Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 9, No. 4, August, 2025

Empoasca Pest Attack Classification on Tea Plantations Using Multispectral Imaging and Deep Learning

Bella K. Kurniawan [1], Fiky Y. Suratman [1], Fani Fauziah [2], Hurianti Vidyaningtyas [1*], Desti M. Saputri [1], Aloysius A. Pramudita [1]

[1] *The University Center of Excellence for Intelligent Sensing-IOT, Telkom University, Bandung, Indonesia.*

[2] *Indonesia Research Institute for Tea and Cinchona (IRITC) Gambung, Bandung 40972, Indonesia.*

Abstract

This study aims to enhance the management of Empoasca pests in tea cultivation, a critical sector for Indonesia's economy, by developing an innovative detection method. The challenge of pest infestations may significantly reduce tea production yields, and the misuse of chemical pesticides further compromises tea quality. We propose a novel approach that integrates multispectral imaging with Convolutional Neural Networks (CNN), specifically employing ResNet-50 and AlexNet architectures to accurately detect Empoasca infestations. We begin with the data collection process, followed by the development of the preprocessing model and evaluation of its performance. We classify tea leaves affected by Empoasca pests using spectral data obtained from a multispectral camera operating across Green, NIR (Near Infrared), REG (Red Edge), and RED channels. We evaluated various spectral channels and identified the green spectrum as the most effective for revealing visual characteristics, such as curled leaves associated with Empoasca damage. Experimental results demonstrated that ResNet-50 outperformed AlexNet, achieving a remarkable accuracy of 99% on the green channel, while AlexNet showed notable accuracy declines on other channel combinations. These findings underscore the effectiveness of the green spectrum and the superiority of ResNet-50 in achieving precise pest detection, offering a reliable technological solution for modern tea plantation management.

Keywords:

Empoasca Pest Detection;
Multispectral Imaging;
Green Spectrum Analysis;
AlexNet;
ResNet-50;
Tea Plantation Pest Management.

Article History:

Received:	11	February	2025
Revised:	25	May	2025
Accepted:	09	June	2025
Published:	01	August	2025

1- Introduction

Tea (Camellia sinensis) flourishes in tropical and subtropical regions characterized by acidic soil conditions and is widely recognized for its health benefits [1]. However, climate variability and change have heightened the risk of infestations from the green leafhopper (Empoasca sp.) in tea plantations throughout Indonesia [2]. Empoasca pests are sap-sucking insects that cause damage to tea leaves, leading to the curling of the affected foliage. Currently, farmers manually inspect tea plants for signs of infestation, but since some symptoms of pests and diseases can appear similar, they often apply pesticides indiscriminately to infested plants. Given that most tea plants are spread across vast areas, timely detection of infested leaves becomes a lengthy process. Accurate and prompt identification is critical for effective management and prevention of infestations, highlighting the need for advanced detection methods using multispectral imaging to safeguard plant production.

Radar-drone systems have shown promise in monitoring soil moisture content across large plantation areas [3]. However, the presence of vegetation covering the soil surface can interfere with the detection of soil moisture using radar systems, as vegetation often influences radar signal propagation. Thus, a method to mitigate this effect is essential. Narmilan et al. (2022) [4] utilized algorithms such as Extreme Gradient Boosting (XGBoost), Random Forest (RF),

* **CONTACT**: huriantividya@telkomuniversity.ac.id

DOI: http://dx.doi.org/10.28991/ESJ-2025-09-04-04

Decision Tree (DT), and K-Nearest Neighbors (KNN), combined with a multispectral camera mounted on an Unmanned Aerial Vehicle (UAV), to detect White Leaf Disease (WLD) in sugarcane leaves, achieving an accuracy of 94% for detecting individual leaves. Convolutional Neural Networks (CNN), have been utilized to process UAV imagery and multispectral data for improving tree species recognition and fruit tree classification, achieving accuracy 92% through effective image fusion techniques [5].

Traditional object recognition techniques, including Scale-Invariant Feature Transform (SIFT) [6], Speeded-Up Robust Features (SURF) [7], and Histograms of Oriented Gradients (HOG) [8], have been extensively applied for detecting and classifying objects. While SIFT is highly effective in object recognition tasks, it involves significant computational complexity [6]. SURF, on the other hand, operates faster than SIFT and can identify key points without overlooking other objects [7]. HOG serves as a feature descriptor for object detection; however, it primarily focuses on computing the gradients present in the image [8]. Additionally, machine learning techniques, such as Support Vector Machines (SVM), have been employed for leaf disease detection [9]. The combination of SVM with K-Nearest Neighbors (KNN) may enhance classification accuracy [10], but tuning parameters with kernel functions in SVM can be challenging for achieving optimal performance. In recent years, the application of machine learning techniques in agriculture has increased, enabling automated solutions for robust plant disease identification and early detection, which are crucial for enhancing productivity and minimizing losses. By employing artificial intelligence and computer vision methodologies, these automated solutions offer a more efficient and accurate disease detection compared to traditional manual monitoring practices [11].

Several CNN architectures, as a subset of machine learning methods, are widely utilized for image recognition and classification [12]. CNN has also been employed to identify diseases across various crops, including apples, corn, grapes, tomatoes, and potatoes, with an accuracy of 86% [13]. However, it only recognizes single-leaf disease identification and has not been used for multi-leaf cases. To address these gaps, this research proposes a method for detecting Empoasca infestations on tea leaves using multispectral imaging with ResNet-50 and AlexNet architectures for data classification to obtain high accuracy. Multispectral cameras were selected for this study due to their ability to capture various spectral features, then facilitating the processing of data with ResNet-50 and AlexNet. These CNN architectures are particularly suitable for addressing issues related to large datasets. Thus, this research will implement ResNet-50 and AlexNet architectures. Unlike previous research focused on single-leaf detection, our approach leverages multispectral imaging to cover larger plantation areas, enhancing detection efficiency and scalability. Through the proposed method, we aim to provide farmers with an accurate and practical tool for identifying pest infestations, enhancing pest management, and sustaining tea production.

This article is organized as follows: Section 2 provides a comprehensive review of relevant literature, detailing previous research on pest detection methods and the application of multispectral imaging in agriculture. In Section 3, we outline the materials and methodologies employed, including the specifics of the multispectral imaging setup and the CNN architectures utilized. Section 4 presents the results and findings of the experiments conducted, highlighting the effectiveness of the proposed method in detecting Empoasca infestations. Finally, Section 5 discusses the conclusions of our findings for pest management in tea plantations, along with recommendations for future research.

2- Literature Review

2-1- Empoasca Infestation on Tea Plants

Currently, Indonesia ranks as the seventh largest tea-producing country in the world, with a tea plantation area of 111,116 hectares and a production total of 129,832 tons. A significant portion of Indonesia's tea production is exported, while the remainder is sold domestically, resulting in a total export value of USD 147 million in 2019 [14]. However, plant pests can adversely affect tea production, which is a vital agricultural commodity in the Indonesian economy. The Empoasca sp. insect, commonly known as the tea leafhopper or green leafhopper, poses a significant threat to tea plants, as illustrated in Figure 1-a. These pests primarily infest the undersides of tea leaves, particularly on new shoots. Under certain environmental conditions, Empoasca may also target the upper surfaces of leaves, although this behavior tends to be transient, as shown in Figures 1-b and Figure 1-c. A high population of Empoasca can cause considerable damage, leading to a color change in tea shoots from pale green to yellowish before ultimately drying out and dying. Such severe infestations can reduce tea shoot production by as much as 50% [15].

(a) (b) (c)

Figure 1. (a) Infestation of Empoasca on tea plants, (b) Empoasca infestation on the lower leaf surface, (c) Empoasca infestation on the upper leaf surface

Empoasca infestations can be classified based on severity, indicated by distinct visual symptoms [16]:

- Moderate infestation: The edges of the leaves curl, and numerous pests are visible on the undersides of the leaves.
- Severe infestation: Young leaves turn dull yellow, curl extensively, and eventually die.

These visual symptoms caused by Empoasca allow for the identification of infestations with the naked eye. However, manual monitoring often encounters challenges due to the extensive size of tea plantations and limited labor resources. Additionally, inconsistent monitoring can impede timely pest management.

To overcome these limitations, digital imagery has emerged as a vital tool for identifying the unique characteristics of tea leaves affected by Empoasca infestations. Digital imaging facilitates more accurate and efficient pest monitoring by capturing high-resolution images of the affected leaves. This approach aids in better tracking of pest movement and spread within plantations, enabling proactive control measures. Consequently, digital imaging has the potential to significantly improve pest management practices and protect tea crops from the detrimental effects of Empoasca. This proposed detection method can also be integrated into drone systems, allowing for large-area inspections of tea plantations.

2-2- Concept of Convolutional Neural Network (CNN)

CNN is the primary model utilized for their proven ability to automatically extract relevant features from images with minimal preprocessing. The architecture of CNN is well-suited to handle multispectral data, as it can simultaneously process spatial and spectral information. CNN is fundamental to deep learning, widely recognized for their capacity to automate feature extraction and classification without manual intervention. Unlike traditional artificial neural networks that require extensive parameter tuning, CNN is designed to process two-dimensional input data, making them particularly effective for image-based tasks. They can recognize new objects within an existing network, providing versatility across various applications. However, one drawback of CNN is their long training time, although this limitation can be mitigated with modern hardware advancements [17]. Furthermore, CNN is more efficient than traditional neural networks, as they require fewer parameters, resulting in faster processing and easier implementation [18]. Specifically designed for two-dimensional input data such as images, CNN architecture employs convolutional operations where weight parameters are organized into sets known as convolutional kernels. These kernels enable CNN to capture spatial hierarchies in data, such as edges, textures, and patterns, allowing for the efficient transformation of raw pixel values into meaningful features, which ultimately leads to accurate classifications [19]. The architecture of neural networks consists of multiple layers, each playing a critical role in processing input data and extracting features. Deep learning has introduced network architectures capable of handling complex input data and transforming it into a more comprehensible format. Modern CNN encompass multiple layers, including convolutional, pooling, and fully connected layers, which map input data to class scores [20-22]. Hyperparameters, such as the weights and biases of neurons, are crucial for optimizing the performance of CNN [23]. Additionally, the depth of the network (the number of layers) and the dimensions (width and height) of each layer significantly influence the model's ability to generalize and accurately classify data [24]. An example of CNN architecture for image classification is illustrated in Figure 2.

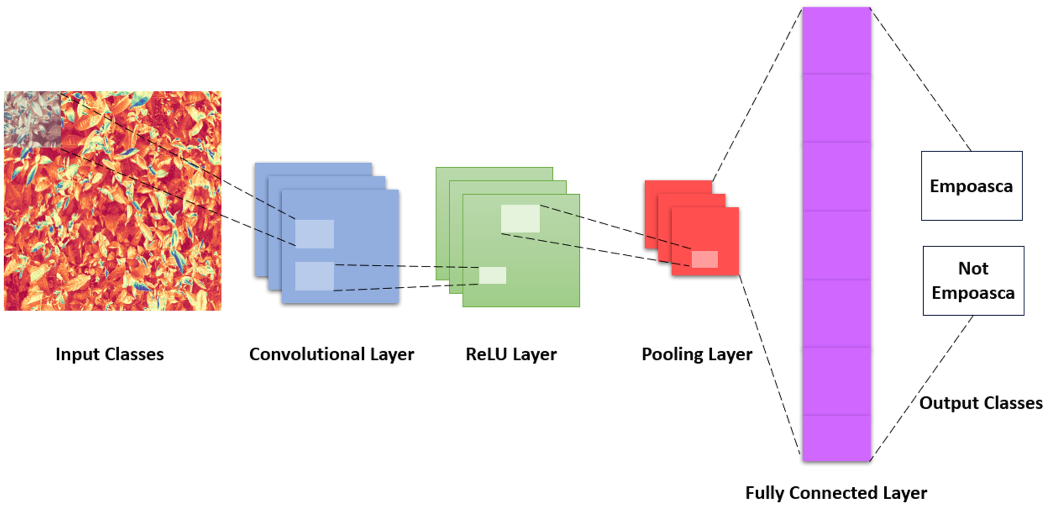

Figure 2. Convolutional Neural Network

2-2-1- AlexNet

One of the significant milestones in the development of CNN is the AlexNet architecture, which was first proposed by Krizhevsky et al. AlexNet demonstrated that deep CNN architectures could be effectively utilized for complex image recognition tasks by increasing network depth and implementing various parameter optimization techniques. It achieved a remarkably high level of classification accuracy on the ImageNet dataset, marking a major breakthrough in the field of machine learning [25]. AlexNet provides a simpler yet effective architecture that has been widely adopted for various image classification tasks [26]. AlexNet architecture comprises convolutional, pooling, and fully connected layers, all designed to extract features and perform classification, enabling efficient processing while retaining important information from the images. Figure 3 illustrates the basic structure of AlexNet.

