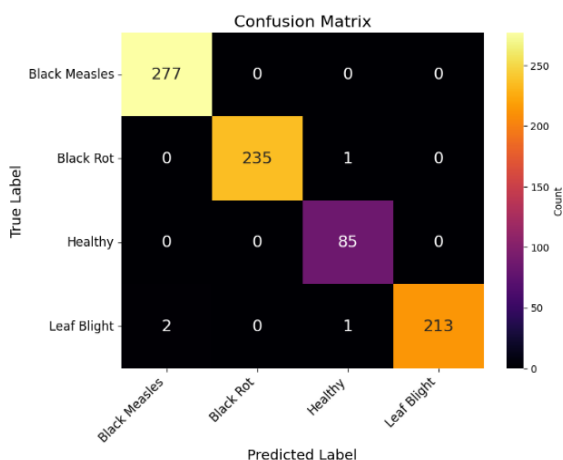

Figure 18: Confusion matrix of HFFIncep-Net for grapes diseases.

For Leaf Blight, the model correctly identified 213 out of 216 samples, proving its strong ability to detect Grape Leaf Blight disease. In the case of healthy grape, the HFFIncep-Net correctly classified all 85 samples which shows valuable model performance.

4.1.5   Model HFFIncep-Net for Apple

Figure 19 and Figure 20 show's how the model performed during training. Figure 19 tracks the training and validation loss, while Figure 20 shows the training and validation accuracy for detecting apple diseases. After the 10th epoch, both the validation loss and accuracy stabilized, meaning the model has learned the important features.
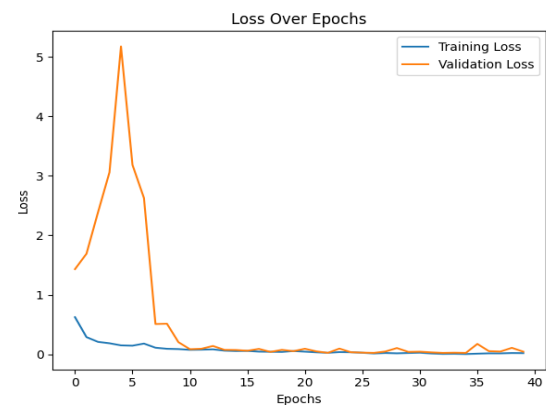

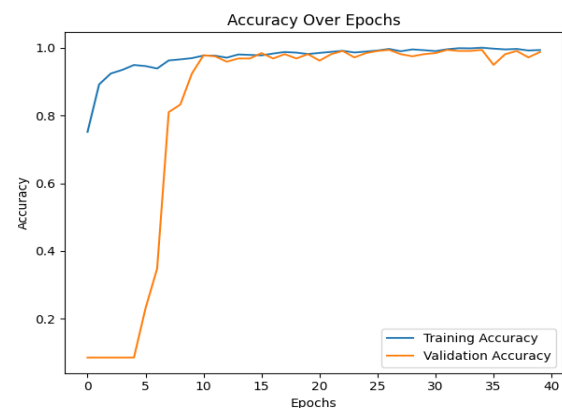Figure 19: Loss curves of HFFIncep-Net for apple diseases classification.

Figure 20: Accuracy curves of HFFIncep-Net for apple diseases classification.

