

Received 11 March 2025, accepted 14 April 2025, date of publication 22 April 2025, date of current version 30 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3563160



RESEARCH ARTICLE

BorB: A Novel Image Segmentation Technique for Improving Plant Disease Classification With Deep Learning Models

T. OZCAN^{ID1} AND E. POLAT^{ID2}

¹Department of Computer Engineering, Erciyes University, Melikgazi, 38039 Kayseri, Türkiye

²Department of Computer Engineering, Niğde Ömer Halisdemir University, Central Campus, 51240 Niğde, Türkiye

Corresponding author: T. Ozcan (tozcan@erciyes.edu.tr)

This work was supported by the Erciyes University Scientific Research Projects Unit under Project FYL-2024-13754.

ABSTRACT Disease detection from leaf images has been among the popular studies in recent years. Classifying leaf diseases using computational methods provides great convenience for farming. In the studies carried out in this field, systems that work with high accuracy and are least affected by environmental factors that can be used in agricultural lands come to the fore. This study investigates the application of deep learning architectures for accurate and efficient plant disease detection within the context of the ongoing digital transformation of the agricultural sector. Recognizing the critical role of AI in modernizing agriculture, this research focuses on enhancing the accuracy of the classification of plant diseases. To facilitate this research, a novel dataset, “EruCauliflowerDB”, was meticulously curated, comprising high-resolution images of cauliflower plants infected with Alternaria Leaf Spot and Black Rot. The obtained EruCauliflower dataset contains 114 images from the Alternaria Leaf Spot disease class and 99 images from the Black Rot disease class. A novel integrated classification system was developed, encompassing three key stages. First, a novel segmentation method, “BorB,” was introduced to effectively isolate diseased leaf regions. This segmentation method enables us to extract features of leaf images in Lab and RGB formats. Combining the features obtained from the two image formats with the OR logical operation separates the leaf region from the background. Second, data augmentation techniques, including geometric transformations, were applied to the segmented images to enhance data diversity and improve model robustness. Finally, four state-of-the-art deep learning models—VGG16, ResNet50, EfficientNetB3, and MobileNetV3 Large—were employed for disease classification. The proposed integrated system demonstrated exceptional performance, achieving 100% classification accuracy on the EruCauliflowerDB dataset across all four models. To assess the system’s robustness, further evaluations were conducted on the independent MangoLeafBD dataset, yielding consistent results with 100% classification accuracy. The proposed Integrated Classifier method was applied by selecting 15 classes from the PlantVillage, another multi-class dataset. As a result of the experiments, PlantVillage plant leaf images were classified with 99.78% accuracy. Experimental results show that the proposed method can be effectively utilized in real-world agricultural settings to assist farmers in early disease detection, thereby reducing crop losses and improving yield quality.

INDEX TERMS Agriculture, BorB image segmentation, data augmentation, deep learning, EruCauliflowerDB, leaf disease detection.

I. INTRODUCTION

Plant production is crucial for nations’ economic growth strategies and food security. As the population in the

The associate editor coordinating the review of this manuscript and approving it for publication was Bhaskar P. Rimal.

developing world rises, food security and planning have emerged as priority challenges for nations. Combating plant diseases is a crucial aspect of ensuring food security. Plant diseases significantly diminish agricultural productivity and crop yield. Diseases affecting plant leaves impede plant development by diminishing photosynthetic efficiency and

compromising product quality. Furthermore, the marks they leave on the leaf diminish light absorption by decreasing the leaf surface area, thereby adversely impacting plant growth [1]. The timely identification of foliar diseases is crucial in this context. Classifying identified leaf diseases and identifying their types enables early intervention in the plant. The capabilities of artificial intelligence in detection and classification facilitate study in this domain.

Applying artificial intelligence to detect and classify plant diseases enables farmers to intervene early. Expert laboratory study of plant leaves is a protracted and expensive endeavor. Farmers may swiftly and consistently make decisions through an easily accessible artificial intelligence system, facilitating early disease intervention and cost reduction [2], [3]. The application of artificial intelligence in agriculture is more significant due to advancing technologies. The adoption of artificial intelligence in agriculture has accelerated due to advancements in image processing and big data. Deep learning and machine learning research assist farmers in making educated decisions by swiftly processing agricultural data. Artificial intelligence systems identify plant diseases and detrimental elements that may adversely impact plant growth [4]. Deep learning techniques yield insights such as disease assessment and plant health evaluation by analyzing numerous plant images. Monitoring the products cultivated on agricultural grounds is essential to enhance the productivity and quality of crops in agriculture. Deep learning techniques are commonly employed to monitor and analyze agricultural products using images [5]. Deep learning techniques have been employed to address intricate issues rapidly and efficiently, facilitating their application in agriculture. The study examined the efficacy of deep learning techniques for image analysis in recognizing barriers encountered by autonomous agricultural machinery and assessing their impact on the safety of such vehicles. The obstacle recognition investigation revealed a prediction accuracy of 99.9% for certain items in the rowed agricultural field and 90.8% in the grass-mowing field; nevertheless, it demonstrated poor predictive capability for certain impediments, such as humans.

The presence of fungi, microorganisms, and bacteria on plants can diminish the productivity of the cultivated crop. If these problems affecting the facility are not addressed promptly, significant economic losses will ensue. Nonetheless, the unintentional application of pharmaceuticals in research aimed at disease prevention in foliage adversely affects the environment and natural ecosystems. The excessive use of pharmaceuticals might adversely impact the natural water and soil cycles. In plant disease prevention, image processing and deep learning are often techniques for detecting plant leaf diseases. One of these studies [6] employed deep learning convolutional neural network designs for disease detection with tomato leaf images. This study involved training images of tomato leaves using the convolutional neural network architectures

AlexNet, ResNet, and GoogLeNet to analyze the characteristics of diseases. The study revealed that the ResNet architecture identified tomato plant diseases with an accuracy of 97.28% during the training experiment utilizing the Stochastic Gradient Descent (SGD) optimization algorithm. A hybrid classifier method proposed for detecting diseases on plant leaves, [7] aims to classify leaf diseases of the bell pepper plant. The investigated method includes image preprocessing, feature extraction, and classification stages. The combined features of local binary pattern (LBP) and VGG-16 structures were used for feature extraction. Random forest (RF) was used in the classification stage. As a result of the experiments, a 99.75% accuracy rate was achieved with random forest using VGG16 and LBP feature extraction methods. A study on olive plant leaf diseases [8] established a specialized structure integrating Convolutional Neural Network architecture with Vision Transformer architecture. This structure aims to discover and classify diseases that may impact olive leaves. This research examined binary and multiple classifications. Binary classification is employed to determine the presence or absence of a disease, while multiple classification is utilized to identify the specific type of disease. The investigation revealed that the proposed structure achieved 96% accuracy in multiple classifications and 97% accuracy in binary classification. A study [9] was undertaken on pixel-based labeling and categorizing deterioration in plant leaves, focusing on automatically identifying deteriorated areas in images captured in a garden setting. To do this, images were transformed into HSV, YUV, Lab, and RGB formats, and the Otsu thresholding method was applied to the channels containing specific information in each image format. This procedure allows for the acquisition of properties of the warped leaf sections to a specific extent. The study employed a two-stage technique to identify and categorize leaf degradation. Initially, a CNN-based pixel classifier was used to distinguish healthy leaves from damaged ones. In the second stage, the compromised leaf pixels were classified internally. A study on disease detection in potato plants [10] examined the categorization of diseases in potato leaves utilizing the Plant Village plant disease dataset, an open resource. This study aimed to differentiate between three distinct classes, comprising images of two potato plant diseases and healthy plants. Leaf images with RGB characteristics were transformed into Lab image format, and the b channel was isolated from the other channels. A threshold value was established on this channel to eliminate the background. The images acquired from the feature extraction process were classified into late blight, early blight, and disease-free categories using a Support Vector Machines classifier. The study suggested a computationally efficient model that attained a 95% accuracy rate.

Different data sets are used in studies conducted to detect and classify leaf diseases. An important problem in these data sets may be the imbalance of the number of samples.

In cases where the number of samples is imbalanced, the accuracy of the class with the majority of data tends to be high. In classes with a small number of data, performance tends to be low due to low diversity and randomness [11]. Data sets obtained in laboratory environments are more suitable for feature extraction with image preprocessing steps. Controlled lighting and a homogeneous background in the environment ease extracting leaf features. Such data sets help deep learning models achieve higher results. In addition to images taken in laboratory environments, field images are taken from the garden or field. These real-world data sets may have different lighting environments, difficult shooting conditions, and background clutter. These situations may make it difficult to detect plant diseases. For this reason, other methods are being developed in the literature for real-world datasets. In a study conducted in this field [12], the SaudiArabiaFlora Dataset belonging to some plant species in Saudi Arabia was created. In this study, a special structure called MIV-PlantNet was developed for the classification of plants. This classifier structure, produced using the features of MobileNet, Inception, and VGG architectures, was improved to classify the created real-world dataset. As a result of the study, a 99% accuracy value was reached. In a study investigating the detection of situations such as ripening and damage in soybean plants with the deep learning method [13], techniques that will increase the performance of InceptionV3 architecture were examined. In this study, five different image classes belong to soybeans. The study investigated the effects of the layers added to the InceptionV3 architecture and the pre-trained model on performance. Compared to other studies on the same dataset, a successful value was achieved by reaching an accuracy rate of 98.73%. In another study on plant diseases, diseases belonging to the sunflower plant were examined [14]. The study compared the classification performance with the AlexNet, VGG16, InceptionV4, MobileNetV3, and EfficientNetB3 deep learning architectures commonly used in the literature. The study also examined the impact of the transfer learning method on disease detection in the sunflower plant. The studies revealed that the EfficientNetB3 architecture outperformed other architectures regarding results.