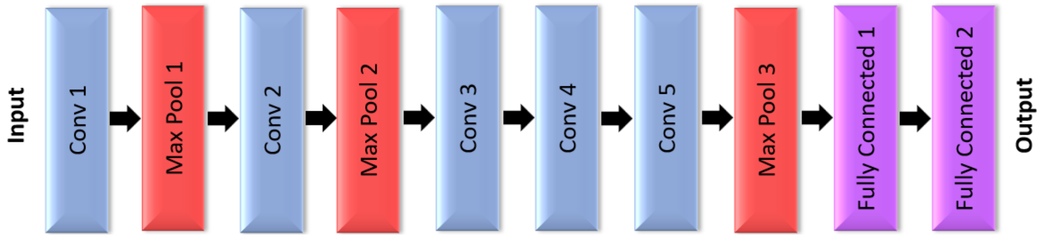

Figure 3. Basic Design of Alexnet Architecture

CNN architectures, such as AlexNet, contain over 60 million parameters, making them robust and effective deep learning systems suitable for tasks involving large datasets. However, in situations with sparse datasets, like the Google Jigsaw Dataset, these models are susceptible to overfitting, which can result in suboptimal performance. To address this issue, various regularization strategies are employed to reduce overfitting [27-29].

2-2-2- ResNet-50

Artificial neural networks are increasingly utilized for image categorization and processing, enhancing both accuracy and efficiency in image classification [30]. Deep learning models consist of multiple layers of artificial neurons that automatically extract significant features from raw data, making them superior to traditional machine learning models in various tasks [31]. One of the most well-known architecture in deep learning is ResNet-50, which comprises 50 layers and is specifically designed to address the vanishing gradient problem often encountered in very deep networks. ResNet-50 was developed to combat accuracy degradation in deeper artificial neural networks, where performance tends to decline as the number of layers increases [32]. Figure 4 illustrates the fundamental concept of ResNet, which involves the use of residual learning blocks that incorporate shortcut connections, minimizing the loss of important features during the convolution process [33].

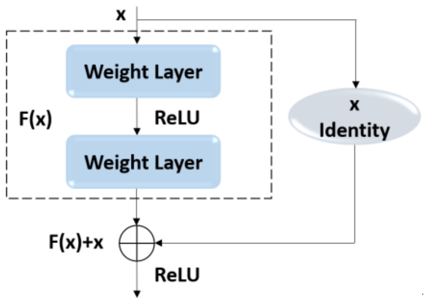

Figure 4. Residual Learning Block

The main components of the ResNet-50 architecture include convolutional layers, identity blocks, convolutional blocks, residual learning blocks, and fully connected layers. Each convolutional layer undergoes batch normalization and applies an activation function. These layers are crucial for detecting significant features in the input images, such as shapes, edges, and textures. Following the convolutional layers, a max pooling layer is employed to reduce the size of the image while retaining essential features. Additionally, convolutional and identity blocks are fundamental elements of the ResNet-50 architecture, with the identity block ensuring that the input and output dimensions remain consistent [34]. The structure of the ResNet-50 model is depicted in Figure 5 [33]. Based on studies [26, 35, 36], the ResNet-50 architecture has been applied to classify different types of leaves, yielding impressive accuracy results. ResNet-50 has demonstrated effectiveness in classifying plant-related datasets with high precision, particularly for tasks that require detailed feature extraction from image datasets.

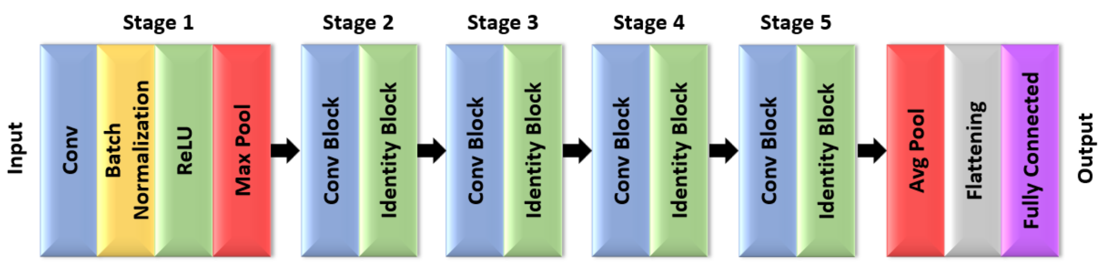

Figure 5. Basic Design of ResNet-50 Architecture

2-3- Related Work

Table 1 presents the state of the art related to this research. Based on a review of studies using cameras and the PlantVillage dataset on a single leaf indicates that it is possible to design a detection method for large tea plantations affected by Empoasca pests, characterized by curling leaves, using multispectral cameras. Data processing techniques employed to analyze images captured using multispectral cameras, transforming them into data that can be processed using AlexNet and ResNet-50 to get accuracy and loss. The method used in this research is the initial phase in developing technological concepts in agriculture, particularly for tea leaves impacted by Empoasca pests.

Table 1. State of the Art

Author	Data Collection	Observed Object	Algorithm	Accuracy
Narmilan et al. (2022) [4]	1. Collected using a drone DJI P4 2. Captured dataset over a large area	Healthy leaves, light infestion, and severe infections	1. XGBoost (Extreme Gradient Boosting) 2. Random Forest (RF) 3. Decision Tree (DT) 4. K-Nearest Neighbors (KNN)	94%
Hari et al. (2019) [13]	1. Dataset from PlantVillage 2. Leaf shooting on single leaves	Apple, corn, grape, tomato, and potato leaves	Convolutional Neural Network (CNN)	86%
Ramdan et al. (2019) [26]	1. Collected using a smartphone camera 2. Leaf shooting on single leaves	GMB series, which GMB 1, GMB 2, GMB 3, GMB 4, and GMB 5	1. ResNet 2. VggNet 3. AlexNet	1. ResNet = 97.84% 2. VGGNet = 32.08% 3. AlexNet = 98.86%
Andrew et al. (2022) [35]	1. Dataset from PlantVillage 2. Leaf shooting on single leaves	14 different plant types (e.g., apple, tomato, orange, grape, potato leaves, etc)	Tranfer learning with DesNet-121, ResNet-50, VGG-16 and Inception V4	99.81%
Kanda et al. (2022) [36]	1. Dataset from PlantVillage 2. Leaf shooting on single leaves	Healthy tomato leaves and tomato leaves affected by diseases	ResNet-18, ResNet-34, ResNet-50, ResNet 101 and ResNet-152 architectures	99.5%
Palma et al. (2022) [37]	1. Collected using a camera (Sony DSC-RX100/13 20.2 MP) 2. Leaf shooting on single leaves	Apple, grape, potato, peach, and healthy leaves	1. Segmentation 2. Disease detection features based on different intensities or colors	98.87%
Fahmi et al. (2019) [38]	1. Collected using a smartphone camera 2. Leaf shooting on single leaves	Disease affected chili leaves	1. Gray Co-occurrence Matrix (GLCM) 2. Support Vector Machine (SVM)	88%
Ahmad et al. (2021) [39]	1. Dataset from PlantVillage 2. Leaf shooting on single leaves	Apples, bell peppers, cherries, corn, grapes, oranges, peaches, potatoes, pumpkins, strawberries, tomatoes leaves	1. Relief algorithm forward 2. Feature selection.	99.13%
Nasra & Gupta (2024) [40]	1. Dataset from ImageNet 2. Leaf shooting on single leaves	Healthy banana leaves and banana leaf spot diseases	ResNet-50	96%
Jain & Sharma (2024) [41]	1. Dataset from AI-Lab Makerere titled "Bean Disease Dataset." 2. Captured in bean fields all over Uganda	Angular leaf spot, bean Rust, Healthy leaves	ResNet-50	88%
Senthil Pandi et al. (2024) [42]	Dataset from PlantVillage	Healthy and diseased leaf (38 different types of plant diseases)	1. CNN 2. ResNet-50 3. VGG-16 4. AlexNet	1. CNN=88.45% 2. ResNet-50 = 91.23% 3. VGG-16 = 89.65% 4. AlexNet = 94.50%
Bharti et al. (2024) [43]	1. Dataset from PlantVillage and Kannauj Village 2. Leaf shooting on single leaves	Potato Pest, Potato Phytophthora, Potato Virus, Potato Late Blight, Potato Early Blight, and Potato Healthy	1. CNN 2. ResNet-50	1. CNN=94.29% 2. ResNet-50 = 98.36%
Arora et al. (2024) [44]	1. Dataset from PlantVillage and Kannauj Village 2. Leaf shooting on single leaves	healthy and diseased plants	1. VGG-16 2. Inception V4 3. AlexNet 4. ResNet-50	1. VGG-16=81.63% 2. Inception V4=98.36% 3. AlexNet=95.10% 4. ResNet-50=99.7%
Present Study	1. Dataset from the Tea and Kina Research Center in Gambung 2. Leaf shooting on groups of leaves	Empoasca pest infestations in tea plantations	1. CNN 2. AlexNet 3. ResNet-50	Single Channel: 1. CNN=76% 2. AlexNet=99% 3. ResNet-50=99% Four Channel: 1. CNN=69% 2. AlexNet=41% 3. ResNet-50=90%

Table 1 presents data from various studies utilizing machine learning and deep learning techniques. The overall trend indicates that deep learning models, particularly those based on the ResNet architecture, significantly outperform traditional machine learning methods in plant disease detection across different methodologies, datasets, and accuracy levels. These studies emphasize the importance of dataset quality and diversity, as well as the effectiveness of imaging techniques in enhancing detection capabilities.

Many previous studies relied on datasets from PlantVillage, ImageNet, and localized sources, suggesting variability in data quality and diversity. Images were captured using drones, smartphone cameras, and standard cameras, which influences the data's quality and scale. ResNet architectures, especially ResNet-50, consistently demonstrate high accuracy, indicating that deeper networks can effectively capture more features.

Most previous methods have focused on detecting pests on single leaves. In this study, we utilize ResNet-50 and spectral imaging to detect Empoasca pest attacks on groups of leaves. We employed a drone to capture images of these groups of leaves and then used ResNet-50 to process the multispectral imaging data for identifying the Empoasca pest attacks. The results indicate that among the evaluated models, ResNet-50 consistently emerged as the most effective model for handling complex multispectral data, achieving a peak accuracy of 99% when using the green channel.

3- Research Methodology

This study focuses on the classification of multispectral camera image data related to Empoasca pest attacks on tea leaves using the ResNet-50 and Alexnet algorithms. In addition, this study also aims to see the most efficient algorithm with the highest accuracy. It begins with the data collection process and then moves on to the preprocessing model development and performance evaluation methods. The implementation process of the ResNet-50 and Alexnet algorithms in this study is explained in Figure 6.

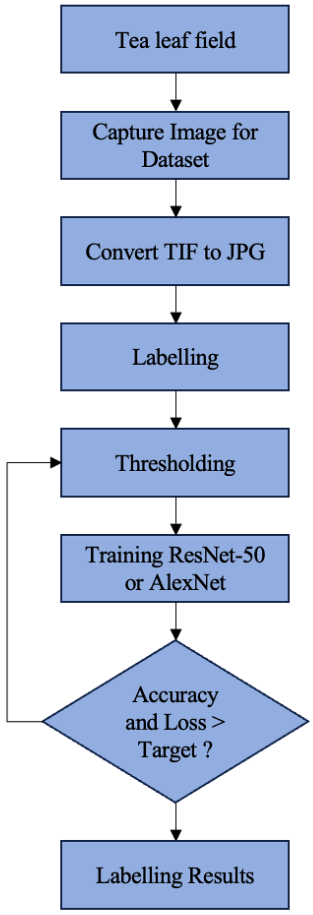

Figure 6. Block Diagram of Data Training Model

3-1- Data Collection

The data collection process for this study was conducted at the Tea and Kina Research Center in Gambung, West Java. The site has extensive tea plantations and diverse environmental conditions, which provide an ideal place to observe different levels of Empoasca infestation. The site offers access to healthy and infested tea plants, making it possible to develop a comprehensive data set. The experimental setup focused on tea leaves from the Camellia sinensis species. Data were collected from several plantation blocks, with sources of healthy leaves, lightly infested leaves, heavily infested leaves, and leaves affected by other pests. Sampling involved experts to determine the sample areas corresponding to each class. The selection of these areas ensures a representative data set covering a wide range of pest conditions. Data were collected under controlled conditions, with consistent lighting and minimal external disturbances.

The study consisted of four classes representing four types of tealeaf conditions: healthy leaves, leaves heavily infested by Empoasca, leaves lightly infested by Empoasca, and leaves with leaf scabs. The dataset of each class contains 300 leaf overlay images where each image is converted into a colour map spectral image to make it easier to distinguish the condition of tea leaves. The colour spectral image in this study consists of 4 channels, namely, green channel, red channel, NIR channel and REG channel. Overall, the total dataset in this research is 4,800 images with each class having 1,200 images. An example of a leaf overlay image for the RGB channel and all four channels is shown in Figure 7.