In Figure 21 presents the confusion matrix for the HFFIncep-Net in apple disease detection, highlights that model is performing perfectly on all the categories of Apple, and only misclassifies 2 sample between Healthy and Apple Scab class due to the high-class variation problem. In the case of Apple Rust and Black Rot, the model correctly classifies all the samples, proving its strong ability to detect apple black rot disease And in the case of Apple Scab the model misclassifies only 1 sample out of 125 samples. In the case of healthy Apple, the HFFIncep-Net correctly classified 329 samples and only misclassified 1 sample.
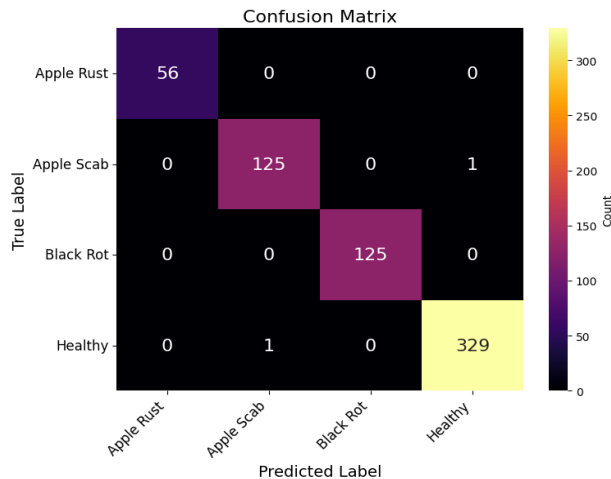
Figure 21: Confusion matrix of HFFIncep-Net for apple diseases.

4.1.6   Model HFFIncep-Net for Cassava

Figure 22 and Figure 23 display the training and validation loss, as well as the training and validation accuracy patterns for Cassava disease classification, providing insights into the model's performance. The training loss gradually decreases over the epochs, while the validation loss also trends downward but exhibits more fluctuations, indicating the model's capacity to learn from the training data. Meanwhile, the training accuracy steadily increases and stabilizes at a high level.
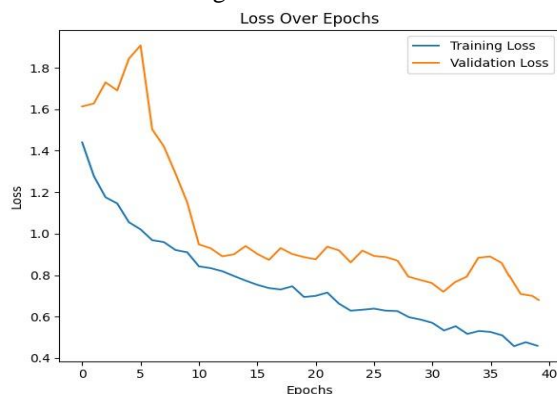

Figure 22: Loss curves of HFFIncep-Net for casava diseases classification.

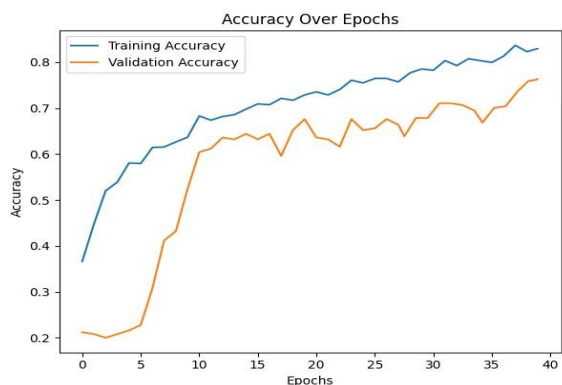
Figure 23: Accuracy curves of HFFIncep-Net for casava diseases classification.

Figure 24 presents the confusion matrix for the HFFIncep-Net in Cassava disease detection, highlighting the model's ability to classify five categories of Cassava leaf conditions. While the model generally performs well, the matrix reveals some misclassifications due to inter- and intra-class variation. The model correctly classifies 88 out of 100 samples for Cassava Bacterial Blight (CBB), demonstrating strong detection of this disease. For Cassava Brown Streak (CBS) case model correctly identifies 70 out of 100 samples, occasionally confusing CBS with CBB. In the case of Cassava Green Mottle (CGM), 75 samples are correctly classified out of 100, indicating moderate confusion with Cassava Mosaic Disease. For Cassava Mosaic Disease (CMD), model accurately classifies 84 out of 100 CMD samples, showing reliable performance on this category. For Healthy Leaves, the model correctly identifies 74 out of 100 samples, with a notable portion misclassified as CGM. Despite these misclassifications, the HFFIncep-Net demonstrates overall strong performance on real-world Cassava images, showcasing its ability to learn discriminative features across multiple disease classes.
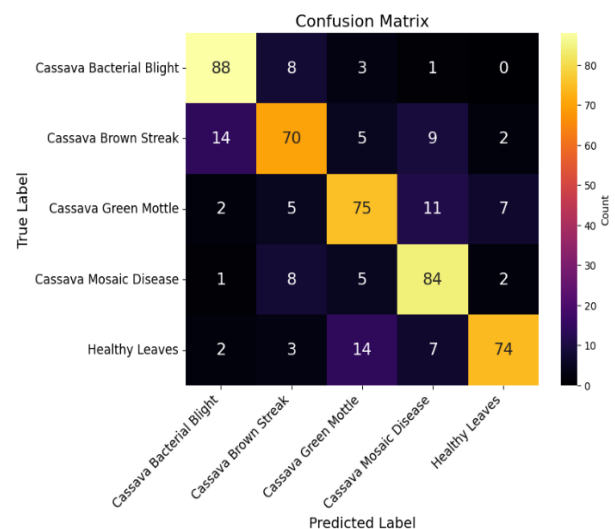