Many different methods have been preferred for classifying plant diseases in the literature. One is to improve the structure of deep learning architectures to increase classification accuracy. In many studies, an increase in accuracy rates has been observed with the proposed new deep learning architectures. Another perspective on plant disease classification has been examining the effect of transfer learning on state-of-the-art models. The effects of the transfer learning method on deep learning architectures, which are frequently preferred in detecting plant diseases, have been examined. Along with these studies, there are also studies in which the analyses are improved by performing operations on the images in plant disease datasets. These studies include systems that separate and analyze the leaf region from the

background. The applicability of these methods for each dataset and environmental factor varies in many studies [15]. This study aimed to investigate the most appropriate method to classify the leaf images in the EruCauliflowerDB dataset obtained. Considering the reviewed literature studies, the proposed method was to analyze the leaf regions desired to be classified by extracting them from the noise. Some segmentation methods in the literature have been tried in this field, and the BorB segmentation method has been proposed. As a result of the experiments, it has been seen that the developed BorB segmentation method is more suitable for the EruCauliflowerDB dataset. Each of the diseased leaves collected for the dataset has its characteristics. Since few diseased leaves are obtained in this area, there are few images in the dataset. The data augmentation technique has been tried in this study to solve one of the problems in plant disease datasets: data scarcity. To increase the effect of the data augmentation technique, the segmentation process of the leaves is performed afterward. The dataset to which preprocessing steps have been applied has been made ready for training with deep learning architectures. At this stage, the transfer learning method has been used for deep learning architectures, considering that it will increase the accuracy rate in training. The effect of each method applied is presented in the results section, and detailed analyses are provided. In this way, each stage of the integrated classifier method presented in the research has become valuable. This study explores the application of deep learning techniques for accurate and efficient plant disease detection, a critical aspect of modern agriculture. In response to the growing global demand for food and the increasing prevalence of plant diseases, this research aims to enhance crop productivity and quality by enabling timely and accurate disease identification. To address this challenge, a novel dataset, "EruCauliflowerDB," was meticulously curated, comprising high-resolution images of cauliflower leaves affected by Alternaria Leaf Spot and Black Rot. This dataset is valuable for training and evaluating plant disease detection models. The proposed methodology integrates several key components: Image Segmentation: A novel segmentation technique, "BorB," was developed to accurately isolate diseased leaf regions from the background, improving the quality of input data for the subsequent stages. Data Augmentation: Geometric transformations were applied to the segmented images to increase data variability and enhance the model's generalization capabilities. Deep Learning Models: Four state-of-the-art deep learning architectures (VGG16, ResNet50, EfficientNetB3, and MobileNetV3 Large) were employed for disease classification. The proposed integrated approach demonstrated exceptional performance, achieving 100% classification accuracy on the EruCauliflowerDB dataset across all four models. Furthermore, the method's robustness was validated on the independent MangoLeafBD dataset, achieving 100% accuracy on the test data. The proposed method was tested with the PlantVillage dataset,

a common dataset on plant diseases. In tests conducted with 15 classes in the PlantVillage dataset, the proposed Integrated Method reached an accuracy rate of 99.78%.

The main contributions of the paper are given below:

- This paper presents a new dataset, **EruCauliflowerDB**, consisting of cauliflower leaf images, to evaluate the proposed methods' effectiveness.
- A novel technique termed **BorB** is introduced for image segmentation. This method introduces a specialized algorithm that distinguishes leaf regions from background information on a pixel-by-pixel basis.
- An integrated method including image segmentation, data augmentation, and deep learning components is proposed.
- The proposed integrated method obtained 100% test accuracy on the EruCauliflowerDB dataset.
- The proposed integrated method, which has the best accuracy rate of 100%, outperforms the existing work [16] that reported a 99.23% accuracy on the MangoLeafBD dataset.
- The proposed integrated method obtained 99.78% test accuracy on the PlantVillage dataset.

The rest of this paper is organized as follows: The materials and methods are presented in Section II. Section III describes the experimental studies, including experimental setup and results. Lastly, Section IV summarizes the conclusion and future works.

II. MATERIALS AND METHODS

A. EXPERIMENTAL SETUP

The hardware and software setups are detailed in the following subsections.

1) HARDWARE SETUP

The images of the dataset were captured using a Redmi Note 9 Pro smartphone, featuring a 64MP primary camera, an 8MP wide-angle lens, and a 2MP depth sensor. Deep learning training was carried out on a computer with the following features: CPU (Intel i7-12700H), 16GB RAM, NVIDIA RTX 3060 GPU.

2) SOFTWARE SETUP

The software, written using Python version 3.7.0, utilized the TensorFlow 2.8.0 library for deep learning and data augmentation techniques. The OpenCV library version 4.9.0 was utilized for image processing and analysis. The Anaconda Spyder application was selected as the Python development and compilation environment.

B. DATASETS

Two distinct data sets were employed for the study. The EruCauliflowerDB dataset, developed for this work, was utilized during the training and testing phases of the suggested approaches. The second dataset comprises images of damaged mango tree leaves. Another publicly available

dataset in the study contains diseased and healthy leaf images of tomato, potato, and pepper plants.

1) EruCauliflowerDB DATASET

The dataset contains images of cauliflower sourced from sick leaves gathered in the cauliflower farming region of Mersin (Silifke Tea Farm). After the pertinent expert engineer conducted studies, the infected leaves were categorized into groups. After this phase, images of each category were captured using the camera of a Redmi Note 9 Pro mobile phone. Two distinct classes were identified in the images of sick leaves captured with this camera. The classes are designated as Alternaria Leaf Spot and Black Rot. The quantity of images acquired from two distinct classes varies. Table 1 presents the number of images in each category.

TABLE 1. The count of the images in EruCauliflowerDB.

Disease	Number of Images
Alternaria Leaf Spot	114
Black Rot	99

Images of diseased leaves of cauliflower plants were obtained in a laboratory environment. Dataset images have background information that does not change. Since the images obtained were taken at different times, there are differences in lighting and shading. The image of plant leaves was taken from an angle from above. The dataset samples are given in Figure 1.



(a) Alternaria Leaf Spot



(b) Black Rot

FIGURE 1. Data samples for EruCauliflowerDB.



(a) Anthracnose disease



(b) Bacterial Canker

FIGURE 2. Data samples for MangoLeafBD.

2) MangoLeafBD DATASET

The MangoLeafBD dataset [17], comprising images of mango leaves, has seven distinct disease categories. The publicly accessible mango leaf disease dataset includes a category for healthy leaves with infected leaves. The collection comprises images from eight categories, each including 500 leaf images. The mango dataset comprises a

total of 4,000 leaf images. The photographs were captured with a mobile phone camera. The identified classes are Anthracnose, Bacterial Canker, Cutting Weevil, Die Back, Gall Midge, Powdery Mildew, Sooty Mold, and Healthy Leaf. The dataset samples are given in Figure 2.

3) PlantVillage DATASET

Plant Village [18] is an image dataset of 54,309 samples depicting leaf diseases on a homogeneous background. To create the PlantVillage dataset, four to seven photographs of each leaf were taken with a standard 20.2-megapixel point-and-shoot camera, the Sony DSC-Rx100/13, in automatic mode. 15 classes were selected from the dataset to evaluate the Integrated classifier method proposed in the study. There are 20,638 images in the selected classes. Each class in the dataset and the number of images belonging to that class are given in Table 2.

The dataset samples are given in Figure 3.

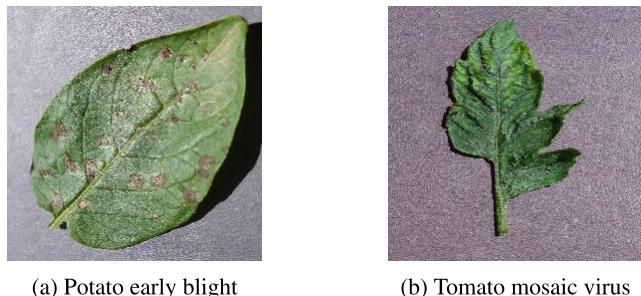


FIGURE 3. Data samples for PlantVillage.

TABLE 2. Image classes and numbers used in this study from the PlantVillage dataset.

Class Name	Number of Images
Pepper bell bacterial spot	997
Pepper bell healthy	1478
Potato Early blight	1000
Potato healthy	152
Potato Late blight	1000
Tomato Target Spot	1404
Tomato mosaic virus	373
Tomato YellowLeaf Curl Virus	3209
Tomato Bacterial spot	2127
Tomato Early blight	1000
Tomato healthy	1591
Tomato Late blight	1909
Tomato Leaf Mold	952
Tomato Septoria leaf spot	1771
Tomato Spider mites Two spotted spider mite	1676

C. PERFORMANCE METRICS

As a result of the training for classification with deep learning, several evaluation metrics are used to evaluate the outputs obtained from the experiments. In this study, the confusion matrix given in Figure 4 was used to compare and evaluate the results of the methods and experiments. In addition to the complexity matrix, metrics such as precision, recall, F1 score, and accuracy rate were used to

analyze the sensitivity of the experiments. The formulae related to these metrics are given in Equations 1-4.

A high precision value indicates that no false positive predictions are made. The recall value indicates how many correct predictions were made from the relevant class. The F1-Score parameter is often used to measure the performance of classes with unbalanced data. Thanks to these parameters calculated in datasets with many classes, success rates are analyzed more accurately. The Precision value given by the equation (1) below provides the ratio of the data predicted as positive to the ones that are actually positive. True Positive (TP) in the equation refers to the number of samples the model predicts as positive and positive, and False Positive (FP) refers to the number of examples that the model predicts as positive but is actually negative. The Recall value given by Equation (2) represents the rate the model correctly predicted among all true positive examples. In the equation, the True Positive (TP) value is divided by the sum of the True Positive (TP) and False Negative (FN) values. False Negative (FN) values are samples that the model predicted were negative but were actually positive.

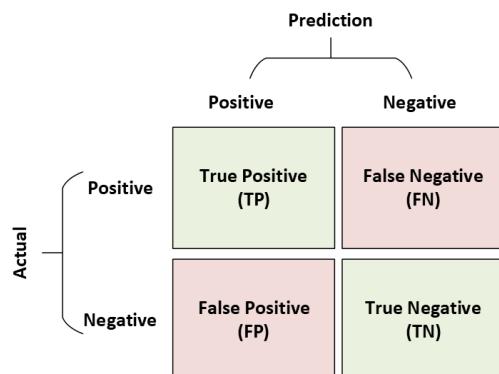


FIGURE 4. Confusion matrix.