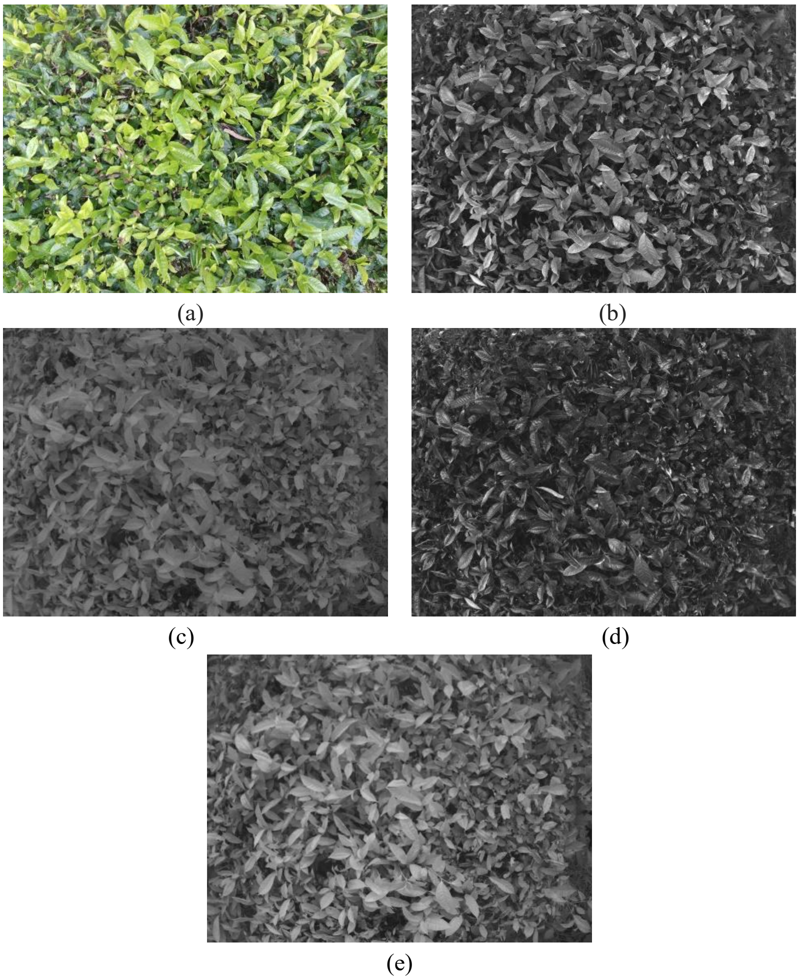

Figure 7. Multispectral camera capture (a) RGB Channel (b) Green Channel (c) NIR Channel (d) RED Channel (e) REG Channel

The image acquisition process for the dataset was carried out using a multispectral camera mounted on a support pole, facing directly down towards the expanse of tea leaves attacked by Empoasca. The results of capturing the dataset using a multispectral camera are shown in Figure 4 which has gone through the conversion process to JPG and through the thresholding process. The final output of each captured image is an image with 4 types of channels, namely Red channel, Green Channel, NIR channel, and REG channel. This multispectral camera is connected to a mobile phone or laptop and is a high-resolution camera that is effective for large-scale image acquisition in tea plantations. An illustration of image capture technique for datasets is shown in Figure 8.

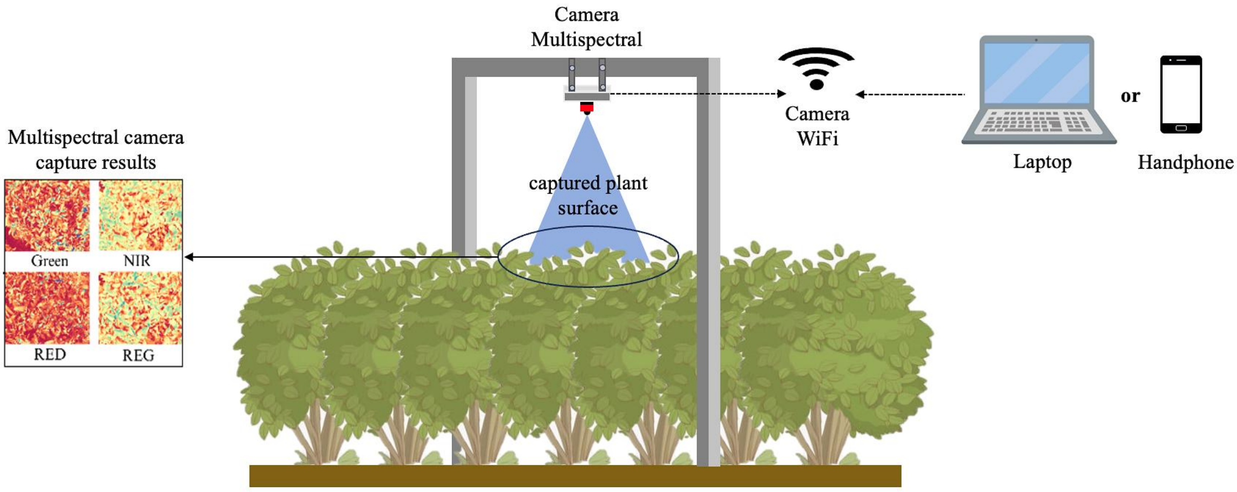

Figure 8. Illustration of image capture technique for datasets

3-2- Preprocessing

Data preprocessing involves converting a raw dataset into a format suitable for training and testing a model. The preprocessing process consists of several stages, including format conversion, dataset labeling, resizing, and spectral image processing.

1. Dataset Format Conversion:

The first stage involves converting the dataset from TIF format to JPG format. This step reduces file size and speeds up computation.

2. Dataset Labeling:

In the second stage, the dataset is labeled by organizing the data into separate folders based on predetermined categories.

3. Dataset Resizing:

In the third stage, all datasets are resized to 256 x 256 pixels. Standardizing the data size ensures the model can more easily recognize patterns in the dataset, resulting in faster computation times.

4. Processing Spectral Images:

The fourth stage, as shown in Figure 9, involves processing spectral images of tea leaves affected by Empoasca pests.

- Channel Extraction: The multispectral camera generates data across four channels: Green, Red, REG, and NIR.
- Spectral Color Map Conversion: The data is converted into a spectral color map to highlight the spectral intensity of each channel, with colors displayed in red, green, and blue.
- Spectral Image Analysis: This step combines information from all spectral channels into a spectral image, which is then analyzed to assess the condition of tea leaves, including detecting signs of *Empoasca* pest infestation

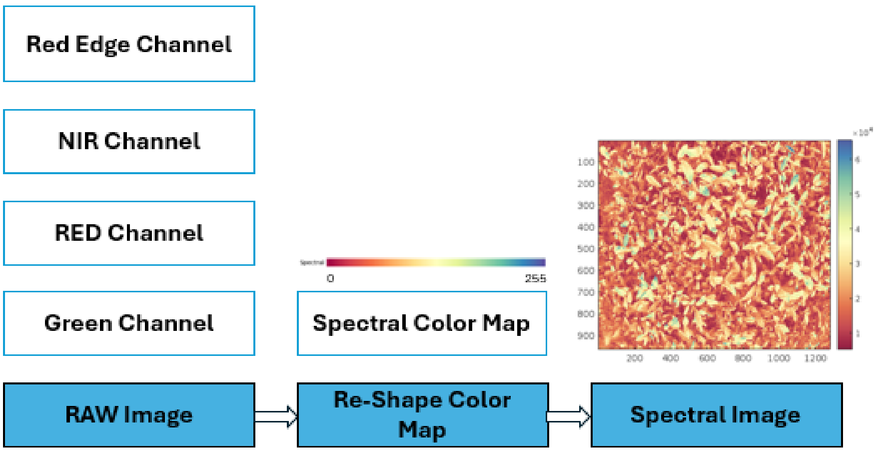

Figure 9. Processing Spectral Image

5. Thresholding Process:

After the image is converted to a spectral format, the fifth step, as shown in Figure 10, is to perform the thresholding process. The purpose of thresholding is to reduce or eliminate noise in the image so that the dataset used in this study contains only images of leaves with visible characteristics. With the proper application of thresholding, important features in the image can be detected and identified more accurately.

- Input Image: The process begins by inputting an image based on four channels: Green, NIR, Red, and Red Edge.
- Determine the Center Value: The center value is calculated to divide pixel intensity into two categories, black and white. A center value of 128 is used as the standard threshold. Pixel intensities less than 128 are set to 0 (black), while intensities equal to or greater than 128 are set to 1 (white).
- Threshold Value for Each Channel: The threshold value for each channel (Green, NIR, Red, and Red Edge) is determined based on the color map values. This threshold forms the basis for the binarization process, which separates the object from the background.

- Application of Thresholding: The thresholding process is applied to each channel (Green, Red, NIR, and Red Edge). For example, when thresholding is applied to the Green channel, an upper bound of 170 is used for the blue color (converted to 1, or white), while a lower bound of 43 is converted to 0 (black).

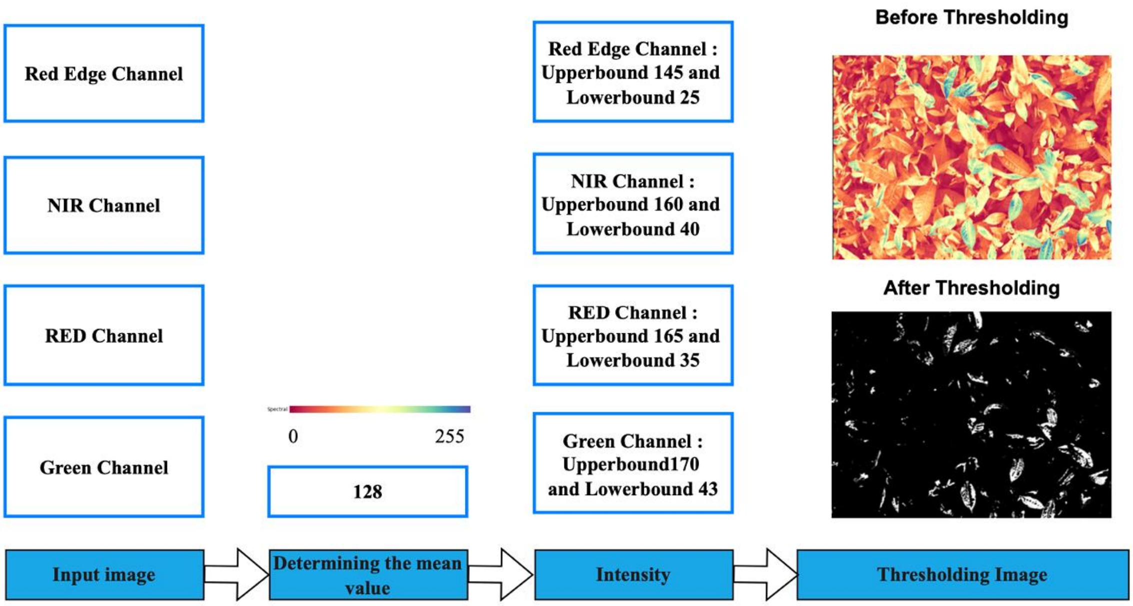

Figure 10. Processing Thresholding

3-3- Classification Models

Intro tentang jenis klasifikasi yang digunakan ResNet-50 and AlexNet are classification models based on Convolutional Neural Networks (CNN) designed to perform pattern recognition tasks in visual data, including classifying tea leaves based on the level of *Empoasca* infestation. As derivatives of CNN, these models utilize convolutional layers to extract essential features from images, such as texture and color, which serve as primary indicators of infestation levels. CNNs are widely recognized for their ability to handle complex image data by leveraging feature hierarchies, ranging from simple features like edges to high-level patterns indicative of infestation on leaves.

ResNet-50, short for Residual Network with 50 layers, addresses the vanishing gradient problem in CNN models by incorporating residual learning mechanisms. This approach enables the network to skip some layers, allowing information to flow more effectively to subsequent layers without losing critical details. This makes ResNet-50 superior in managing more complex and extensive datasets. On the other hand, AlexNet, one of the pioneering CNN models, employs a simpler architecture with only eight layers but remains effective for medium-sized datasets. In this context, ResNet-50 tends to be more suitable for detecting intricate patterns associated with *Empoasca* infestation, while AlexNet offers faster training times but may be less precise in capturing subtle details in tea leaf images.

3-3-1- ResNet-50

One of the models used to classify images of tea leaf fields based on the level of *Empoasca* pest infestation is ResNet-50, a deep learning model based on the Convolutional Neural Network (CNN). This model accepts input in the form of images with a size of 256×256 pixels, which are then processed through several stages of convolutional layers, pooling, and residual blocks. The process begins with an initial 7×7 convolutional layer with 64 filters, which functions to extract initial features from the input image. This is followed by max pooling to reduce the spatial dimensions of the image while retaining critical information.

The architecture consists of 4 main stages, each containing several residual blocks with a certain number of loops. Stage 1 involves 3 loops of residual blocks that use a combination of 1×1 and 3×3 convolutional layers with 64 filters, followed by a 1×1 convolutional layer with 256 filters. Stage 2 increases the feature complexity with 4 loops, utilizing a combination of 1×1 and 3×3 convolutions with 128 filters, ending with a 1×1 convolution containing 512 filters. The residual blocks in each stage allow the model to pass information between layers through shortcut connections, which help address the vanishing gradient problem in deep learning networks.