Figure 24: Confusion matrix of HFFIncep-Net for cassava diseases.

4.2   Performance comparative analysis

To evaluate the effectiveness of the proposed HFFIncep-Net, a comprehensive comparison was conducted against several well-established state-of-the-art architectures, focusing on key performance metrics such as precision, recall, F1-score, and accuracy.

MobileNet is significantly lightweight compared to AlexNet, ResNet-50, and DenseNet-201, making it highly efficient for deployment on mobile devices. In terms of performance, MobileNet achieves a precision of 81.04%, a recall of 76.74%, an F1-score of 77.13%, and an accuracy of 90.94%. It provides a strong balance between efficiency and accuracy, making it suitable for real-time applications. AlexNet architecture, which consists of 8 layers in which 5 are convolutional and 3 are fully connected layers. It utilizes 3×3 filters with a stride of 1 to capture the features in the image. AlexNet has approximately 61 million parameters and the dropout

layer is used to minimize the risk of overfitting. In the metrics provided, AlexNet achieves a precision of 81.55%, a recall of 78.02%, F1 score of 76.14%, and an accuracy of 90.98%.

ResNet-50 addresses the challenges of training deeper networks by introducing residual blocks with skip connections, effectively handling the gradient vanishing problem and class variation problem (Inter and Intra class variation) by extracting features at multiple levels with different strides. And Resnet-50 has a total of 25 million parameters which means it is lighter than Alexnet and requires less computation. Resnet-50 is preforming effectively than the Alexnet with the precision of 82.08%, recall of 77.76%, F1 score of 78.51%, and accuracy of 91.28%. This improvement is because of multilevel feature extraction which offers better learning and accuracy.

Densely Connected Convolutional Networks (DenseNet) feature a dense connectivity pattern where each layer is connected to other layers, which enhances feature propagation and reuse. This architecture also addresses the class variation and gradient vanishing problem as features are extracted at multiple levels with different strides (s-1, s-2, e.g.) DenseNet-201 achieved a precision of 85.36%, recall of 79.49%, F1 score of 79.31%, and accuracy of 92.16%, outperforming ResNet-50.

PDD-Net is built on the VGG16 architecture, which is known for its straightforward design with 16 layers, including convolutional filters of 3 × 3. By using VGG16 as its base, PDD-Net benefits from strong feature extraction capabilities while keeping the model efficient. The PDD-Net outperforms the other architecture, with a precision of 86.27%, recall of 86.41%, F1 score of 86.29%, and accuracy of 94.59%.

Inception V3 architecture utilizes Inception modules, which perform parallel convolutions and max pooling operations with different filters to extract features at multiple scales that address the RoI problem. The Inception V3 has approximately 23.9 million parameters. Inception V3 focuses on computational efficiency while maintaining high performance. It achieves a precision of 88.22%, a recall of 86.08%, F1 score of 82.21%, and accuracy of 94.74%.