In the equation given by Equation (3), the F1-score value is calculated. It is preferred that the overall performance of the model be measured. In this equation, the value ranges between 0 and 1, and the closer it is to 1, the better the model's performance. Equation (4) shows the proportion of examples the model correctly predicted in the entire data set. The Accuracy value is obtained by dividing the sum of the True Positive (TP) and True Negative (TN) values, which indicate negative correct predictions, by the sum of the True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) values.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

$$F1 - score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

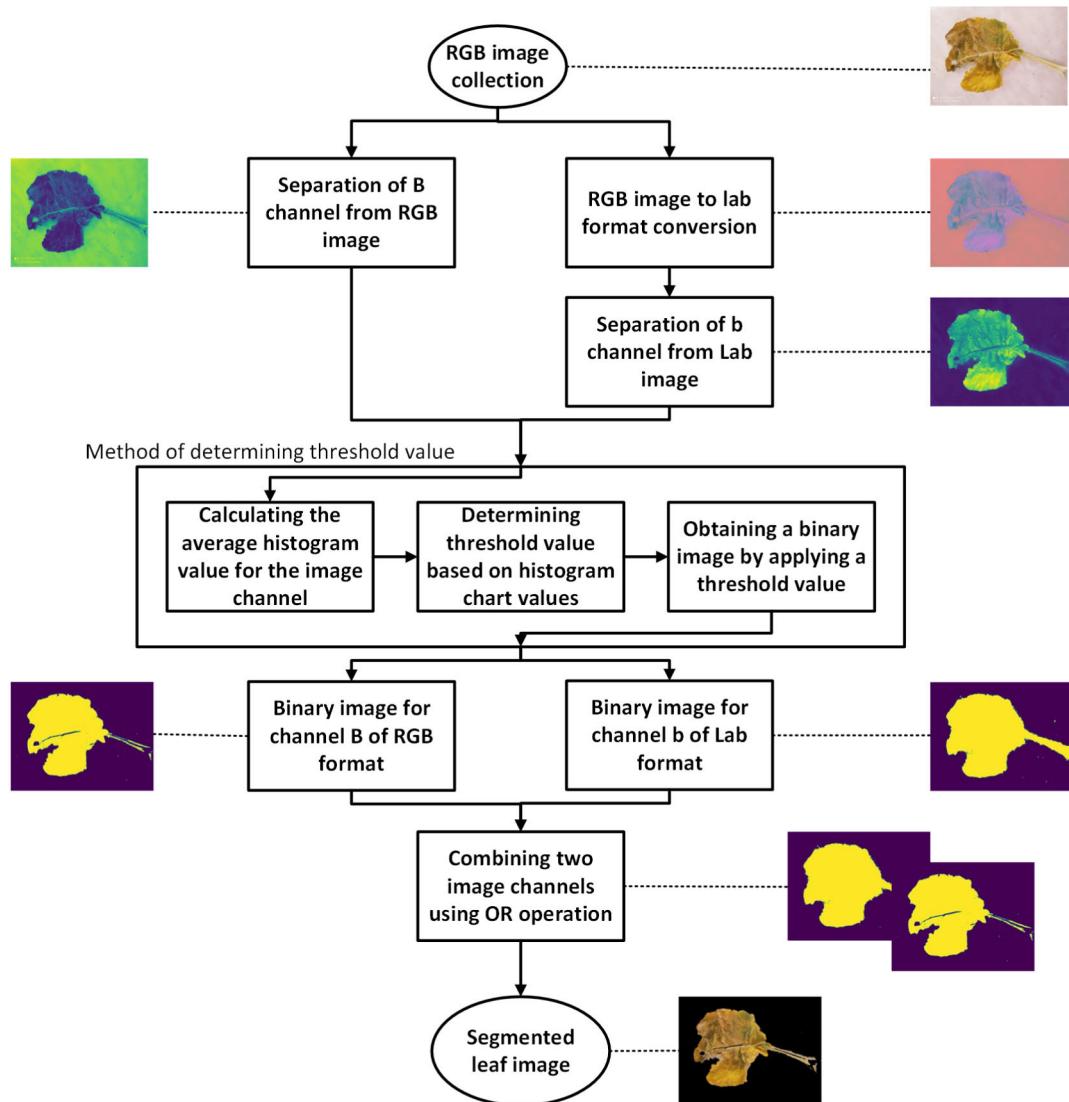


FIGURE 5. Proposed BorB segmentation method.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (4)$$

D. BORB: A NOVEL IMAGE SEGMENTATION METHOD

The datasets obtained to classify leaf diseases contain a small number of images specific to the problem. This requires the deep learning architecture to perform the analysis and learn with a small amount of data. In this case, background regions other than leaf regions are also considered in the study. This situation causes classifier deep learning architectures to analyze unnecessary regions in the image. The BorB segmentation method, which separates leaf regions from the background, has been added to the preprocessing step in the Integrated Method to improve this situation.

This paper proposes a segmentation method utilizing image processing techniques. Images of sick leaves are extracted from the background before the classification phase. As depicted in the algorithm flowchart in Figure 5, the

initial phase is acquiring RGB images. In this phase, images of sick leaves are captured utilizing the camera specifications outlined in the dataset section. Subsequently, RGB images are transformed into Lab format, and this format's three channels (l, a, b) are extracted. The RGB images were concurrently processed by isolating their channels (R, G, B). During this investigation phase, channel b of the Lab format and channel B of the RGB format underwent histogram mean computation for analysis and segmentation processing. At this stage, the pertinent channel of each image in the dataset was retrieved and incorporated into the histogram mean computations. The histogram mean value graph for the B channel in RGB format and the b channel in Lab format was derived from these processes for the leaf disease dataset. Subsequently, a study was conducted to ascertain a uniform threshold value inside the dataset, utilizing the histogram value graph. This technique enabled the identification of the critical threshold value for the segmentation step.

The threshold settings were applied to the Blue channel of the RGB image and the b channel of the Lab image according to the results obtained from the histogram mean value graph given in Figure 6 and Figure 7. The graphs show the selected threshold value with a red dashed line [19]. The green area was chosen to remove leaf pixels from the image. The histogram graph was meticulously analyzed, and as a consequence of the experiments, a threshold value of 135 was applied to the b channel of the Lab format, resulting in the effective extraction of relevant leaf information, as illustrated in Figure 8 (e) and (f). A threshold value of 130 was designated for the Blue channel extracted from the RGB image by histogram analysis, resulting in the generation of binary images based on this threshold. Figures 8 (c) and (d) illustrate the binary images generated after the application of the threshold to the Blue channel of the RGB image. These procedures ensured that the intended leaf region was accurately separated during the segmentation process.

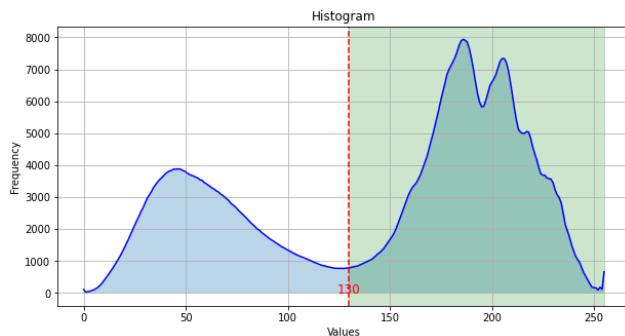


FIGURE 6. Histogram graph for B channel of RGB image.

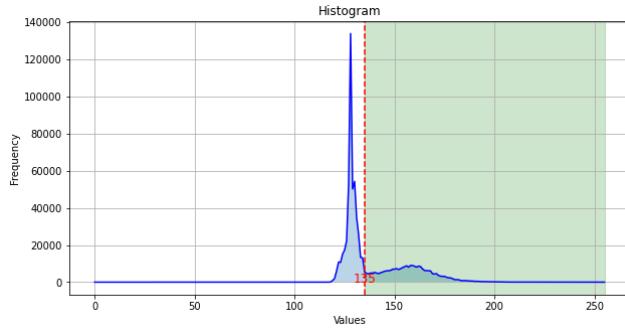


FIGURE 7. Histogram graph for b channel of Lab image.

The images generated using the specified threshold values were overlaid onto the binary channel images, Blue for RGB and b for Lab (blue to yellow), with each pixel position assigned a value of 1 designated as leaf information. Consequently, the selected pixel locations in the composite image were retained in the actual RGB image, while all other pixels were converted to black ($R=0$, $G=0$, $B=0$). The results of this method are depicted in Figures 8 (g) and (h), illustrating that the leaf region is effectively segmented while other areas are omitted. This phase is essential to maintain

the segmentation process's integrity and segregate the leaf information distinctly.

To compare the proposed BorB segmentation method with state-of-the-art segmentation methods, experimental studies were carried out by using the Otsu threshold and K-means clustering methods, which are frequently used in the literature [20], [21]. Figure 9 compares the result images obtained with BorB segmentation with the results obtained with Otsu thresholding and K-means clustering.

When Figure 9 is examined, it is seen that pixels showing leaf features are obtained more losslessly with the BorB method for this dataset. According to Figure 9, the loss of leaf parts is the highest with the Otsu method.

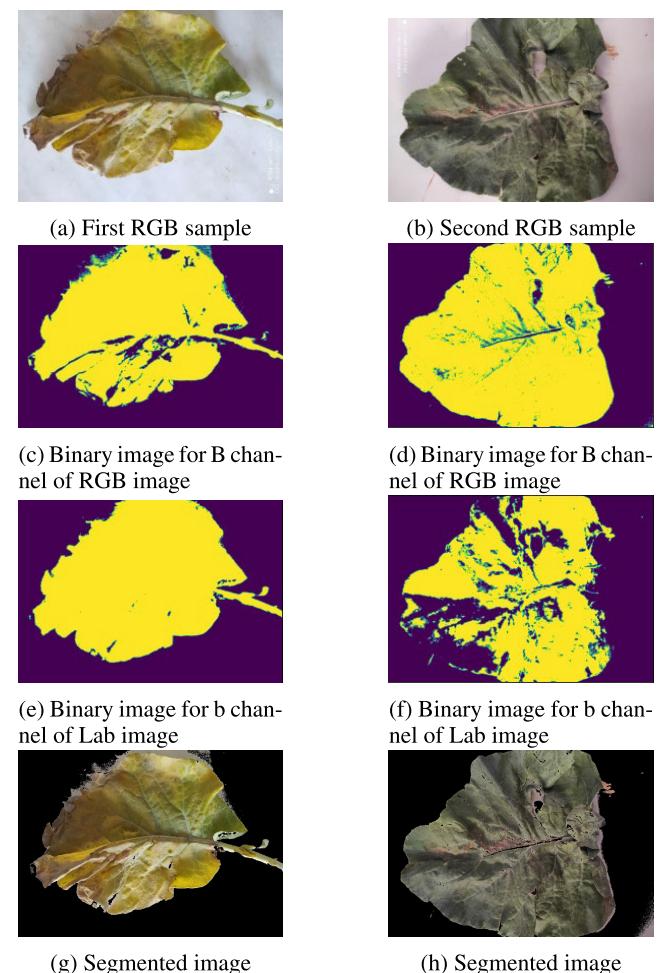


FIGURE 8. The segmentation examination of two distinct samples.