Stage 3 becomes the most complex phase with 6 loops of residual blocks. This stage includes a 1×1 convolutional layer with 512 filters and a 3×3 convolutional layer with 512 filters, followed by a 1×1 convolution containing 2048 filters. In Stage 4, the number of loops decreases to 3, but the complexity remains high with a combination of 1×1 and

3x3 convolutional layers that have 256 and 1024 filters, respectively. After passing through all stages, the extracted features are processed through an average pooling layer and then passed to a fully connected layer for classification. Finally, the softmax layer is used at the final stage to produce the output, which predicts the infestation category of tea leaves by the *Empoasca* pest. The ResNet-50 architecture used is illustrated in Figure 11.

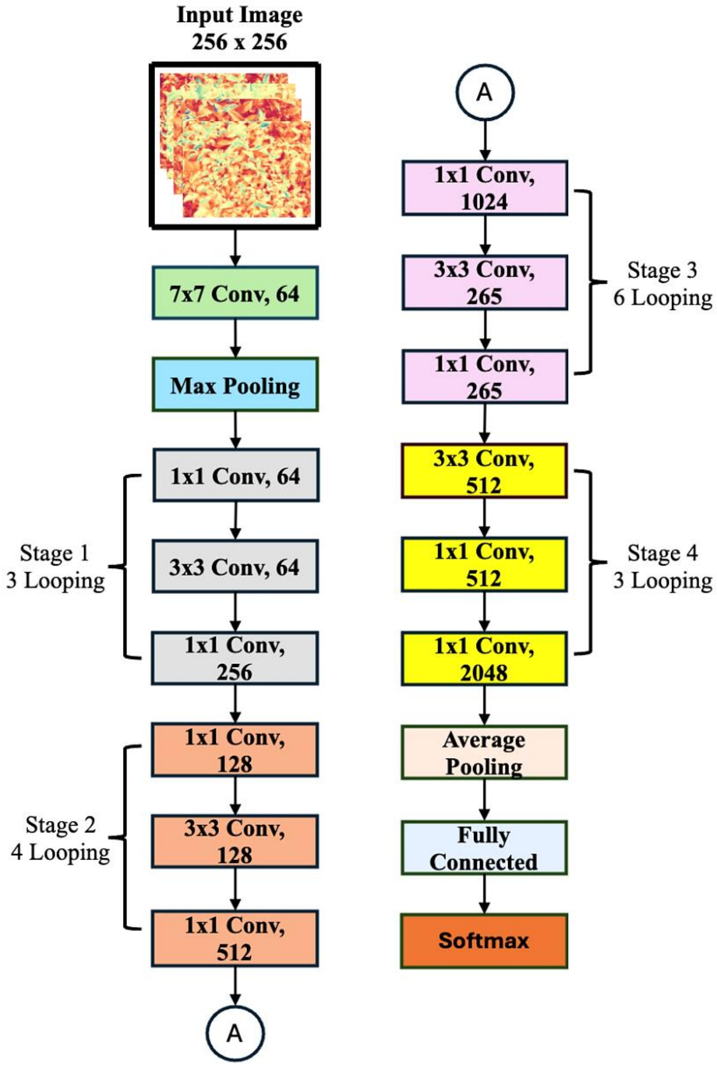

Figure 11. **ResNet-50 Architecture**

3-3-2- AlexNet

The AlexNet architecture is simple yet effective in solving image classification tasks. In this study, AlexNet is used to classify images of tea leaf canopies based on the level of Empoasca pest infestation, utilizing a combination of convolutional layers, pooling, ReLU activation, and fully connected layers. The model receives input in the form of 256×256 pixel images, which are then processed through a series of convolutional layers. The process begins with the first convolutional layer of size 2×2 with an output dimension of 55×55×96, which aims to extract initial features from the input image. Next, the ReLU (Rectified Linear Unit) operation is applied to introduce non-linearity, followed by max pooling, which serves to reduce the spatial dimensions of the image while retaining important features. This stage helps to speed up computational processes while avoiding overfitting.

The next stage involves several additional convolutional layers of size 3×3, where the number of filters increases to 256 and 384. ReLU activation continues to be used after each convolutional layer to ensure efficient non-linear learning, while max pooling is applied again to reduce the data dimensions. After passing through the convolutional and pooling processes, the data is transformed into a vector form through a flattening process. Two dense layers, each with 4096 neurons, are then used to combine the features that have been previously extracted. ReLU activation is followed by a final dense layer with 1000 neurons, which is subsequently passed to the softmax layer to produce probability outputs corresponding to the classification categories of tea leaf infestation levels. The illustration of the AlexNet model architecture used is shown in Figure 12.

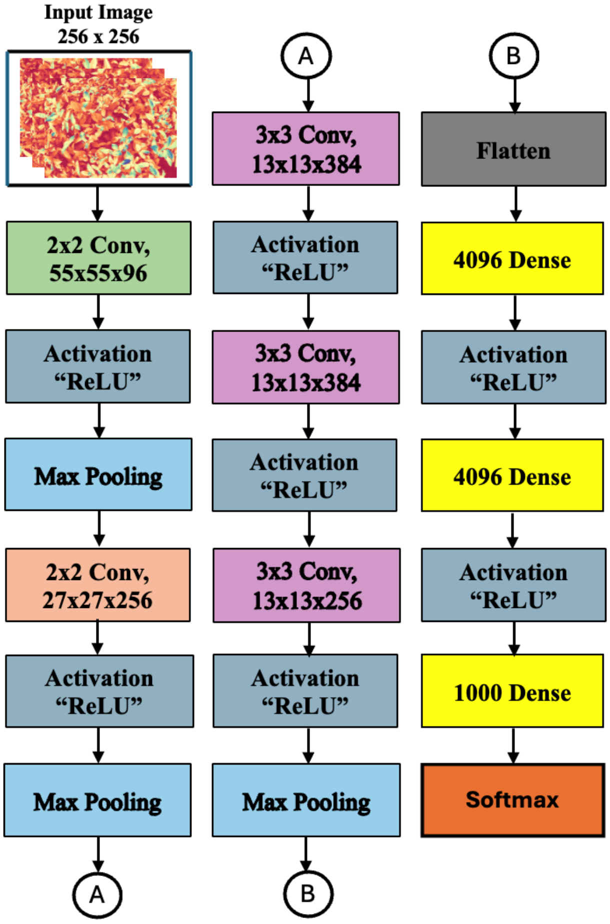

Figure 12. AlexNet Architecture

3-4- Performance Evaluation

After the training and testing stages, evaluation was conducted by analyzing the confusion matrix, which provides detailed information about the distribution of the model's predictions. The confusion matrix classifies prediction results into four main categories: true positive (TP), which indicates the number of correctly detected cases of infected leaves; true negative (TN), which refers to the number of non-infected leaves that were accurately classified. On the other hand, false positive (FP) represents cases where non-infected leaves were mistakenly classified as infected, while false negative (FN) refers to cases where infected leaves were not successfully detected by the model.

A systematic performance analysis of the ResNet-50 and AlexNet models in classifying tea leaves based on the level of *Empoasca* infestation was conducted to comprehensively evaluate the effectiveness of the classification models used in this study. Performance evaluation was carried out using key indicators, namely accuracy, precision, recall, F1-Score, and loss values, each of which provides an in-depth understanding of the model's ability to process and recognize patterns in the data. Accuracy measures the overall percentage of correct predictions, while precision indicates how well the model avoids errors by detecting only truly infected leaves. Recall reflects the model's ability to detect all cases of infected leaves, and F1-Score combines precision and recall to produce a balanced metric, especially important when the class distribution in the data is uneven. Additionally, analyzing the loss value is crucial to ensure that the model has been trained optimally without experiencing overfitting or underfitting.

3-4-1- Accuracy

Accuracy is an indicator used in testing to determine how accurately the ratio of correct predictions (negative and positive) can perform the classification process correctly. Equation 1 is used to calculate the accuracy value.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{1}$$

3-4-2- Precision

Precision is an indicator used in testing that describes how accurately a model predicts positive occurrences. Equation 2 is used to calculate the precision value by dividing the number of true positives by the total number of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{2}$$

3-4-3- Recall

Recall is a testing indicator that shows the success rate of correctly identifying the types of images after the identification process. Equation 3 is used to obtain the recall value:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

3-4-4- F1-Score

The harmonic mean of the precision and recall values calculated for the positive class. Equation 4 is used to determine the F1-Score.

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Recall}*\text{Precision}}{\text{Recall+Precision}} \tag{4}$$

3-4-5- Thresholding

Thresholding is an image segmentation technique used to separate objects from the background in an image. This method works by relying on differences in brightness or darkness in the image. By determining a certain threshold, thresholding can identify and separate parts of the image that have different levels of brightness or darkness, so that objects can be isolated from their backgrounds more clearly [40]. Parts of the image that have a higher level of darkness will be turned darker until they reach perfect black with an intensity value of 0. Meanwhile, parts of the image that tend to be light will be processed to become brighter, until they reach perfect white with an intensity value of 1. Thus, this process effectively emphasizes the contrast between dark and light areas in the image, making it easier to identify and analyze the objects contained therein [41].

As a result of the segmentation process using the thresholding method, the resulting image is a binary image with pixel intensity values consisting only of 0 or 1. After the object in the image is successfully separated from the background through segmentation, this binary image can be used as a mask for cropping so that the original image is obtained without a background or with a changeable background [42]. Furthermore, Equation 1 shows the process of grouping the pixel values into two categories, and Equation 2 shows the process of determining the threshold value T, as described by the following formula [43].

$$f(x, y) = \frac{G_{min}, f_0(x,y)<T}{G_{max}, f_1(x,y)\geq T} \tag{5}$$

where G_{min} is Intensity value 0, G_{max} is Intensity value 255, f_0 is The pixel intensity value is less than the threshold T, f_1 is The pixel intensity value is equal to or greater than the threshold T., and T = Threshold value.

$$T = \frac{G_{max} + G_{min}}{2} \tag{6}$$

In this process, pixels with intensity values less than are assigned a value of 0 (black), while those with values equal to or greater than are assigned 1 (white). This approach provides a simplified yet effective method for enhancing object contrast, making subsequent tasks such as feature extraction and classification more efficient.

4- Results and Discussion

This research focuses on classifying tea leaves affected by Empoasca pests using spectral data captured by a multispectral camera across four channels: Green, NIR (Near Infrared), REG (Red Edge), and RED. The primary objective was identifying the most effective channel for distinguishing healthy leaves from Empoasca-infested leaves. The process began with dataset collection, where images of healthy and infested leaves were captured, categorized, and labeled. This dataset, created under the guidance of field technicians, consisted of four categories: healthy leaves, leaves with mild Empoasca infestation, leaves with severe Empoasca infestation, and leaves affected by other pests. Each category included 300 images across the four spectral channels, resulting in a comprehensive dataset of 4,800 images. The dataset was divided into 80% for training and 20% for testing, allowing the models to learn patterns during the training phase and evaluate their performance on unseen data during testing.

Tests were conducted to achieve optimal classification accuracy using three deep learning models: Convolutional Neural Network (CNN), ResNet-50, and AlexNet. These tests involved single-channel, double-channel, three-channel, and four-channel combinations to classify tea leaves based on the severity of Empoasca infestation. The models were trained with 200 epochs, a batch size of 4, the Adam optimizer, a learning rate of 0.00001, and a decay rate of 0.000001. The performance of each model was evaluated by analyzing accuracy and loss across multiple channel configurations to ensure robust generalization to new data.

4-1-Single Channel Analysis

The first testing scenario was conducted to evaluate the performance of each model in classifying tea leaf canopies infested with *Empoasca* using a single-channel dataset, namely the Green, Red, NIR, and ReG channels. Performance evaluation was carried out using the Precision, Recall, *and* F1-Score metrics, which take into account the level of *Empoasca* infestation in the tea leaf canopies. The infestation levels were categorized into four classes: healthy leaves, leaves with mild *Empoasca* infestation, leaves with severe Empoasca infestation, and leaves affected by other pests.

Meanwhile, accuracy was evaluated without distinguishing the severity of the infestation in the tea leaf canopies to provide an overall picture of the model's performance. The accuracy evaluation was conducted to measure the proportion of correct predictions globally, reflecting the model's ability to recognize patterns from the given data. This testing is crucial to determine how well the models can adapt to various spectral channels without considering variations in infestation severity. The analysis also provides insights into the consistency of the model's performance when processing single-channel data from the Green, Red, NIR, and REG channels. The accuracy results for each model tested on the single-channel evaluation are presented in Figure 13 as part of the model performance analysis.