The proposed HFFIncep-Net is designed for plant disease classification, employing multilevel and multiscale feature extraction techniques to avoid problems like RoI scale variation, gradient vanishing, and class variation problem and then flows into a GAP layer, which compresses it into a more manageable size [20]. The proposed model has approximately 1.2 million parameters, making it lightweight than other state-of-the-art architectures. To mitigate overfitting, a dropout layer with a rate of 0.2 is applied before the final classification layer. The proposed model achieves the highest performance across all other architectures, with a precision of 94.61%, recall of 94.31%, F1 score of 94.31%, and accuracy of 97.80%. Table 3 shows the comparison of proposed HFFIncep-Net model with current state-of-the-art classifiers.

5   Conclusions and future work

This study presents the HFFIncep-Net model, designed to address critical challenges in plant disease detection, including variables like RoI scale variations and intra-class variations. This model leverages a combination of advanced feature extraction techniques, and an activation function optimized to mitigate the issue of inactive neurons. As a result, it demonstrates superior accuracy compared to traditional architecture such as VGG16, AlexNet, and conventional Inception-based models. The findings reveal that the hybrid approach effectively captures essential features across varying scales and depths, enabling precise identification of plant diseases even in complex and dynamic real-world environments. This innovation holds significant potential for empowering farmers and agricultural experts to diagnose crop issues swiftly, safeguard yields, and contribute to a stable and secure food supply chain. Future research directions could involve incorporating additional environmental data or experimenting with state-of-the-art attention mechanisms. By continuously enhancing these models, automated plant disease detection can become faster, more reliable, and widely accessible, benefiting farming communities worldwide.

Table 3: Comparative analysis of proposed model with different State-of-the-art CNN Architecture.

Framework	CNN Architecture	Precision %	Recall %	F1 Score %	Accuracy %
[23]	MobileNet [36]	81.04	76.74	77.13	90.94
[21], [29]	AlexNet [37]	81.55	78.02	76.14	90.98
[26]	ResNet-50 [38]	82.08	77.76	78.51	91.28
[22]	DenseNet-201 [39]	85.36	79.49	79.31	92.16
[24], [25],[19], [27], [32]	VGG16 [40]	86.27	86.41	86.29	94.59
[24], [28], [23], [29]	Inception V3 [35]	88.22	86.08	82.21	94.74
Proposed Model	HFFIncep-Net	94.61	94.31	94.31	97.80

References

[1] F. Taghikhah, A. Voinov, N. Shukla, T. Filatova, and M. Anufriev, "Integrated modeling of extended agro-food supply chains: A systems approach," *Eur J Oper Res*, vol. 288, no. 3, pp. 852–868, Feb. 2021, doi: 10.1016/j.ejor.2020.06.036.

[2] D. Imami, "DRINI IMAMI VLADISLAV VALENTINOV ENGJELL SKRELI Food Safety and Value Chain Coordination in the Context of a Transition Economy: The Role of Agricultural Cooperatives," *Int J Commons*, vol. 15, no. 1, pp. 21–34, doi: 10.5334/DRINI.

[3] Y. Shang, M. Kamrul Hasan, G. J. Ahammed, M. Li, H. Yin, and J. Zhou, "Applications of nanotechnology in plant growth and crop protection: A review," 2019, *MDPI AG*. doi: 10.3390/molecules24142558.

[4] D. Bass, G. D. Stentiford, H. C. Wang, B. Koskella, and C. R. Tyler, "The Pathobiome in Animal and Plant Diseases," Nov. 01, 2019, *Elsevier Ltd*. doi: 10.1016/j.tree.2019.07.012.

[5] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease detection and classification by deep learning," Nov. 01, 2019, *MDPI AG*. doi: 10.3390/plants8110468.

[6] U. Shruthi, V. Nagaveni, and B. K. Raghavendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," in *2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019*, Institute of Electrical and Electronics Engineers Inc., Mar. 2019, pp. 281–284. doi: 10.1109/ICACCS.2019.8728415.

[7] M. Nagaraju and P. Chawla, "Systematic review of deep learning techniques in plant disease detection," *International Journal of System Assurance Engineering and Management*, vol. 11, no. 3, pp. 547–560, Jun. 2020, doi: 10.1007/s13198-020-00972-1.