E. DATA AUGMENTATION METHODS

This study employed geometric transformation for data augmentation to enhance the efficacy of deep learning models in classifying diseases in cauliflower leaves. Data augmentation is a technique that enhances a constrained data set by producing diverse variations and enlarging the dataset. This method seeks to improve the generalization capacity of deep learning models employed in classification tasks and to achieve greater accuracy across various scenarios [22].

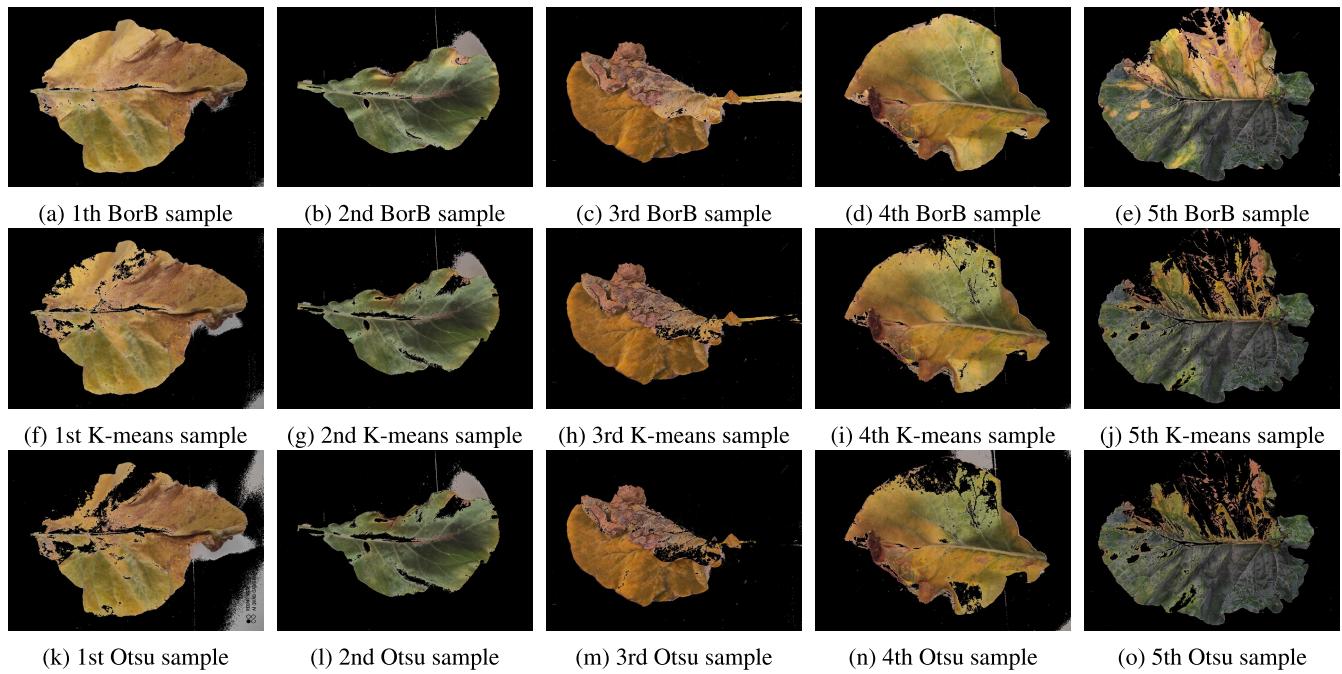


FIGURE 9. Results of different segmentation methods for EruCauliflowerDB dataset images.

In this study, different geometric transformation methods were applied within the scope of data augmentation. The transformations used include random rotation, horizontal and vertical rotation, zooming, random panning, and random brightness change. As a result of these operations, the areas where empty pixels occur in the images are filled with values close to the closest pixel values in the original image, and it is aimed to preserve the features of the actual image. In this study, data augmentation methods are applied dynamically only in the training phase, not to directly increase the number of images in the dataset. At each training step, the available images in the dataset were diversified and presented to the deep learning models so that the models could learn over a wider variation. Table 3 shows the random value ranges of the geometric operations applied. The rotation, pan, and zoom ratios in this table are expressed in degrees. The brightness value range is randomized between 0.5 and 1.5, representing the increase or decrease of brightness. Sample images of the data augmentation method are presented in Figure 10, visually demonstrating the applied methods' diversity.

TABLE 3. Geometric transformation parameters.

Parameters	Values
Rotation degree	40
Zoom degree	11.5
Shear degree	11.5
Brightness range	0.5-1.5

F. DEEP LEARNING METHODS

Deep learning (DL) is a subset of machine learning that employs multi-layer artificial neural networks. The deep learning model emphasizes acquiring more intricate features



(a) Image with data augmentation applied
(b) Segmented image with data augmentation applied

FIGURE 10. Data augmentation method examples.

than traditional machine learning. In contrast to machine learning, it eliminates the necessity for manual data preparation. It is often utilized because of its capability for feature extraction. Convolutional Neural Networks (CNNs) are the most prevalent deep neural networks. CNN is an artificial neural network that has transformed the image processing domain. Throughout the training process, the network identifies the features of the images by executing a form of “tapping” operation on the input data. CNN can yield more sensitive and precise outcomes than conventional image processing techniques in face recognition, object detection, and image classification applications. The layers, categorized as input, hidden, and output, facilitate learning by identifying patterns and relationships within the data.

The deep learning architectures used in the study were preferred as pre-trained models. Since the number of images was relatively small, the pre-trained model was thought to contribute to the training accuracy. The preferred optimizer

function for the training stages was the Adam optimizer, which is frequently used in the literature. Effective results in the field of plant diseases have been observed in a study [23].

1) VGG16

The VGG16 deep learning model, developed by Simonyan and Zisserman in 2014, is a significant Convolutional Neural Network (CNN) within deep learning models. It is extensively utilized in numerous computer vision applications, including image classification and object recognition. The model comprises 16 convolutional layers and features a deep network architecture. This model has convolutional layers, each featuring filters of dimensions 3×3 [24]. The selection of diminutive filter sizes enables the model to possess fewer parameters. This accelerates the training process and facilitates effective outcomes with less computational expense. The VGG16 model possesses pre-trained weights derived from the ImageNet dataset [25]. The pre-trained weights facilitate rapid and efficient training of the model in transfer learning applications. Consequently, VGG16 is extensively utilized in numerous image processing and classification endeavors due to its superior performance. The success of the VGG16 architecture in extracting features from images is frequently encountered in studies in the literature. In a study using three-dimensional (3D) images for early lung cancer diagnosis, VGG16 was used as a feature extractor in the proposed model [11].

2) ResNet50

The ResNet50 deep learning architecture possesses a deeper network topology than alternative architectures. A more profound network architecture results in diminishing gradients at each layer. The gradient is computed at each layer of the neural network during backpropagation. While a more significant number of layers is anticipated to facilitate the learning of more intricate features, it is occasionally evident that the values derived from the activation functions are diminished. To address this scenario, the ResNet50 architecture includes a component known as the Residual Building Block [26]. This topology resolves the vanishing gradient problem by bypassing the convolutional layers. Figure 11 illustrates the architectural composition of the Residual Building Block.

The ResNet50 deep learning model comprises 50 layers and is trained on the ImageNet dataset. Using this pre-trained network structure, successful results have been obtained on small data sets [27]. This CNN model comprises three convolutional layers arranged in the sequence of 1×1 , 3×3 , and 1×1 . This triadic configuration is referred to as a bottleneck. This structure accelerates the training period.

3) EfficientNetB3

The EfficientNetB3 deep learning architecture was first presented in 2019 [28]. For classical deep learning models,

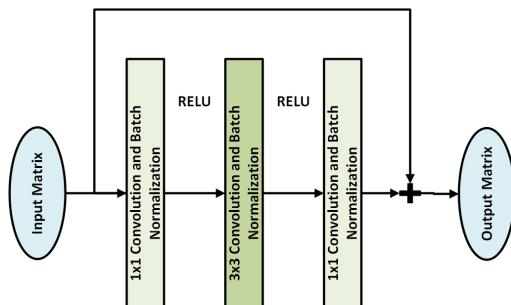


FIGURE 11. Residual building block.

scaling the number of neurons and layers is applied to increase accuracy or training speed. Some deep learning models aim to analyze more information by keeping the resolution of the images obtained high. In EfficientNet deep learning architectures, Compound Coefficient is used for this scaling process. Instead of randomly scaling the number of layers, neurons, or image resolution with composite scaling, each dimension is balanced with a specific fixed scaling coefficient. As a result, it was observed that the EfficientNet model achieved successful results with many datasets [29]. In the convolution block of this model, there is a structure called Mobile Inverted Bottleneck. This structure first appeared in the MobileNet architecture and was also used in this architecture. Inverted residual connections and linear bottlenecks used in MobileNetV2 are included in the structure [30]. In the EfficientNet architecture, Swish and Global Average Pooling between layers are used as activation functions.

4) MobileNetV3 LARGE

MobileNetV3 [31] is a deep learning architecture designed for mobile and resource-limited environments. This architecture has come to the forefront with its low processing power requirement and fast operation. There are two different utilized architectures: MobileNetV3 Large and MobileNetV3 Small. MobileNetV3 Large structure is used in applications where higher accuracy is required and was also used in this research. MobileNetV3 Small is lighter than Large. This feature makes it suitable for real-time applications [32]. In this architecture, it uses the Hard-Swish function, which is a more efficient version of the Swish activation function. MobileNetV3 Large architecture has a bottleneck block like MobileNetV2. Each bottleneck block includes three layers. These are layers that perform 1×1 , 3×3 , and 1×1 convolution operations.

In the MobileNetV3 architecture, in the first part, dimension reduction is performed in the image channel to reduce the computational cost of the input image matrix. In the size Reduction Block that follows, the image size is reduced, but the number of image channels is not reduced. The last layer block expands the image channel and makes it available in subsequent layers [33].