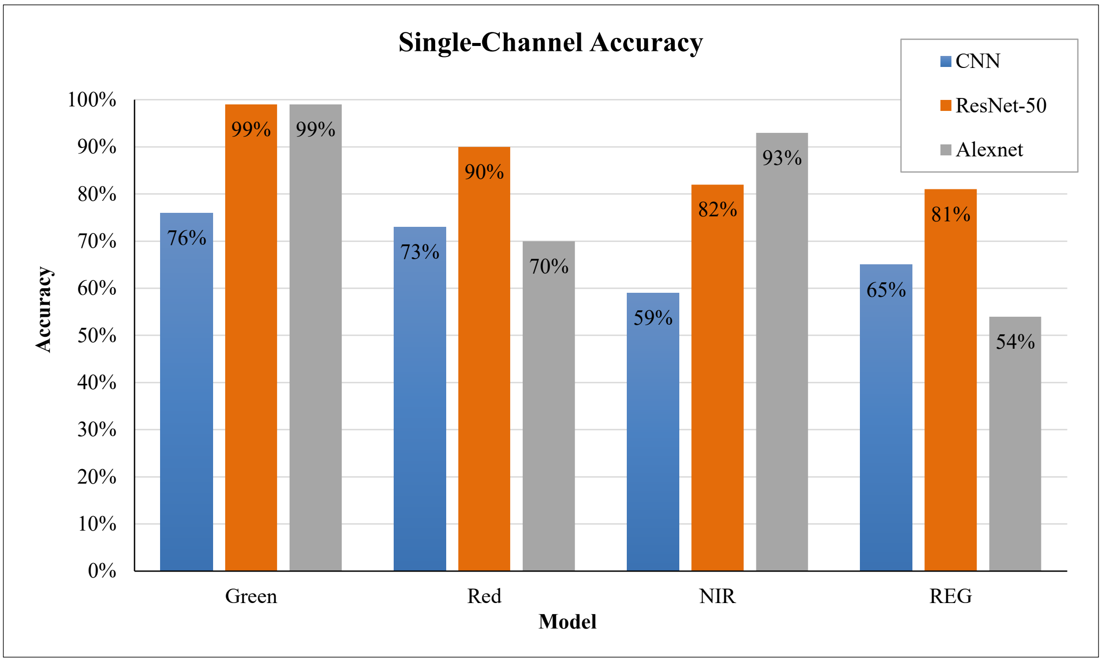

Figure 13. Comparison of Accuracy for Single-Channel

Based on the testing results shown in the two figures, the comparison of accuracy and loss for the three models—CNN, ResNet-50, and AlexNet—was analyzed across four single-channel datasets (Green, Red, NIR, and REG) to classify tea leaf canopies infested with Empoasca. The results in Figure 13 demonstrate that ResNet-50 and AlexNet consistently outperformed CNN in terms of accuracy. The Green channel showed the best performance for both ResNet-50 and AlexNet, with accuracy reaching 99%, while CNN achieved only 76%. This underscores the capability of deeper architectures, such as ResNet-50 and AlexNet, to extract more complex features from tea leaf images.

On the Red channel, ResNet-50 maintained strong performance with 90% accuracy, followed by CNN with 73% and AlexNet with 70%. Although ResNet-50 performed well on this channel, AlexNet's results were slightly lower than on the Green channel, suggesting that the Red channel may contain less informative features compared to Green. The NIR channel provided interesting results, with AlexNet achieving the highest accuracy of 93%, followed by ResNet-50 at 82%, and CNN at only 59%. This indicates that the NIR channel has significant potential for detecting patterns related to the level of tea leaf infestation, although CNN's performance on this channel was relatively weaker.

The REG channel produced more varied results. ResNet-50 achieved the highest accuracy of 81%, while CNN recorded 65%, and AlexNet only 54%. AlexNet's lower performance on the REG channel compared to the other channels suggests that this model is less effective at extracting features from the REG channel, while ResNet-50 demonstrated consistent stability.

Overall, these results show that ResNet-50 has the most consistent performance with high accuracy across all single-channel datasets. AlexNet excelled on certain channels, particularly NIR, but demonstrated limitations on the REG channel. Meanwhile, CNN exhibited lower performance compared to the other two models but remains relevant for scenarios requiring limited computational resources. The Green channel proved to be the most informative, followed by NIR and Red, while the REG channel displayed more varied performance across models.

Based on Table 2, the performance of CNN, ResNet-50, and AlexNet on single channels indicates that the Green channel consistently delivers the best results compared to other channels. On the Green channel, ResNet-50 recorded a very high F1-Score, reaching 99 percent for all classes except for the "Others" class, where the F1-Score was 100 percent. AlexNet also showed outstanding performance on this channel with a perfect F1-Score (100 percent) across almost all classes, demonstrating its ability to leverage the spectral features of the Green channel. CNN, while having lower performance compared to the other two models, remained stable with the highest F1-Score of 85 percent in the "Others" class. This shows that the Green channel provides highly rich and relevant spectral information for detecting levels of Empoasca infestation.

Table 2. Comparison of Precision, Recall, and F1-Score for CNN, ResNet-50, and AlexNet Across Single Channels (Green, Red, NIR, and REG) on Empoasca Infestation Classification

Channel	Class	CNN			ResNet-50			AlexNet		
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
GREEN	Healthy	83%	89%	86%	98%	**100%**	99%	**100%**	**100%**	**100%**
	Heavy	70%	55%	62%	98%	98%	88%	**100%**	**100%**	**100%**
	Light	74%	65%	69%	**100%**	97%	99%	98%	**100%**	99%
	Others	76%	**97%**	85%	**100%**	**100%**	**100%**	**100%**	98%	99%
RED	Healthy	**100%**	89%	**94%**	**100%**	**100%**	**100%**	80%	69%	74%
	Heavy	55%	85%	67%	98%	93%	96%	67%	70%	68%
	Light	73%	45%	56%	98%	98%	98%	65%	65%	65%
	Others	78%	74%	76%	93%	98%	96%	70%	76%	73%
NIR	Healthy	46%	53%	50%	82%	85%	83%	**100%**	94%	97%
	Heavy	70%	63%	67%	88%	82%	84%	83%	95%	88%
	Light	60%	60%	60%	81%	85%	83%	95%	87%	90%
	Others	61%	58%	60%	80%	78%	79%	95%	95%	95%
REG	Healthy	72%	71%	72%	83%	72%	77%	37%	95%	53%
	Heavy	79%	50%	61%	82%	83%	83%	81%	57%	67%
	Light	55%	75%	63%	82%	84%	83%	90%	66%	76%
	Others	64%	66%	65%	78%	84%	81%	**100%**	2%	3%

The Red channel, although yielding decent results, showed limitations compared to the Green channel. ResNet-50 remained the best-performing model on this channel with an average F1-Score of around 96 percent across all classes. However, AlexNet recorded significantly lower performance in the Light and Heavy classes, with F1-Score of 65 percent and 68 percent, respectively. CNN also showed lower results compared to the Green channel, particularly in the Light class, with an F1-Score of only 56 percent. This indicates that the Red channel is less informative in detecting complex infestation patterns compared to the Green channel.

The NIR channel produced mixed results, with ResNet-50 maintaining good F1-Score across all classes, averaging 83 percent. AlexNet performed well in the "Others" class with an F1-Score of 95 percent but experienced significant declines in the Light and Heavy classes, with F1-Scores of only 66 percent and 67 percent, respectively. CNN showed the lowest performance on this channel, with its highest F1-Score reaching only 67 percent in the Heavy class. This indicates that the NIR channel can provide additional useful information but is more effective when utilized by more complex models like ResNet-50.

The ReG channel showed the weakest results among all channels, with low F1-Score for CNN and AlexNet. ResNet-50 still demonstrated good performance with an F1-Score of around 83 percent in the Heavy class and 81 percent in the "Others" class. However, AlexNet recorded very poor results in the "Others" class, with an F1-Score of only 3 percent, indicating that the ReG channel has limitations in providing relevant information for classifying Empoasca infestation. Overall, the Green channel is the best single channel for detecting infestations, followed by NIR, while the Red and ReG channels provide more limited information. ResNet-50 exhibited the best performance across all channels, highlighting its capability to utilize spectral features from various sources effectively.

4-2- Double Channel Analysis

The second testing scenario was conducted to evaluate the performance of each model in classifying tea leaf canopies infested with Empoasca using a double-channel dataset, which consisted of combinations of Green and Red, Green and NIR, Green and ReG, Red and NIR, Red and ReG, as well as NIR and ReG channels. Performance evaluation was carried out using the Precision, Recall, and F1-Score metrics, which take into account the level of Empoasca infestation in the tea leaf canopies. The infestation levels were categorized into four classes: healthy leaves, leaves with mild Empoasca infestation, leaves with severe Empoasca infestation, and leaves affected by other pests. Additionally, accuracy evaluation was conducted to provide an overall picture of the model's ability to recognize patterns from the double-channel combinations. The accuracy results for each model tested in the double-channel scenario are presented in Figure 14.

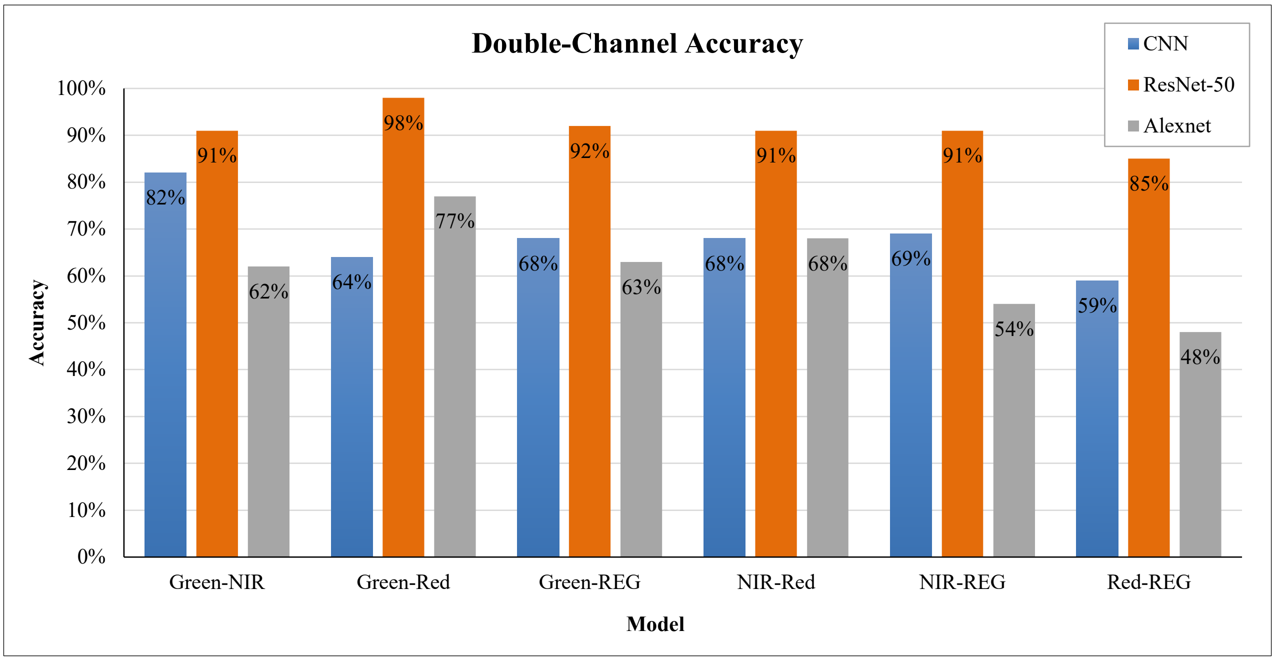

Figure 14. **Comparison of Accuracy for Double-Channel**

Based on Figure 14, the comparison of accuracy across six double-channel combinations reveals varying performance among the three classification models—CNN, ResNet-50, and AlexNet—in detecting Empoasca infestation in tea leaf canopies. The Green-Red channel combination achieved the highest accuracy, particularly for ResNet-50, which reached 98 percent, underscoring the model's superiority in utilizing complementary spectral information. CNN and AlexNet achieved accuracies of 64 percent and 77 percent, respectively, on this combination, showing that although AlexNet could leverage the information better than CNN, ResNet-50 significantly outperformed both.

In the Green-NIR combination, ResNet-50 again demonstrated excellent performance with 91 percent accuracy, followed by CNN with 82 percent and AlexNet with 62 percent. This combination appears to provide rich spectral information for detecting Empoasca infestation patterns, which are better utilized by deeper architectures like ResNet-50. AlexNet showed a considerable drop in accuracy on this combination, indicating its limitations in leveraging information from the NIR channel compared to Green.

The Green-ReG combination showed a similar trend, with ResNet-50 achieving 92 percent accuracy, while CNN and AlexNet recorded 68 percent and 63 percent, respectively. This indicates that the Green channel remains a significant contributor to classification, even though the ReG channel appears less informative for AlexNet and CNN. ResNet-50, with its ability to extract complex features, remained stable in utilizing information from this combination.

For the NIR-ReG combination, ResNet-50 achieved 91 percent accuracy, far surpassing CNN (69 percent) and AlexNet (54 percent). This combination indicates that the NIR channel is more dominant in providing crucial information compared to ReG. AlexNet exhibited a noticeable weakness in this combination, indicating that this model is less effective in utilizing spectral features from the ReG channel. In contrast, ResNet-50 maintained its high performance.