[8] X. E. Pantazi, D. Moshou, and A. A. Tamouridou, "Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers," *Comput Electron Agric*, vol. 156, pp. 96–104, Jan. 2019, doi: 10.1016/j.compag.2018.11.005.

[9] M. Muhathir, W. Hidayah, and D. Ifantiska, "Utilization of Support Vector Machine and Speeded up Robust Features Extraction in Classifying Fruit Imagery," *Computer Engineering and Applications Journal*, vol. 9, no. 3, pp. 183–193, Oct. 2020, doi: 10.18495/COMENGAPP.V9I3.347.

[10] F. Mohameth, C. Bingcai, K. A. Sada, F. Mohameth, C. Bingcai, and K. A. Sada, "Plant Disease Detection with Deep Learning and Feature Extraction Using Plant Village," *Journal of Computer and Communications*, vol. 8, no. 6, pp. 10–22, Jun. 2020, doi: 10.4236/JCC.2020.86002.

[11] S. Iniyan, R. Jebakumar, P. Mangalraj, M. Mohit, and A. Nanda, "Plant Disease Identification and Detection Using Support Vector Machines and Artificial Neural Networks," *Advances in Intelligent Systems and Computing*, vol. 1056, pp. 15–27, 2020, doi: 10.1007/978-981-15-0199-9_2.

[12] K. Hanbay, "Hyperspectral image classification using convolutional neural network and two-dimensional complex Gabor transform," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 35, no. 1, pp. 443–456, 2020, doi: 10.17341/gazimmfd.479086.

[13] B. S. Kusumo, A. Heryana, O. Mahendra, and H. F. Pardede, "Machine Learning-based for Automatic Detection of Corn-Plant Diseases Using Image Processing," *2018 International Conference on Computer, Control, Informatics and its Applications: Recent Challenges in Machine Learning for Computing Applications, IC3INA 2018 - Proceeding*, pp. 93–97, Jul. 2018, doi: 10.1109/IC3INA.2018.8629507.

[14] K. Golhani, S. K. Balasundram, G. Vadamalai, and B. Pradhan, "A review of neural networks in plant disease detection using hyperspectral data," Sep. 01, 2018, *China Agricultural University*. doi: 10.1016/j.inpa.2018.05.002.

[15] Z. Tianyu, M. Zhenjiang, and Z. Jianhu, "Combining CNN with Hand-Crafted Features for Image Classification," *International Conference on Signal Processing Proceedings, ICSP*, vol. 2018-August, pp. 554–557, Feb. 2019, doi: 10.1109/ICSP.2018.8652428.

[16] S. Zhang, S. Zhang, C. Zhang, X. Wang, and Y. Shi, "Cucumber leaf disease identification with global pooling dilated convolutional neural network," *Comput Electron Agric*, vol. 162, pp. 422–430, Jul. 2019, doi: 10.1016/J.COMPAG.2019.03.012.

[17] J. Zhang, Y. Rao, C. Man, Z. Jiang, and S. Li, "Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things," *Int J Distrib Sens Netw*, vol. 17, no. 4, 2021, doi: 10.1177/15501477211007407.

[18] Y. Zhong, M. Zhao, Y. Zhong, and M. Zhao, "Research on deep learning in apple leaf disease recognition," *CEAgr*, vol. 168, p. 105146, Jan. 2020, doi: 10.1016/J.COMPAG.2019.105146.

[19] H. Alghamdi and T. Turki, "PDD-Net: Plant Disease Diagnoses Using Multilevel and Multiscale Convolutional Neural Network Features," *Agriculture (Switzerland)*, vol. 13, no. 5, May 2023, doi: 10.3390/agriculture13051072.

[20] M. H. Ashraf and H. Alghamdi, "HFF-Net: A hybrid convolutional neural network for diabetic retinopathy screening and grading," *Biomedical Technology*, vol. 8, pp. 50–64, Dec. 2024, doi: 10.1016/j.bmt.2024.09.004.