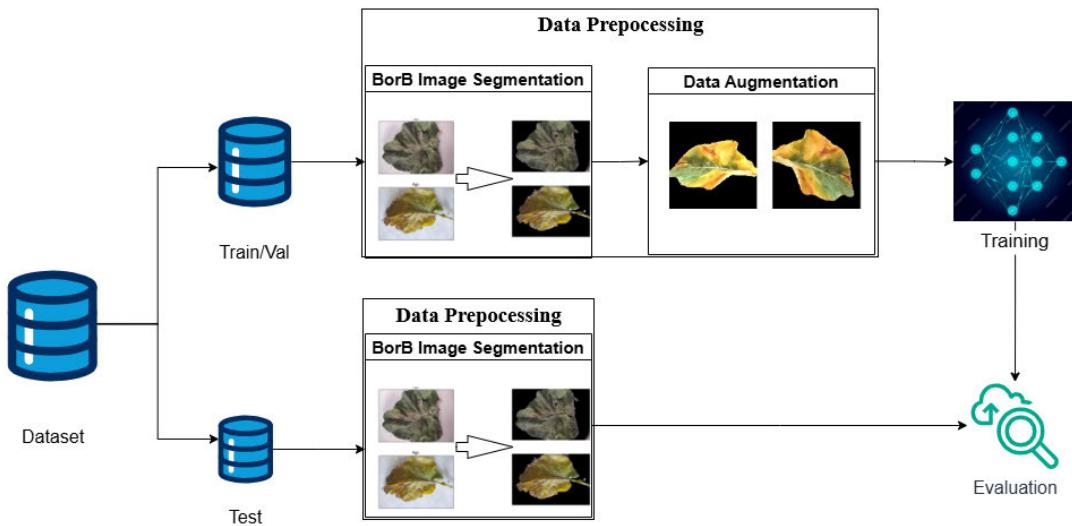


FIGURE 12. Proposed integrated classification system.

G. PROPOSED METHOD

This paper proposes an integrated classification method consisting of three main stages, including a new image segmentation method named BorB. The first step of the method is image segmentation. The second phase is data augmentation, and the last phase is training with random search-based hyperparameter-tuned pre-trained CNN models. The stages of the proposed method are given in Figure 12.

Image segmentation is applied to the leaf images separated from the background, and image augmentation methods are performed. The obtained data is given to the training steps in four different deep learning architectures. After the training steps, the test images are given to the trained models, and the results are obtained.

III. EXPERIMENTAL RESULTS

Simulations were performed on EruCauliflowerDB, MangoLeafBD, and PlantVillage datasets to test the proposed method's success. The experimental results obtained are presented in detail in the subsections.

A. SIMULATION RESULTS FOR THE ERUCAULIFLOWERDB DATASET

The EruCauliflowerDB dataset of the cauliflower plant was created to detect diseases in cauliflower leaves. Data augmentation and image segmentation processes were applied to this data set using image processing methods. The data obtained after image processing methods were given to deep learning architectures such as ResNet50, EfficientNetB3, VGG16, and MobileNetV3 Large for classification purposes. Test and validation images were kept the same in this study for all methods. In this way, the effect of each method on the accuracy value was tested on the same images, and the results were obtained. The training was conducted under the

same conditions and in the same environment to compare different deep-learning models and techniques. As a result of the training, it was aimed to predict the correct classes of diseased cauliflower leaf images.

1) SELECTION OF THE SPLIT RATIO FOR THE DATASET

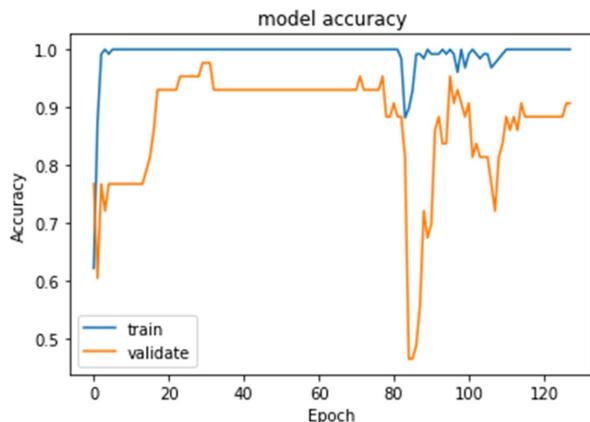
In the initial phase of this study, meticulous research was conducted to determine the optimal data-splitting ratio for the experimental setup. The data separation process significantly influences the learning process, as the data allocated for training directly impacts the model's generalization ability. To this end, various approaches were evaluated, including scenarios with a high proportion of data assigned to the training set and scenarios with balanced data distribution across training, validation, and test sets. Through rigorous experimentation, the most effective data-splitting ratio was identified.

The results obtained from deep learning architectures for the data set separated by different percentages are shown in Table 4 to evaluate these other approaches. Training and testing phases were performed with VGG16, EfficientNetB3, ResNet50, and MobileNetV3 Large models for each percentage data separation.

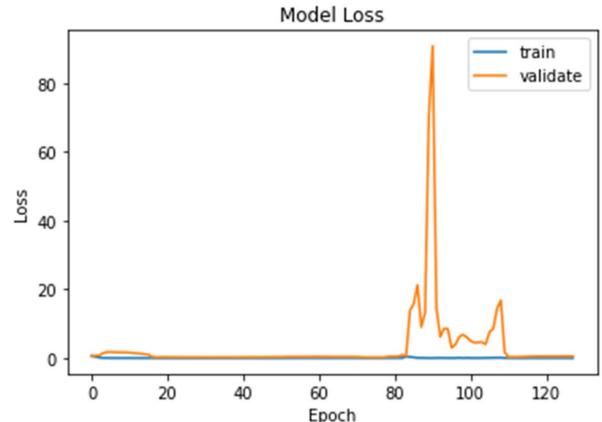
The ResNet50 model gave the highest accuracy rate, with 95.34% in the training. Considering these results, splitting as 60% training, 20% validation, and 20% test was preferred for the following stages of the study.

2) RESULTS FOR THE PROPOSED METHODS

Since the EruCauliflowerDB dataset was created just for this research, variant methods containing subcomponents of the proposed method were used in experimental studies to compare the proposed one. Variant models are presented in Table 5.



(a) Accuracy



(b) Loss

FIGURE 13. ResNet50 training steps with actual (nonprocessed) dataset.**TABLE 4.** Experimental results to determine the dataset partition ratio.

Data Split Ratio	CNN	Test Accuracy
80%, 10%, 10%	VGG16	86.39
	EfficientNetB3	81.81
	MobileNet V3	86.36
	ResNet50	90.91
70%, 15%, 15%	VGG16	84.84
	EfficientNetB3	87.87
	MobileNet V3	93.93
	ResNet50	90.9
60%, 20%, 20%	VGG16	93.02
	EfficientNetB3	90.69
	MobileNet V3	88.37
	ResNet50	95.34

TABLE 5. Variant methods for comparison.

Variant Methods	Subcomponents
Method 1	Base Dataset & Pre-trained Model
Method 2	Base Dataset & Data Augmentation & Pre-trained Model
Method 3	Segmented Dataset & Pre-trained Model

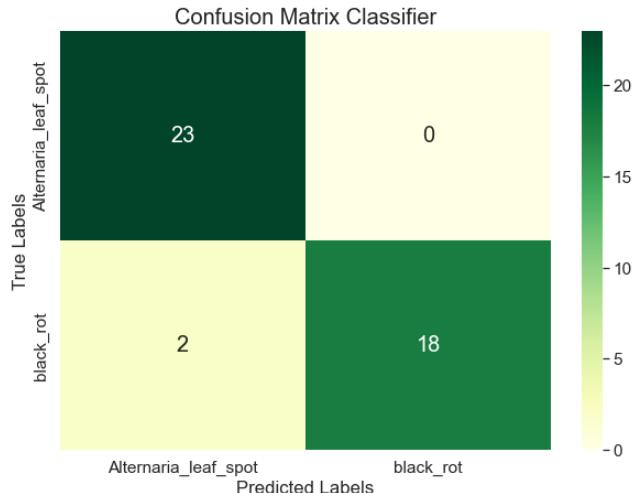
The training parameters given in Table 6 are standard for all methods. The validation and test images used for the methods are selected similarly for each deep learning model.

TABLE 6. Training parameters.

Parameters	Values
Input image size	224x224
Learning rate	0.0001
Batch size	16
Training steps	150
Optimizer	Adam
Optimizer decay	0.0001
Loss	Categorical
Early stop	50

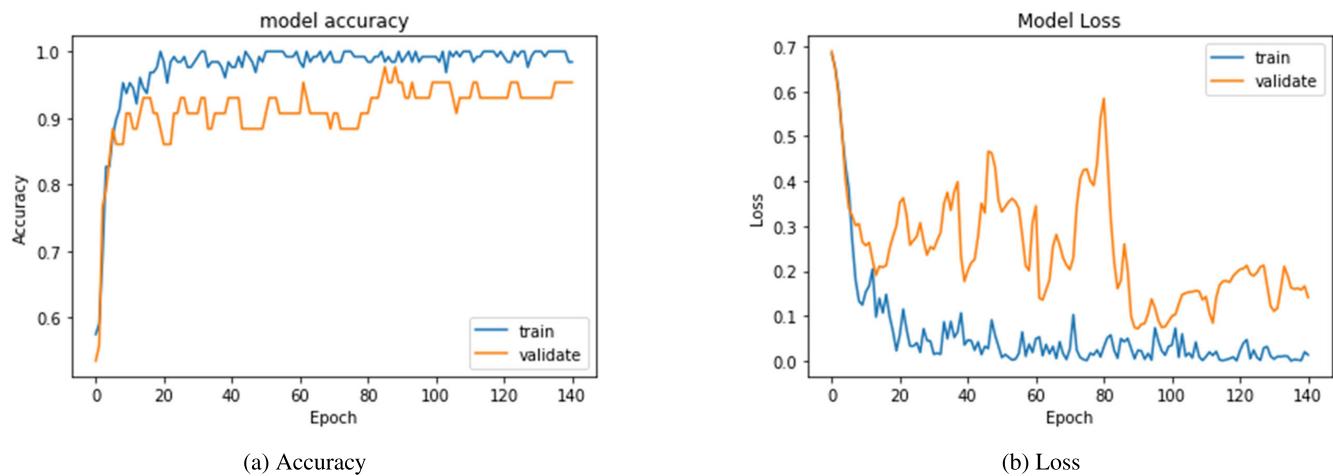
a: RESULTS FOR METHOD 1

This method gives the dataset to deep learning models for training without applying image processing and data

**FIGURE 14.** Confusion matrix for Method 1.

augmentation steps. Table 4 shows the results obtained for this method, divided into 60% training, 20% validation and 20% testing. No processing was applied to the data in this part of the study. Since image processing steps were not applied, the values obtained in Table 4 gave valid results for this method. When the results are analyzed, the best accuracy value is achieved by the ResNet50 model with 95.34%. In Figure 13, the accuracy and loss values observed in the training steps are visualized on the graph.