In the Red-ReG combination, ResNet-50 achieved 85 percent accuracy, while CNN and AlexNet recorded 59 percent and 48 percent, respectively. This combination shows that the ReG channel has significant limitations in providing informative features, especially for AlexNet. Although the Red channel is somewhat informative, the superior performance of ResNet-50 underscores its ability to utilize additional information from the channel combination.

Overall, ResNet-50 demonstrated the best performance across all double-channel combinations, with consistently high accuracy across all pairs. CNN showed relatively competitive performance in certain combinations, such as Green-NIR, but lagged significantly in others. AlexNet, while showing strength in certain combinations like Green-Red, displayed weaknesses in combinations involving ReG. These results highlight the importance of selecting the optimal channel combination and the appropriate classification model to effectively detect Empoasca infestations.

Based on the analysis in Table 3, the Green-Red double-channel combination demonstrated the best performance among all tested channel combinations. ResNet-50 achieved the highest F1-Score of 99 percent across nearly all classes, including healthy leaves, mild infestations, and severe infestations. This combination integrates spectral information from the Green channel, which was proven to be the best in the single-channel tests, with additional features from the Red channel, improving classification accuracy. CNN and AlexNet also showed improved performance with this combination compared to others, although they still lagged behind ResNet-50. These results indicate that the Green-Red channel combination provides optimal synergy for detecting Empoasca infestation patterns.

Table 3. Comparison of Precision, Recall, and F1-Score for CNN, ResNet-50, and AlexNet Across Double Channels on Empoasca Infestation Classification

Channel	Class	CNN			ResNet-50			AlexNet		
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
GREEN-NIR	Healthy	**96%**	**94%**	**95%**	91%	94%	92%	97%	**98%**	**98%**
	Heavy	76%	78%	77%	92%	90%	91%	25%	1%	2%
	Light	75%	76%	75%	88%	93%	91%	42%	90%	58%
	Others	82%	81%	82%	93%	87%	90%	68%	59%	63%
GREEN-RED	Healthy	68%	83%	75%	99%	**99%**	**99%**	91%	87%	89%
	Heavy	62%	49%	55%	**100%**	98%	**99%**	79%	70%	74%
	Light	61%	71%	65%	94%	**99%**	97%	70%	77%	73%
	Others	63%	51%	57%	97%	94%	96%	70%	74%	72%
GREEN-REG	Healthy	71%	67%	69%	91%	93%	92%	96%	52%	68%
	Heavy	56%	62%	59%	93%	93%	93%	**98%**	50%	66%
	Light	70%	66%	68%	91%	87%	89%	94%	52%	67%
	Others	74%	74%	74%	92%	87%	93%	40%	**98%**	57%
NIR-RED	Healthy	85%	68%	76%	93%	90%	92%	98%	71%	82%
	Heavy	58%	77%	66%	90%	95%	92%	62%	90%	74%
	Light	61%	61%	61%	94%	90%	92%	75%	10%	18%
	Others	76%	67%	71%	88%	90%	89%	59%	98%	73%
NIR-REG	Healthy	57%	55%	56%	91%	91%	91%	34%	75%	47%
	Heavy	68%	73%	71%	91%	90%	90%	73%	65%	69%
	Light	69%	73%	71%	94%	90%	92%	87%	79%	83%
	Others	79%	74%	77%	88%	93%	90%	0%	0%	0%
RED-REG	Healthy	47%	54%	50%	60%	87%	87%	61%	11%	19%
	Heavy	63%	64%	64%	87%	85%	84%	35%	74%	48%
	Light	71%	47%	57%	83%	81%	83%	65%	51%	55%
	Others	60%	69%	64%	86%	85%	84%	61%	11%	61%

The Green-NIR combination emerged as the second-best option, especially for ResNet-50, with an average F1-Score of 99 percent for healthy and mildly infested leaves. The NIR channel complements the strength of the Green channel by providing additional spectral information, particularly useful for detecting more severe infestations. CNN performed fairly well with this combination, achieving an average F1-Score of approximately 75 percent, while AlexNet recorded lower results, particularly for severe infestations, with an F1-Score of only 2 percent. This combination suggests that Green-NIR can deliver competitive results but is more effective when used with deep architectures like ResNet-50.

The NIR-Red combination also showed promising results for ResNet-50, with an average F1-Score of 92 percent across most classes. This combination provides additional information from the NIR channel, which is valuable for detecting complex infestation patterns. However, AlexNet again exhibited inconsistent performance, especially for mild and severe infestations, with relatively low F1-Score. CNN showed stable performance with an average F1-Score of approximately 64 percent, though it still lagged behind ResNet-50. This combination indicates that NIR-Red provides sufficient spectral information but is more effective when utilized by models with strong feature extraction capabilities.

Overall, the Green-Red channel combination proved to be the best for detecting Empoasca infestations. This combination delivered the highest F1-Score across nearly all classes for ResNet-50 and performed relatively well for

CNN and AlexNet compared to other combinations. The superiority of Green-Red is supported by the rich spectral information provided by the Green channel, enhanced by the contributions of the Red channel. The Green-NIR combination emerged as the second-best option, while combinations involving the ReG channel tended to show lower performance, especially for models with simpler architectures like AlexNet.

Based on the results in Table 3 and Figure 14, ResNet-50 consistently demonstrated the best performance among the three models tested, achieving higher accuracy, precision, recall, and F1-Score across nearly all channel combinations. The Green-Red channel combination proved to be the most superior, with an accuracy of 98 percent for ResNet-50 and the highest F1-Score across all classes in the table. This channel combination integrates the spectral strengths of Green, which has been shown to excel in single-channel testing, with the contributions of Red to enrich feature information.

4-3- Triple Channel Analysis

The third testing scenario was conducted to evaluate the performance of each model in classifying tea leaf canopies infested with Empoasca using a triple-channel dataset. This dataset consisted of combinations of Green-NIR-Red, Green-NIR-REG, NIR-Red-REG, and Red-REG-Green channels. Performance evaluation was carried out using the Precision, Recall, and F1-Score metrics, which take into account the level of Empoasca infestation in the tea leaf canopies. The infestation levels were categorized into four classes: healthy leaves, leaves with mild Empoasca infestation, leaves with severe Empoasca infestation, and leaves affected by other pests. This evaluation aimed to assess the extent to which the models could recognize more complex patterns by leveraging the combination of three spectral channels.

In addition, an accuracy evaluation was conducted to provide a general overview of the model's ability to recognize patterns from the triple-channel combinations. Unlike the previous testing scenarios, this evaluation was designed to measure how much the model's accuracy improves with the additional spectral information from the third channel. This analysis also provides insights into the consistency of the model's performance when processing more complex data combinations. The accuracy results for each model tested in the triple-channel scenario are presented in Figure 15 as part of the model performance analysis

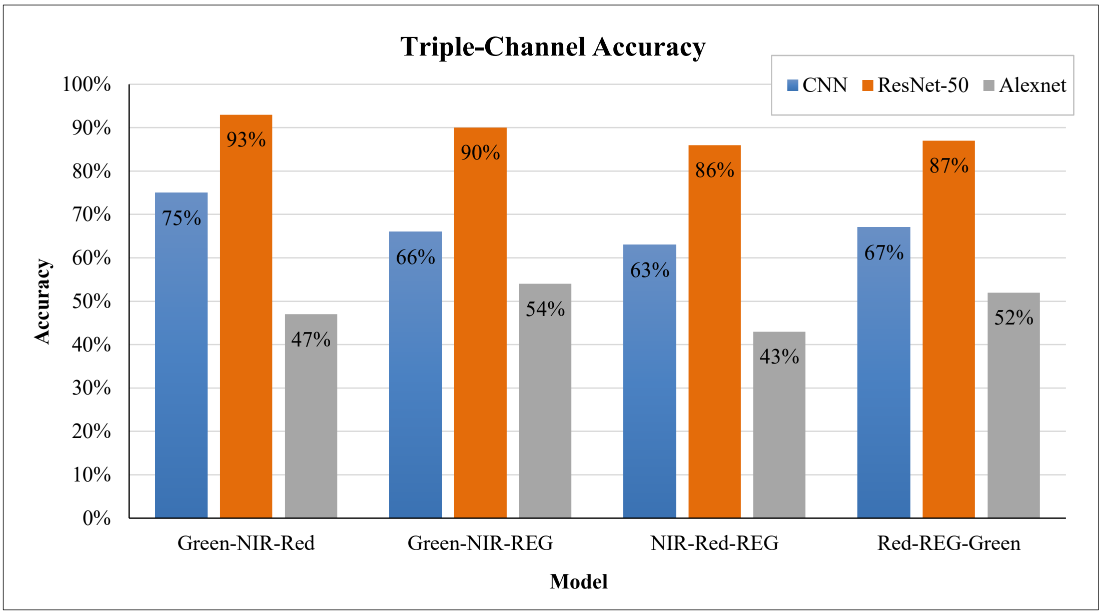

Figure 15. Comparison of Accuracy for Triple-Channel

Based on Figure 15, the accuracy of the three classification models—CNN, ResNet-50, and AlexNet—on four triple-channel combinations shows that ResNet-50 consistently outperforms the other models. Among all combinations, the Green-NIR-Red channel achieved the highest accuracy for ResNet-50, indicating its strong ability to leverage the combined spectral information from these three channels. CNN also performed relatively well on this combination, while AlexNet exhibited significantly lower accuracy, suggesting that it struggles to process the added complexity of triple-channel data effectively.

The second-best combination for ResNet-50 is Green-NIR-REG, which shows slightly lower accuracy compared to Green-NIR-Red. This indicates that while the addition of the REG channel contributes valuable information, it is not as effective as the contribution from the Red channel. For CNN, accuracy on Green-NIR-REG is notably lower compared to its performance on Green-NIR-Red, reflecting its limitations in utilizing information from the REG channel. AlexNet, on the other hand, showed a consistent pattern of lower accuracy across all combinations, indicating its limited ability to generalize in this scenario.

The NIR-Red-REG combination produced the third-best accuracy for ResNet-50, demonstrating its ability to adapt to various spectral combinations. However, for CNN, the performance dropped significantly compared to the previous combinations, suggesting that this combination may be less informative for simpler models. AlexNet again recorded the lowest accuracy for this combination, emphasizing its inability to fully utilize the spectral features of this channel combination.

The Red-REG-Green combination showed comparable performance to the NIR-Red-REG combination for ResNet-50, maintaining its position as the best-performing model overall. CNN achieved slightly better accuracy on this combination compared to NIR-Red-REG, while AlexNet's performance remained consistently low. This pattern highlights the robustness of ResNet-50 in adapting to various combinations and its ability to extract relevant features effectively.

In conclusion, the Green-NIR-Red combination emerged as the best triple-channel configuration, particularly for ResNet-50, due to its ability to capture complementary spectral features from these channels. ResNet-50 is consistently the best-performing model across all combinations, while CNN offers moderate performance, and AlexNet struggles to maintain accuracy in this more complex scenario. The results underscore the importance of selecting optimal channel combinations and robust models to achieve accurate classification in tea leaf canopy analysis.

Based on the results presented in Table 4, the combination of triple channels shows varying performances across different models and classes. ResNet-50 consistently demonstrates superior performance, achieving the highest precision, recall, and F1-Score across almost all channel combinations and classes. Among the four channel combinations, Green-NIR-Red stands out as the best-performing combination for ResNet-50, with exceptionally high metrics across all classes, particularly in the Healthy and Others categories. This performance highlights the strength of combining Green and Red channels, previously identified as the best single and double channels, respectively, with the additional spectral information provided by the NIR channel.

Table 4. Comparison of Precision, Recall, and F1-Score for CNN, ResNet-50, and AlexNet Across Triple Channels on Empoasca Infestation Classification

Channel	Class	CNN			ResNet-50			AlexNet		
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
GREEN-NIR-RED	Healthy	82%	79%	80%	93%	**95%**	94%	89%	63%	74%
	Heavy	67%	74%	70%	**95%**	93%	94%	50%	4%	8%
	Light	68%	68%	68%	91%	**93%**	92%	32%	92%	48%
	Others	84%	80%	82%	**91%**	89%	90%	53%	28%	36%
GREEN-NIR-REG	Healthy	82%	73%	77%	90%	88%	89%	79%	70%	74%
	Heavy	74%	39%	51%	88%	**91%**	89%	69%	35%	46%
	Light	68%	64%	66%	**92%**	91%	91%	34%	55%	42%
	Others	52%	88%	66%	**92%**	**92%**	**92%**	49%	53%	51%
NIR-RED-REG	Healthy	83%	56%	67%	**87%**	**87%**	**87%**	87%	22%	35%
	Heavy	55%	58%	57%	**89%**	85%	87%	0%	0%	0%
	Light	54%	72%	62%	84%	**88%**	86%	30%	98%	46%
	Others	66%	67%	66%	**85%**	**85%**	**85%**	79%	59%	68%
RED-REG-GREEN	Healthy	72%	75%	73%	84%	**90%**	87%	74%	68%	71%
	Heavy	61%	62%	62%	**87%**	**87%**	**87%**	0%	0%	0%
	Light	61%	67%	64%	**86%**	**86%**	**86%**	53%	65%	59%
	Others	72%	61%	66%	**93%**	86%	90%	38%	73%	50%

The second-best combination is Green-NIR-REG, which also yields strong metrics for ResNet-50, though slightly lower than Green-NIR-Red. For CNN, the performance on Green-NIR-REG is moderate, with lower metrics compared to Green-NIR-Red. AlexNet, however, struggles to utilize the combined spectral information from these three channels effectively, resulting in significantly lower precision, recall, and F1-Score, especially for classes such as Heavy and Light infestation. These results suggest that AlexNet lacks the capacity to generalize effectively with more complex combinations, limiting its utility in this scenario.