[21] A. S. Keceli, A. Kaya, C. Catal, and B. Tekinerdogan, "Deep learning-based multi-task prediction system for plant disease and species detection," *Ecol Inform*, vol. 69, Jul. 2022, doi: 10.1016/j.ecoinf.2022.101679.

[22] W. Chen, J. Chen, R. Duan, Y. Fang, Q. Ruan, and D. Zhang, "MS-DNet: A mobile neural network for plant disease identification," *Comput Electron Agric*, vol. 199, Aug. 2022, doi: 10.1016/j.compag.2022.107175.

[23] D. Sutaji and O. Yıldız, "LEMOXINET: Lite ensemble MobileNetV2 and Xception models to predict plant disease," *Ecol Inform*, vol. 70, Sep. 2022, doi: 10.1016/j.ecoinf.2022.101698.

[24] T. Sanida, A. Sideris, M. V. Sanida, and M. Dasygenis, "Tomato leaf disease identification via two–stage transfer learning approach," *Smart Agricultural Technology*, vol. 5, Oct. 2023, doi: 10.1016/j.atech.2023.100275.

[25] S. G. Paul *et al.*, "A real-time application-based convolutional neural network approach for tomato leaf disease classification," *Array*, vol. 19, Sep. 2023, doi: 10.1016/j.array.2023.100313.

[26] C. Zhou, S. Zhou, J. Xing, and J. Song, "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network," *IEEE Access*, vol. 9, pp. 28822–28831, 2021, doi: 10.1109/ACCESS.2021.3058947.

[27] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, Jan. 2022, doi: 10.1016/j.aiia.2021.12.002.

[28] R. Pan *et al.*, "A two-stage feature aggregation network for multi-category soybean leaf disease identification," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 8, Sep. 2023, doi: 10.1016/j.jksuci.2023.101669.

[29] S. Zhang, D. Wang, and C. Yu, "Apple leaf disease recognition method based on Siamese dilated Inception network with less training samples," *Comput Electron Agric*, vol. 213, Oct. 2023, doi: 10.1016/j.compag.2023.108188.

[30] G. Dai, Z. Tian, J. Fan, C. K. Sunil, and C. Dewi, "DFN-PSAN: Multi-level deep information feature fusion extraction network for interpretable plant disease classification," *Comput Electron Agric*, vol. 216, Jan. 2024, doi: 10.1016/j.compag.2023.108481.

[31] G. Dai, J. Fan, Z. Tian, and C. Wang, "PPLC-Net:Neural network-based plant disease identification model supported by weather data augmentation and multi-level attention mechanism," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 5, May 2023, doi: 10.1016/j.jksuci.2023.101555.

[32] P. S. Thakur, T. Sheorey, and A. Ojha, "VGG-ICNN: A Lightweight CNN model for crop disease identification," *Multimed Tools Appl*, vol. 82, no. 1, pp. 497–520, Jan. 2023, doi: 10.1007/s11042-022-13144-z.

[33] M. H. Ashraf, F. Jabeen, H. Alghamdi, M. S. Zia, and M. S. Almutairi, "HVD-Net: A Hybrid Vehicle Detection Network for Vision-Based Vehicle Tracking and Speed Estimation," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 8, Sep. 2023, doi: 10.1016/j.jksuci.2023.101657.

[34] A. D. Mohammed and D. Ekmekci, "Breast Cancer Diagnosis Using YOLO-Based Multiscale Parallel CNN and Flattened Threshold Swish," *Applied Sciences 2024, Vol. 14, Page 2680*, vol. 14, no. 7, p. 2680, Mar. 2024, doi: 10.3390/APP14072680.

[35] C. Szegedy *et al.*, "Going Deeper with Convolutions."

[36] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017, [Online]. Available: http://arxiv.org/abs/1704.04861

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." [Online]. Available: http://code.google.com/p/cuda-convnet/

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03385

[39] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," Aug. 2016, [Online]. Available: http://arxiv.org/abs/1608.06993

[40] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.1556