After the training stages, the results were obtained with the test folder containing the same images used in other methods. A confusion matrix was created for the correct analysis of the results, in which the number of correct and incorrect predictions for each class could be analyzed. Figure 14 shows the confusion matrix output of the ResNet50 model training of Method 1. Here, all predictions are correct except for two samples. The possible reasons for misclassifications are visual similarity, image quality, and limited training data.

**FIGURE 15.** EfficientNetB3 training steps with augmented dataset.**b: RESULTS FOR METHOD 2**

In this second method, data augmentation and variations were performed using image processing methods on actual data. The images of each class in the training folder from the EruCauliflowerDB dataset were diversified using geometric transformations before proceeding to the training phase. These are geometric operations such as random rotation, random shear, random zoom, random horizontal flip, random brightness, and random vertical flip. As a result of these operations, if there are empty parts in the images, the empty areas are filled with values similar to the nearest pixel values. In this way, it is aimed to provide diversity in each training step without losing the features of the image. The data augmentation method was applied for all four CNN architectures, and training was conducted. With these operations, the amount of data was not increased in the folder using the data augmentation method; only the data diversified at each step in the training phase was given to the training. Diversified data were included and shuffled in the training steps with actual data. The training results based on these conditions are presented in Table 7.

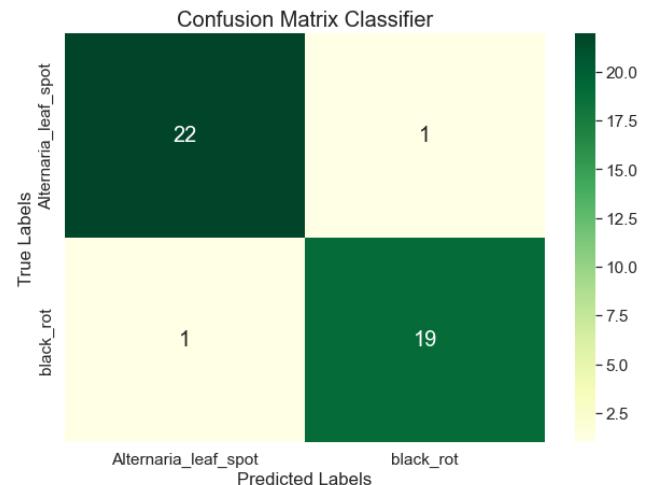
TABLE 7. Results obtained with Method 2.

Model	Early Stop	Training Accuracy	Test Accuracy
VGG16	100	95.28	95.34
EfficientNetB3	141	98.43	95.34
MobileNetV3 Large	83	98.43	93.02
ResNet50	128	100	93.02

The accuracy rates shown in Table 7 show that EfficientNetB3 and VGG16 models reach the same value with a 95.34% test accuracy rate. In addition, the EfficientNetB3 model reached a 98.43% accuracy rate in the training data and reached the highest result compared to other models. In the training, 141 training steps were performed until this result was reached, and when the early stopping condition was reached, the training was stopped by storing the weights.

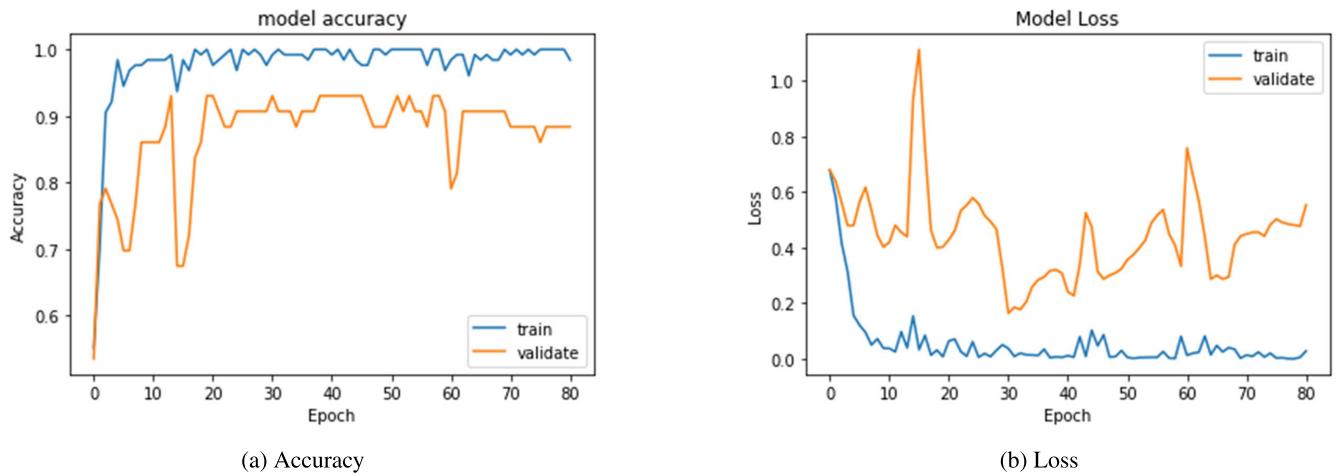
For EfficientNetB3, the accuracy and loss values graph in the training steps is given in Figure 15.

As a result of the training using this method, a confusion matrix was created for the EfficientNetB3 model and is given in Figure 16.

**FIGURE 16.** Confusion matrix for Method 2.**c: RESULTS FOR METHOD 3**

This part of the study aims to find the pixels with leaf features in the images in the EruCauliflowerDB dataset. For this purpose, background information was extracted by applying some image processing methods to the images obtained, and diseased leaf pixel locations were obtained. In this method, the diseased leaf images obtained by distinguishing the background were given to VGG16, EfficientNetB3, MobileNetV3 Large, and ResNet50 models and put into training and testing stages. This section performed training without applying image augmentation and variation methods.

After the image segmentation method, the training and test steps were given to the images. Using the training parameters given in Table 6, training was performed on the four CNN

**FIGURE 17.** EfficientNetB3 training steps with segmented dataset.

models used in the other methods. As a result of this training, the results shown in Table 8 were achieved.

TABLE 8. Results obtained with Method 3.

Model	Early Stop	Training Accuracy	Test Accuracy
VGG16	66	96.06	95.34
EfficientNetB3	81	98.43	97.67
MobileNetV3 Large	59	100	86.04
ResNet50	70	98.43	90.69

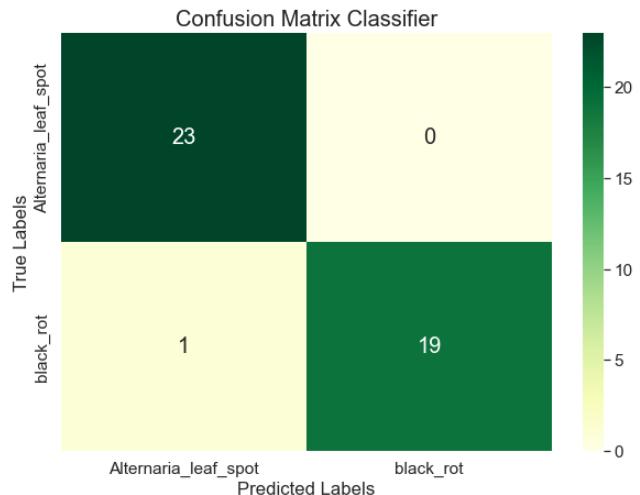
According to the results in the table, as a result of the training steps with the images generated using the segmentation method, the EfficientNetB3 model reached the highest accuracy rate with 97.67% in 81 steps and was stopped because the early stopping condition was met. Figure 17 shows the accuracy and loss values of the EfficientNetB3 model in the training step.

The confusion matrix shown in Figure 18 is achieved due to the tests performed with the test images used in the other methods. The test data was also segmented in this method, and the background data was converted to black pixels.

d: RESULTS FOR PROPOSED INTEGRATED METHOD

This section of the study proposes a method that combines image augmentation and image segmentation, previously tested in different contexts. In this part of the study, in the first stage, background information was extracted from the images in the dataset using the image segmentation method described in Section II. The segmentation images were saved with diseased leaf information in the training, validation, and test folders against a black background. This method utilized the same training, validation, and test images as the other three methods.

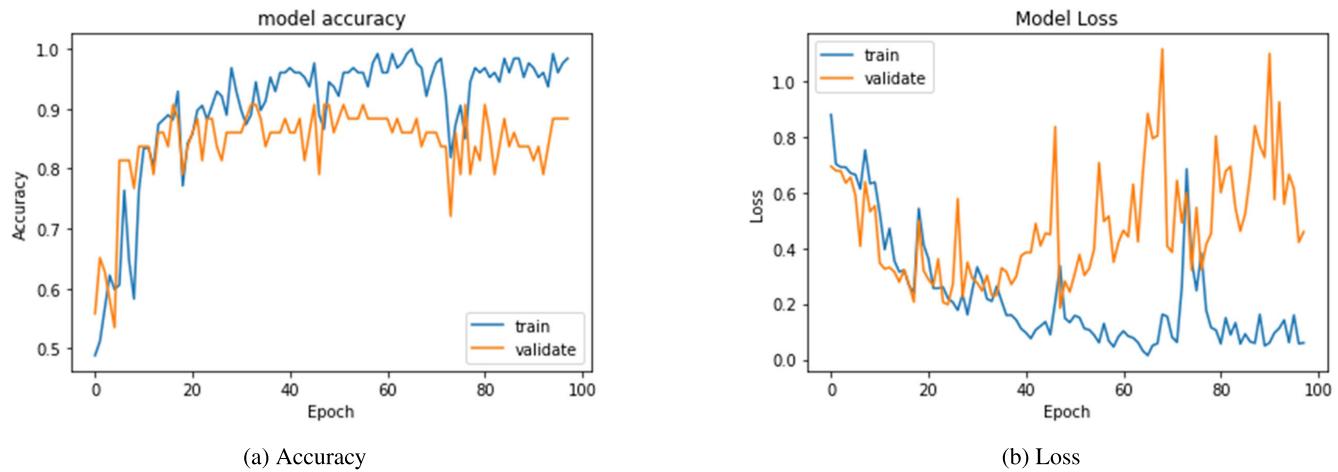
When using the proposed method, data augmentation methods were used only for the folder containing the training images obtained due to the BorB segmentation process. Before starting the training steps, the methods and values to change the image morphology given in Table 3 were used on

**FIGURE 18.** EfficientNetB3 confusion matrix with Method 3.**TABLE 9.** Results obtained with the Proposed Integrated Method.