The NIR-Red-REG and Red-REG-Green combinations rank lower in terms of overall performance. ResNet-50 remains consistent, though its performance for these combinations is notably lower compared to Green-NIR-Red. CNN shows a significant drop in metrics, particularly for the Heavy and Light classes, across both channel combinations.

AlexNet again records poor results, emphasizing its limitations in handling multi-channel data effectively. The results indicate that these combinations lack the complementary spectral information seen in Green-NIR-Red and Green-NIR-REG, which are essential for accurate classification.

In conclusion, the Green-NIR-Red channel combination emerges as the best configuration for detecting Empoasca infestation, particularly when using ResNet-50. This combination leverages the strengths of the Green channel, the best single-channel performer, and the Green-Red combination, the best double-channel configuration, while enhancing it further with the NIR channel. The results reaffirm the importance of selecting optimal channel combinations to maximize classification accuracy, with ResNet-50 proving to be the most robust and consistent model across all combinations.

Based on Table 4 and Figure 15, ResNet-50 consistently emerges as the best-performing model across all triple-channel combinations, with the highest accuracy, precision, recall, and F1-Score values. Among the combinations, Green-NIR-Red proves to be the most effective, leveraging the spectral richness of the Green channel, identified as the best single-channel performer, combined with the complementary features of the NIR and Red channels. This combination enables ResNet-50 to achieve optimal classification results, particularly for detecting healthy leaves and various levels of Empoasca infestation. The second-best combination is Green-NIR-REG, which still delivers strong performance for ResNet-50 but slightly lags behind Green-NIR-Red. CNN demonstrates moderate accuracy for Green-NIR-Red and Green-NIR-REG, although its performance is less consistent for other combinations. AlexNet, on the other hand, struggles to generalize across all triple-channel combinations, indicating its limitations in processing complex spectral data. Overall, the Green-NIR-Red combination is the most effective for classification, and ResNet-50 is the most robust model for accurately identifying infestation levels in tea leaf canopies.

4-4- Four Channel Analysis

The fourth testing scenario was conducted to evaluate the performance of the model in classifying tea leaf canopies infested with Empoasca using a four-channel dataset, namely the combination of Green, NIR, Red, and ReG channels. Performance evaluation was carried out using the Precision, Recall, and F1-Score metrics, taking into account four infestation categories: healthy leaves, leaves with mild Empoasca infestation, leaves with severe Empoasca infestation, and leaves affected by other pests. Model accuracy was also analyzed to provide an overall picture of the model's ability to recognize patterns from this spectral combination without distinguishing the severity of infestation. This testing aims to determine the extent to which the model can optimally and consistently utilize all four channels in the classification task. The accuracy results for the model in the four-channel testing are presented in Figure 16 as part of the model performance analysis

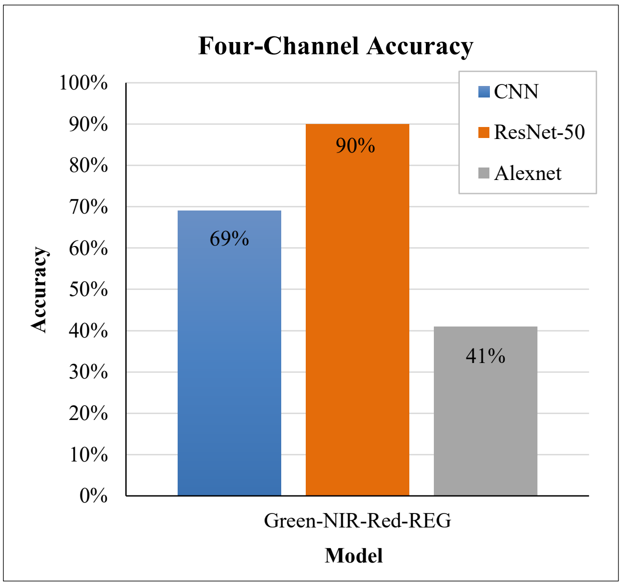

Figure 16. **Accuracy for Four-Channel**

The results displayed in Figure 16 highlight the accuracy of three classification models—CNN, ResNet-50, and AlexNet—on the four-channel dataset, which combines Green, NIR, Red, and ReG channels. ResNet-50 demonstrates the highest accuracy, showcasing its superior ability to leverage the complementary spectral information provided by all four channels. This outcome further reinforces ResNet-50's capacity to generalize effectively across complex datasets, making it an ideal choice for tasks that involve diverse spectral inputs. CNN follows with moderate accuracy, which, although lower than ResNet-50, indicates its ability to extract meaningful patterns even with simpler architectural depth.

AlexNet, on the other hand, shows the lowest accuracy among the three models, further affirming its limitations when processing datasets with increased spectral complexity. The drop in performance may result from its simpler architecture, which struggles to fully utilize the added features provided by the four-channel combination. Despite this, the inclusion of the Green channel, which consistently performed well in previous single-channel and double-channel evaluations, continues to contribute significantly to the overall performance, especially for ResNet-50. These findings emphasize the importance of pairing advanced models like ResNet-50 with rich spectral data combinations for optimal classification outcomes.

The analysis of Table 5 highlights the performance of the three models—CNN, ResNet-50, and AlexNet—when applied to the four-channel combination (Green, NIR, Red, and REG). ResNet-50 consistently demonstrates superior results across all metrics (Precision, Recall, and F1-Score) for every class. This model achieves high recall and F1-Score values, particularly for the healthy and heavy infestation classes, reflecting its robustness in capturing features across complex datasets. CNN performs moderately well, with relatively balanced scores, particularly in the "others" and "healthy" classes, but it lags behind ResNet-50 in effectively utilizing the combined spectral information. AlexNet, however, struggles significantly with this combination, particularly in the "light" and "others" classes, where the F1-Score values are remarkably low, indicating an inability to adapt to the increased complexity of four-channel data. This pattern underscores the importance of deeper architectures, like ResNet-50, in leveraging complementary features from multiple channels.

Table 5. Comparison of Precision, Recall, and F1-Score for CNN, ResNet-50, and AlexNet Using Four Channels (Green, NIR, Red, and REG) on Empoasca Infestation Classification

Channel	Class	CNN			ResNet-50			AlexNet		
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
GREEN-NIR-RED-REG	Healthy	79%	72%	75%	**91%**	**91%**	**91%**	85%	67%	75%
	Heavy	66%	51%	57%	89%	**93%**	91%	27%	87%	42%
	Light	70%	70%	70%	88%	**89%**	88%	50%	81%	2%
	Others	61%	80%	69%	**91%**	86%	89%	42%	10%	16%

Integrating the results from Table 5 and Figure 16, it is evident that ResNet-50 is the best-performing model for the four-channel combination (Green-NIR-Red-REG), achieving the highest accuracy and balanced scores across all classes. This aligns with previous findings that highlighted the strength of the Green channel in single-channel experiments and the Green-Red combination in double-channel tests. ResNet-50's deep architecture enables it to effectively process and extract meaningful features from the combined spectral information provided by the four channels. In contrast, CNN shows moderate performance, reflecting its limitations in handling the complexity of multi-channel data, while AlexNet exhibits substantial challenges, with its accuracy and F1-Score significantly lower for several classes. These findings confirm that ResNet-50, combined with the Green-NIR-Red-REG channel configuration, provides the most reliable and accurate model for classifying Empoasca infestations in tea leaf canopies, leveraging the complementary features of all four channels effectively.

The comprehensive analysis of single-channel, double-channel, triple-channel, and four-channel combinations highlights the critical role of spectral data in enhancing the accuracy of Empoasca pest detection in tea plantations. Among all the tested configurations, the Green channel consistently emerged as the most informative single-channel, providing rich spectral features that significantly contributed to classification performance. This was further validated in combined-channel analyses, where the integration of the Green channel with other spectral bands, particularly Red and NIR, demonstrated superior results. For instance, the Green-Red combination achieved 98% accuracy with ResNet-50, while the Green-NIR-Red triple-channel combination delivered the highest overall accuracy of 94%. These findings underscore the importance of leveraging complementary spectral information to maximize classification outcomes. Notably, the stepwise addition of channels—such as progressing from single to double and triple combinations—resulted in measurable improvements in accuracy, with the Green-NIR-Red combination outperforming simpler configurations like Green-Red or Green-NIR. This progressive enhancement demonstrates the value of combining multiple spectral bands to capture nuanced features for improved pest detection.

ResNet-50 consistently outperformed CNN and AlexNet across all channel configurations, reaffirming its suitability for agricultural applications involving multispectral imaging. Its ability to effectively process and extract meaningful features from combined spectral data, even in the most complex four-channel configuration (Green-NIR-Red-REG), highlights the strength of its deep architecture and residual learning mechanism. For example, ResNet-50 achieved a peak accuracy of 99% on the Green channel, compared to CNN's 76% and AlexNet's slightly lower performance in multi-channel scenarios. The improvement in accuracy when transitioning from simpler models like AlexNet to ResNet-50 underscores the importance of advanced architectures in handling high-dimensional and complex datasets. While CNN showed moderate performance, it struggled to match ResNet-50's accuracy, particularly in multi-channel scenarios, emphasizing the need for robust models to achieve reliable pest classification. These results collectively suggest that ResNet-50, paired with optimized spectral channel combinations, offers a scalable and accurate solution for modern tea plantation management.

The implications of these findings extend beyond Empoasca pest detection, offering a framework for broader agricultural applications. By integrating UAV-mounted multispectral cameras with deep learning models like ResNet-50, this approach has the potential to revolutionize pest monitoring and management practices on a large scale. However, challenges such as computational requirements, battery life, and data transmission must be addressed to ensure efficient deployment in real-world scenarios. For instance, optimizing energy consumption through low-power sensors and strategic flight planning can help extend UAV operational time. Additionally, enhancing the interpretability of model outputs through explainable AI techniques could empower farmers and agricultural experts to make informed decisions, ultimately contributing to sustainable and efficient pest management strategies. These advancements highlight the importance of leveraging advanced technologies to address critical challenges in modern agriculture.

4-5- Discussion

The study highlights the effectiveness of the green channel in detecting Empoasca pest infestations on tea plantations, leveraging its unique sensitivity to chlorophyll degradation and surface pigmentation changes caused by hama. The green spectrum is particularly well-suited for this task due to several key factors. First, it captures significant variations in reflectance that occur when chlorophyll content decreases, making it highly effective for identifying early signs of stress caused by Empoasca feeding. Second, the green channel provides clear visual contrast between healthy and infested leaves, which is crucial for distinguishing subtle differences in leaf texture and color. Additionally, as part of the visible light spectrum, the green channel enables early detection of infestations, even when changes may not yet be apparent in other spectral bands like NIR or red.

However, relying solely on the green channel has limitations. It lacks structural information provided by other bands such as NIR, which can be critical for differentiating damage caused by Empoasca from other environmental factors like drought or diseases. In cases of severe infestation, where spectral changes involve multiple bands, focusing only on the green channel may overlook complex patterns present in other channels, potentially reducing detection accuracy. Furthermore, environmental conditions such as lighting variability and weather effects can influence green channel reflectance, introducing noise or bias if not complemented by data from additional channels.

ResNet-50 demonstrates superior performance compared to AlexNet, particularly in handling complex multispectral data. This superiority stems from fundamental architectural differences. AlexNet, with its relatively shallow structure of 8 layers, struggles to capture intricate patterns in high-dimensional datasets like multispectral images. In contrast, ResNet-50's deeper architecture (50 layers) incorporates residual connections (skip connections), which mitigate vanishing gradient issues during backpropagation. These skip connections enable ResNet-50 to effectively process hierarchical and abstract features in multispectral data, making it better suited for fine-grained analysis tasks such as detecting subtle spectral variations caused by Empoasca infestations.

ResNet-50's ability to handle complex channel combinations and extract discriminative features is a key advantage. Its deep architecture allows progressive learning of both low-level features (e.g., edges and textures) and high-level features (e.g., specific spectral signatures of infested leaves). Bottleneck layers and global average pooling further enhance its robustness by reducing overfitting and improving generalization, especially in variable agricultural settings. Pre-trained weights on large datasets also facilitate transfer learning, enabling effective performance even with limited labeled data. These architectural features collectively contribute to ResNet-50's higher accuracy and reliability compared to AlexNet.

5- Conclusion

This study demonstrates the efficacy of multispectral imaging combined with deep learning models for detecting Empoasca pest infestations in tea plantations. Among the models evaluated, ResNet-50 consistently emerged as the most effective, achieving superior accuracy across all channel combinations, with a peak accuracy of 99% when utilizing the Green channel. The Green channel proved to be the most informative single channel, providing significant spectral features for identifying pest infestations. Its importance was further highlighted in combined-channel analyses, where it consistently enhanced model performance, particularly in the Green-Red and Green-NIR combinations. These combinations leveraged the complementary strengths of the Green channel and other channels, further improving detection accuracy and demonstrating the robustness of ResNet-50 in handling complex multispectral data. AlexNet, while performing well on the Green channel, exhibited significant limitations in processing more complex channel combinations, further underscoring the strength of ResNet-50.

The findings underline the potential of integrating advanced convolutional neural network architectures, particularly ResNet-50, with optimized multispectral imaging for scalable, efficient, and accurate pest management in tea plantations. ResNet-50's ability to effectively process complex datasets through its residual learning mechanism makes it a reliable solution for agricultural applications. The Green-NIR-Red-REG channel combination further demonstrated the system's potential to maximize pest detection accuracy by incorporating complementary spectral features. This integration offers a practical tool for improving pest monitoring and detection in real-world agricultural settings.

Future research could focus on real-time deployment of this system using UAV-mounted multispectral cameras to facilitate large-scale monitoring and detection of pest infestations. Additionally, expanding the applicability of this approach to other crops and pest species could validate its utility and adaptability across diverse agricultural contexts. Exploring methods to reduce computational requirements while maintaining accuracy is critical to ensuring widespread adoption, especially in resource-limited environments. Finally, enhancing the interpretability of the model's outputs through explainable AI techniques could help farmers and agricultural experts make informed decisions, ensuring this technology effectively addresses the challenges of sustainable pest management

6- Declarations

6-1- Author Contributions

Conceptualization, B.K.K., F.F., and A.A.P.; methodology, B.K.K., F.Y.S., and A.A.P.; software, B.K.K., H.V., and D.M.S.; validation, F.F., F.Y.S., and A.A.P.; formal analysis, F.F. and A.A.P.; investigation, F.Y.S.; resources, F.F.; data curation, A.A.P.; writing—original draft preparation, B.K.K., H.V., and D.M.S.; writing—review and editing, F.Y.S. and A.A.P.; visualization, H.V. and D.M.S.; supervision, F.F., F.Y.S., and A.A.P.; project administration, H.V. All authors have read and agreed to the published version of the manuscript.

6-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6-3- Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6-4- Institutional Review Board Statement

Not applicable.

6-5- Informed Consent Statement

Not applicable.

6-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

[1] Hajiboland, R. (2017). Environmental and nutritional requirements for tea cultivation. Folia Horticulturae, 29(2), 199–220. doi:10.1515/fhort-2017-0019.

[2] Adelianingsih, D., Hidayati, R., & Sugiarto, Y. (2019). Potential of Green Leafhopper Attack (Empoasca sp.) in Tea Plantation Based on Climate Change Scenarios. Agromet, 33(2), 84–95. doi:10.29244/j.agromet.33.2.84-95.

[3] Pramudita, A. A., Wahyu, Y., Rizal, S., Prasetio, M. D., Jati, A. N., Wulansari, R., & Ryanu, H. H. (2022). Soil Water Content Estimation with the Presence of Vegetation Using Ultra Wideband Radar-Drone. IEEE Access, 10, 85213–85227. doi:10.1109/ACCESS.2022.3197636.

[4] Narmilan, A., Gonzalez, F., Salgadoe, A. S. A., & Powell, K. (2022). Detection of White Leaf Disease in Sugarcane Using Machine Learning Techniques over UAV Multispectral Images. Drones, 6(9), 230. doi:10.3390/drones6090230.

[5] Hooshyar, M., Li, Y. S., Chun Tang, W., Chen, L. W., & Huang, Y. M. (2024). Economic Fruit Trees Recognition in Hillsides: A CNN-Based Approach Using Enhanced UAV Imagery. IEEE Access, 12, 61991–62005. doi:10.1109/ACCESS.2024.3391371.

[6] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94.

[7] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding, 110(3), 346–359. doi:10.1016/j.cviu.2007.09.014.

[8] Tsolakidis, D.G., Kosmopoulos, D.I., Papadourakis, G. (2014). Plant Leaf Recognition Using Zernike Moments and Histogram of Oriented Gradients. Artificial Intelligence: Methods and Applications. SETN 2014, Lecture Notes in Computer Science, 8445, Springer, Cham, Switzerland. doi:10.1007/978-3-319-07064-3_33.

[9] Hossain, S., Mou, R. M., Hasan, M. M., Chakraborty, S., & Razzak, M. A. (2018). Recognition and detection of tea leaf's diseases using support vector machine. 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 150–154. doi:10.1109/cspa.2018.8368703.

[10] Muralidharan, R., & Chandrasekar, C. (2011). Object recognition using SVM-KNN based on geometric moment invariant. International Journal of Computer Trends and Technology, 1(1), 215-220.

[11] Qadri, S. A. A., Huang, N. F., Wani, T. M., & Bhat, S. A. (2025). Advances and Challenges in Computer Vision for Image-Based Plant Disease Detection: A Comprehensive Survey of Machine and Deep Learning Approaches. IEEE Transactions on Automation Science and Engineering, 22, 2639–2670. doi:10.1109/TASE.2024.3382731.

[12] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(1), 53. doi:10.1186/s40537-021-00444-8.

[13] Hari, S. S., Sivakumar, M., Renuga, P., karthikeyan, S., & Suriya, S. (2019). Detection of Plant Disease by Leaf Image Using Convolutional Neural Network. 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 1–5. doi:10.1109/vitecon.2019.8899748.

[14] Zuhdi, F., Rambe, K. R., & Rahmadona, L. (2022). Analysis of Competitiveness and Forecasting of Indonesian Tea Exports to Main Destination Countries. Media Ekonomi Dan Manajemen, 37(2), 240. doi:10.24856/mem.v37i2.2888.

[15] Wang, F., Zhang, B., Wen, D., Liu, R., Yao, X., Chen, Z., Mu, R., Pei, H., Liu, M., Song, B., & Lu, L. (2022). Chromosome-scale genome assembly of Camellia sinensis combined with multi-omics provides insights into its responses to infestation with green leafhoppers. Frontiers in Plant Science, 13, 1004387. doi:10.3389/fpls.2022.1004387.

[16] Fornasiero, D., Pavan, F., Pozzebon, A., Picotti, P., & Duso, C. (2016). Relative infestation level and sensitivity of grapevine cultivars to the leafhopper empoasca vitis (Hemiptera: Cicadellidae). Journal of Economic Entomology, 109(1), 416–425. doi:10.1093/jee/tov313.

[17] Krichen, M. (2023). Convolutional Neural Networks: A Survey. Computers, 12(8), 151. doi:10.3390/computers12080151.

[18] Tran, D. T., Iosifidis, A., & Gabbouj, M. (2018). Improving efficiency in convolutional neural networks with multilinear filters. Neural Networks, 105, 328–339. doi:10.1016/j.neunet.2018.05.017.

[19] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017, 1–6. doi:10.1109/ICEngTechnol.2017.8308186.

[20] Ghiasi-Shirazi, K. (2019). Generalizing the Convolution Operator in Convolutional Neural Networks. Neural Processing Letters, 50(3), 2627–2646. doi:10.1007/s11063-019-10043-7.

[21] Xu, W., He, J., Shu, Y., & Zheng, H. (2020). Advances in Convolutional Neural Networks. Advances and Applications in Deep Learning, IntechOpen, London, United Kingdom. doi:10.5772/intechopen.93512.

[22] Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. Artificial Intelligence Review, 57(4), 99. doi:10.1007/s10462-024-10721-6.

[23] Yoo, J., & Kang, S. (2023). Class-Adaptive Data Augmentation for Image Classification. IEEE Access, 11, 26393–26402. doi:10.1109/ACCESS.2023.3258179.

[24] Haryono, K. B., Purnomo, H. C., Ferdinand, R., Lucky, H., & Suhartono, D. (2022). Investigating the Influence of Layers Towards Speed and Accuracy of Neural Networks. 2022 International Conference on Informatics Electrical and Electronics (ICIEE), 1–6. doi:10.1109/iciee55596.2022.10010322.

[25] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. doi:10.1145/3065386.

[26] Ramdan, A., Suryawati, E., Kusumo, R. B. S., Pardede, H. F., Mahendra, O., Dahlan, R., Fauziah, F., & Syahrian, H. (2019). Deep CNNBased Detection for Tea Clone Identification. Jurnal Elektronika Dan Telekomunikasi, 19(2), 45. doi:10.14203/jet.v19.45-50.

[27] Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning, 6-11 July, 2015, Lille, France.

[28] Mikolajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW), 117–122. doi:10.1109/iiphdw.2018.8388338.

[29] Reghunath, A., Nair, S. V., & Shah, J. (2019). Deep learning based Customized Model for Features Extraction. 2019 International Conference on Communication and Electronics Systems (ICCES), 1406–1411. doi:10.1109/icces45898.2019.9002299.

[30] Basha, S. H. S., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing, 378, 112–119. doi:10.1016/j.neucom.2019.10.008.

[31] Chen, H. (2023). Studies Advanced in Image Classification based on Deep Learning. Applied and Computational Engineering, 8(1), 623–628. doi:10.54254/2755-2721/8/20230287.

[32] Mukti, I. Z., & Biswas, D. (2019). Transfer Learning Based Plant Diseases Detection Using ResNet50. 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 1–6. doi:10.1109/eict48899.2019.9068805.

[33] Hassan, S. M., & Maji, A. K. (2024). Pest Identification Based on Fusion of Self-Attention with ResNet. IEEE Access, 12, 6036–6050. doi:10.1109/ACCESS.2024.3351003.

[34] Sumit, S. S., Anavatti, S., Tahtali, M., Mirjalili, S., & Turhan, U. (2024). ResNet-Lite: On Improving Image Classification with a Lightweight Network. Procedia Computer Science, 246(C), 1488–1497. doi:10.1016/j.procs.2024.09.597.

[35] Andrew, J., Eunice, J., Popescu, D. E., Chowdary, M. K., & Hemanth, J. (2022). Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications. Agronomy, 12(10), 2395. doi:10.3390/agronomy12102395.

[36] Kanda, P. S., Xia, K., Kyslytysna, A., & Owoola, E. O. (2022). Tomato Leaf Disease Recognition on Leaf Images Based on Fine-Tuned Residual Neural Networks. Plants, 11(21), 2935. doi:10.3390/plants11212935.

[37] Palma, D., Blanchini, F., & Montessoro, P. L. (2022). A system-theoretic approach for image-based infectious plant disease severity estimation. PLOS ONE, 17(7), e0272002. doi:10.1371/journal.pone.0272002.

[38] Fahmi, H., Zarlis, M., Mawengkang, H., Zendrato, N., & Sulindawaty. (2019). The Using of Thresholding and Region Merging Algorithm for Correcting the Multiple Choice Answer Sheets. Journal of Physics: Conference Series, 1255(1), 12047. doi:10.1088/1742-6596/1255/1/012047.

[39] Ahmad, N., Asif, H. M. S., Saleem, G., Younus, M. U., Anwar, S., & Anjum, M. R. (2021). Leaf Image-Based Plant Disease Identification Using Color and Texture Features. Wireless Personal Communications, 121(2), 1139–1168. doi:10.1007/s11277-021-09054-2.

[40] Nasra, P., & Gupta, S. (2024). ResNet50: Deep Learning Method for Automated Detection of Banana Leaf Spot Diseases. 2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC), 1–6. doi:10.1109/icicec62498.2024.10808847.

[41] Jain, E., & Sharma, P. (2024). Advanced Detection of Bean Leaf Disease with ResNet50: A Cutting-Edge Method for Precise Classification. 2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART), 79–84. doi:10.1109/smart63812.2024.10882591.

[42] Senthil Pandi, S., Suguna Devi, R., Bala Subramanian, C., & Kannaiah, S. K. (2024). Automatic Plant Leaf Disease Detection using AlexNet and Image Processing Techniques. 2024 International Conference on Cybernation and Computation, CYBERCOM 2024, 441–446. doi:10.1109/CYBERCOM63683.2024.10803225.

[43] Bharti, R., Srivastava, V., Bajpai, A., & Sahu, S. (2024). Comparative Analysis of Potato Leaf Disease Classification Using CNN and ResNet50. 2024 International Conference on Data Science and Its Applications (ICoDSA), 87–91. doi:10.1109/icodsa62899.2024.10651649.

[44] Arora, U., Mishra, U., Singh, S., & Singh, V. (2024). Comparative analysis of VGG16, Inception V4, AlexNet, and ResNet 50 for Plant Disease Identification. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 1–7. doi:10.1109/icccnt61001.2024.10725169.