Model	Early Stop	Training Accuracy	Test Accuracy
VGG16	98	98.43	100
EfficientNetB3	150	98.43	95.34
MobileNetV3 Large	70	99.21	88.37
ResNet50	58	100	97.67

training data. The results obtained after the training were the accuracy values given in Table 9 for the four deep learning models. In training, the VGG16 model achieved the highest accuracy rate of 100%. In the proposed method, where BorB image segmentation and data augmentation processes are applied together, the deep learning model's accuracy and loss graphs that give the highest accuracy in the training stages are shown in Figure 19.

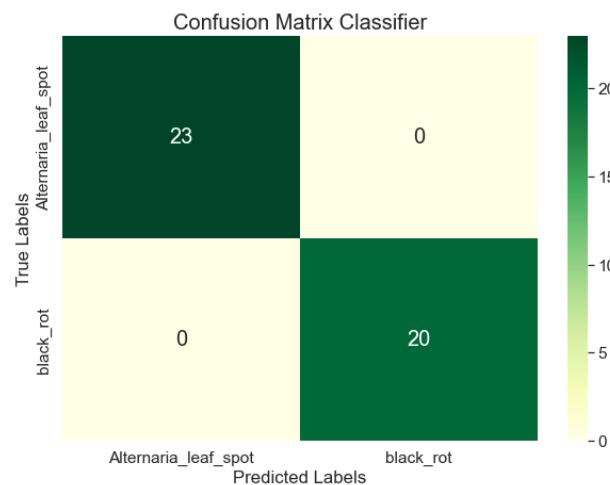
For the VGG16 model, which reached the highest accuracy value as a result of the training, it was observed that there was no incorrect prediction between the classes in the test data, and it is shown on the confusion matrix in Figure 20.

**FIGURE 19.** VGG16 training steps for the proposed integrated method.

Inference times of deep learning architectures used with the proposed Integrated Method are given in Table 10. The inference times and input image sizes of the trained models for 1 image are also specified in the table.

TABLE 10. Inference time of deep learning architectures.

Model	Image Size	Batch Size	Inference Time (ms)
VGG16	224x224	1	9
EfficientNetB3	224x224	1	18
MobileNetV3 Large	224x224	1	10
ResNet50	224x224	1	11

**FIGURE 20.** Confusion matrix for the proposed integrated method.

e: RESULTS OF DIFFERENT SEGMENTATION TECHNIQUES IN THE PROPOSED INTEGRATED METHOD

The Otsu threshold and K-means cluster methods were used to analyze the BorB image segmentation performance in the proposed integrated classification method. The classification success rates obtained by using three different methods for image segmentation are presented in Table 11.

TABLE 11. Results obtained from the proposed integrated method with different segmentation techniques.

Segmentation Methods	CNN	Test Accuracy
Otsu threshold	EfficientNetB3	97.67
	MobileNetV3 Large	90.69
	ResNet50	90.69
	VGG16	97.67
K-means clustering	EfficientNetB3	95.34
	MobileNetV3 Large	90.69
	ResNet50	86.04
	VGG16	97.67
Proposed BorB	EfficientNetB3	95.34
	MobileNetV3 Large	88.37
	ResNet50	97.67
	VGG16	100

According to Table 11, the best result was achieved in the integrated classification method created with the BorB method proposed for image segmentation.

B. COMPARISON RESULTS

Four different methods (The Proposed Integrated Method and its three variants) were analyzed for classifying diseases in cauliflower leaves with deep learning methods, which is the main subject of this research; for each of these methods, training, and testing steps were carried out with four different deep learning architectures. This section gives the table of the highest accuracy values obtained for each method and some evaluation metrics.

Table 12 shows some evaluation metrics for the methods used in the experiments. These values were obtained from the test data. Table 13 shows the deep learning architectures that achieve the highest accuracy for each method.

C. SIMULATION RESULTS FOR THE MANGOLEAFBD DATASET

To prove the effectiveness of the proposed method, experiments were conducted using diseased leaf images of a different plant. For this purpose, the publicly available

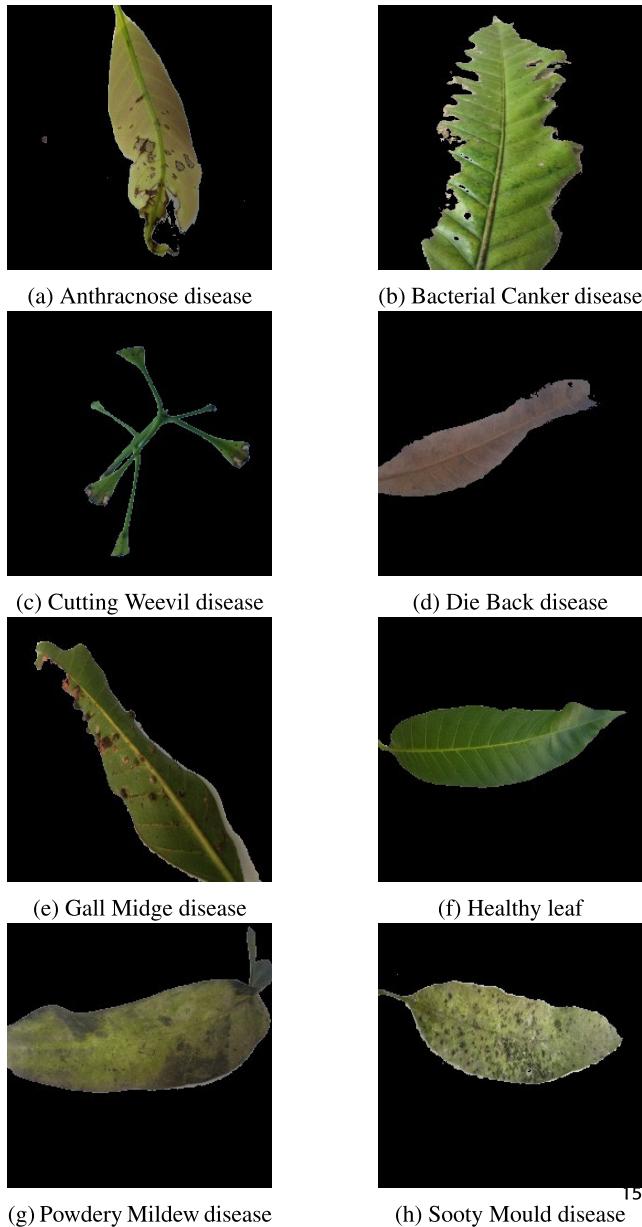


FIGURE 21. Segmented MangoLeafBD images.

MangoLeafBD dataset was tested using the proposed method. The proposed BorB image segmentation method was applied to the leaf images in the dataset. The segmented leaf images are presented in Figure 21.

As shown in Figure 21, the leaf images obtained were prepared for the training step using the image augmentation techniques described in Section II. To prepare for the training steps, the parameters given in Table 6 were set, and the training steps were completed with the VGG16 model. As a result of the training, the testing steps were performed on the MangoLeafDB data set, and the results were compared with other studies in the literature in Table 14. From the table, the accuracy rate for the run of the proposed method is 100%, which is better than the existing methods

in [16] (which have reported performances of 99.23%). When other performance metrics were examined, a 100% rate was achieved in all metrics with the proposed method. The confusion matrix, which is formed according to the performance of the proposed method, is shown in Figure 22.

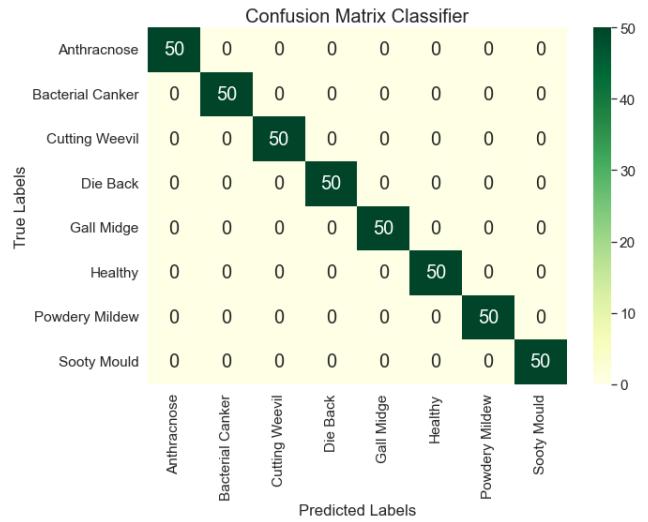


FIGURE 22. Confusion matrix for the MangoLeafBD.

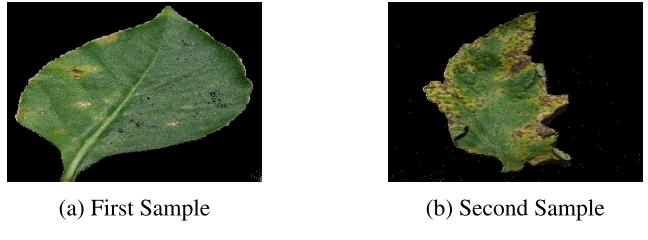


FIGURE 23. Examples of PlantVillage images with BorB segmentation.

D. SIMULATION RESULTS FOR THE PLANTVILLAGE DATASET

The proposed Integrated Classifier Method for classifying leaf diseases was tested with the open-to-use available PlantVillage leaf disease dataset. To implement the proposed method, 15 classes were selected from the PlantVillage dataset.

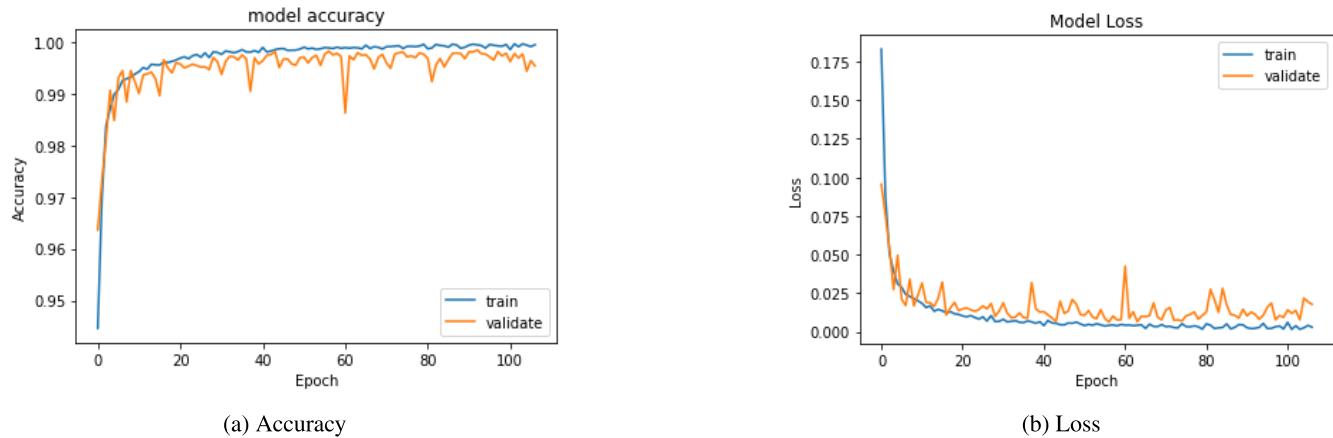
The steps shown in Figure 12 were followed, and the results given in Table 15 were obtained. Precision, Recall, and F1-Score values are given in the table. These parameters are important for the evaluation of the multi-class PlantVillage dataset. Calculating these evaluation metrics is important for accurate analysis in data sets with unbalanced data numbers between classes. Output examples of the BorB segmentation method used to distinguish leaf regions from background information in the preprocessing step are given in Figure 23. The proposed integrated classifier method was used to train the PlantVillage dataset. In this method, VGG16 architecture was used to perform deep learning stages.

TABLE 12. Evaluation metrics for four different methods.

Method	CNN	Class Names	Precision	Recall	F1-Score
Real Data	ResNet50	Alternaria Leaf Spot	92	100	95.83
		Black Rot	100	90	94.74
Data augmentation	EfficientNetB3	Alternaria Leaf Spot	95.65	95.65	95.65
		Black Rot	95	95	95
Image segmentation	EfficientNetB3	Alternaria Leaf Spot	95.83	100	97.87
		Black Rot	100	95	97.44
Proposed Method	VGG16	Alternaria Leaf Spot	100	100	100
		Black Rot	100	100	100

TABLE 13. The highest accuracy values were obtained for the methods.

Method	Model	Early Stop	Training Accuracy	Test Accuracy
Method 1	ResNet50	128	100	95.34
Method 2	EfficientNetB3	141	98.43	95.34
Method 3	EfficientNetB3	81	98.43	97.67
Proposed Method	VGG16	98	98.43	100

**FIGURE 24.** Training steps for PlantVillage dataset.**TABLE 14.** Comparison of the proposed method for the MangoLeafBD dataset.

Competing method	Precision	Recall	F1-Score	Accuracy
(Salamai, 2023) [16]	99.01	99.03	99.02	99.23
(Hou et al., 2021) [34]	97.92	98.17	98.04	98.12
(Hasani et al., 2021) [35]	95.03	98.12	96.55	97.62
(Thai et al., 2023) [36]	96.20	95.59	95.89	97.20
(Guo et al., 2022) [37]	96.11	96.41	96.26	96.64
Proposed	100	100	100	100

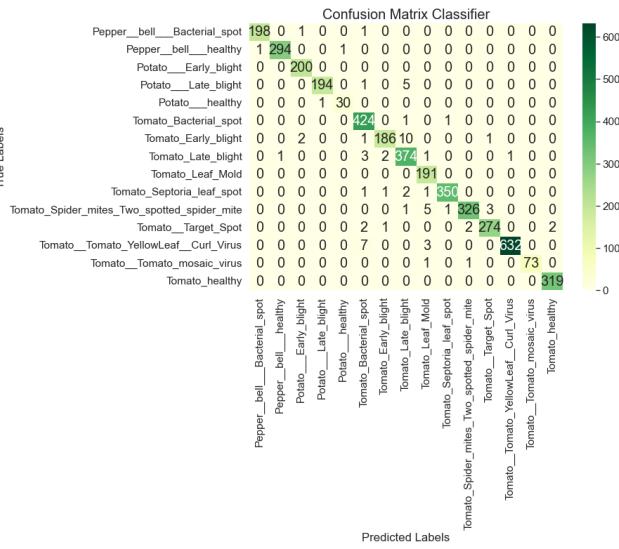
The results given in Table 15 are important to examine how balanced the training performed for the multi-class data set is. Additionally, to conclude that there is no overfitting, the training and validation graphs are given in Figure 24. As a result of the training steps performed, the test accuracy rate for 15 classes was 99.78%. In order to analyze the prediction errors between classes, the confusion matrix is given in Figure 25. When the confusion matrix is examined,

TABLE 15. Results obtained by applying the proposed method using the PlantVillage dataset.

Class name	Precision	Recall	F1-Score
Pepper bell bacterial spot	99.50	99.00	99.25
Pepper bell healthy	99.66	99.32	99.49
Potato Early blight	98.52	100	99.26
Potato Late blight	99.49	97.00	98.23
Potato healthy	96.77	96.77	96.77
Tomato Bacterial spot	96.36	99.53	97.92
Tomato Early blight	97.89	93.00	95.38
Tomato Late blight	95.17	97.91	96.52
Tomato Leaf Mold	94.55	100	97.20
Tomato Septoria leaf spot	99.43	98.59	99.01
Tomato Spider mites	99.09	97.02	98.05
Tomato Target Spot	98.56	97.51	98.03
Tomato Yellow Leaf Curl	99.84	98.44	99.14
Tomato mosaic virus	100	97.33	98.65
Tomato healthy	99.38	100	99.69

it is seen that balanced predictions are made even though the number of images in the classes is different.

True labels

**FIGURE 25.** Confusion matrix for the PlantVillage dataset.

IV. CONCLUSION

Artificial intelligence studies in agriculture have become a trend in recent years. In particular, detecting and intervening diseases in leaves increases the productivity and quality of the crop. This work employed deep learning, data augmentation, and image segmentation techniques for detecting and categorizing illnesses in leaves. The proposed integrated method includes the proposed BorB image segmentation technique, image augmentation steps, and training with a deep learning model. Three different datasets are used to perform the experimental studies. The EruCauliflowerDB dataset was built exclusively for this research to identify and classify two distinct illnesses in cauliflower leaves. The dataset was split into train, validation, and test folders to illustrate the significance of each stage in the suggested methodology. Three methods (Method 1, Method 2, and Method 3), sub-variants of the method, were used to compare the proposed method. According to the experimental results, the proposed method obtained 100% test accuracy when utilizing the VGG16 deep learning architecture for training. Experimental studies have also been carried out on the MangoLeafBD dataset to examine the robustness of the proposed method. The proposed method achieved 100% accuracy on the test data. The proposed Integrated Classifier method was applied by selecting 15 classes from the PlantVillage dataset. As a result of the experiments, PlantVillage plant leaf images were classified with 99.78% accuracy. Experimental results have shown that a successful method has been developed by achieving 100% success rates on both data sets. On the other side, there are the limitations of the study. These are: dataset size and diversity, environmental factors such as variations in lighting and humidity, and computational resources.

Based on the context of this study, future research should focus on expanding the 'EruCauliflowerDB' dataset with diverse images from various geographical locations and environmental conditions to enhance model robustness.

Cross-dataset validation on publicly available plant disease datasets is crucial for assessing generalizability. Developing a real-time detection system deployable on mobile or embedded devices would enable practical applications. Integrating hyperspectral imaging or thermal cameras can provide richer plant health information. Exploring advanced segmentation techniques like instance or semantic segmentation can further improve diseased region delineation. Long-term studies across multiple growing seasons and the inclusion of additional disease classifications will also contribute to the advancement of robust agricultural AI systems.

ACKNOWLEDGMENT

The authors would like to thank the Proofreading and Editing Office of the Dean for Research at Erciyes University for the copyediting and proofreading service for this manuscript.

REFERENCES

- [1] E. Hudik, Y. Yoshioka, S. Domenichini, M. Bourge, L. Soubigou-Taconnat, C. Mazubert, D. Yi, S. Bujaldon, H. Hayashi, L. De Veylder, C. Bergounioux, M. Benhamed, and C. Raynaud, "Chloroplast dysfunction causes multiple defects in cell cycle progression in the arabidopsis crumpled leaf mutant," *Plant Physiol.*, vol. 166, no. 1, pp. 152–167, Aug. 2014.
- [2] D. S. Joseph, P. M. Pawar, and K. Chakradeo, "Real-time plant disease dataset development and detection of plant disease using deep learning," *IEEE Access*, vol. 12, pp. 16310–16333, 2024.
- [3] V. E. González-Rodríguez, I. Izquierdo-Bueno, J. M. Cantoral, M. Carbú, and C. Garrido, "Artificial intelligence: A promising tool for application in phytopathology," *Horticulturae*, vol. 10, no. 3, p. 197, Feb. 2024.
- [4] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: A review," *Plant Methods*, vol. 17, no. 1, pp. 1–18, Dec. 2021.
- [5] K. Dokic, L. Blaskovic, and D. Mandusic, "From machine learning to deep learning in agriculture—The quantitative review of trends," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 614, no. 1, Dec. 2020, Art. no. 012138.
- [6] K. Zhang, Q. Wu, A. Liu, and X. Meng, "Can deep learning identify tomato leaf disease?" *Adv. Multimedia*, vol. 2018, pp. 1–10, Sep. 2018.
- [7] M. Bhagat, D. Kumar, and S. Kumar, "Bell pepper leaf disease classification with LBP and VGG-16 based fused features and RF classifier," *Int. J. Inf. Technol.*, vol. 15, no. 1, pp. 465–475, Jan. 2023.
- [8] H. Alshammari, K. Gasmi, I. Ben Ltaifa, M. Krichen, L. Ben Ammar, and M. A. Mahmood, "Olive disease classification based on vision transformer and CNN models," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Jul. 2022.
- [9] L. C. Ngugi, M. Abdelwahab, and M. Abo-Zahhad, "A new approach to learning and recognizing leaf diseases from individual lesions using convolutional neural networks," *Inf. Process. Agricult.*, vol. 10, no. 1, pp. 11–27, Mar. 2023.
- [10] M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine," in *Proc. IEEE 30th Can. Conf. Electr. Comput. Eng. (CCECE)*, Apr. 2017, pp. 1–4.
- [11] S. V. Moravvej, R. Alizadehsani, S. Khanam, Z. Sobhaninia, A. Shoeibi, F. Khozeimeh, Z. A. Sani, R.-S. Tan, A. Khosravi, S. Nahavandi, N. A. Kadri, M. M. Azizan, N. Arunkumar, and U. R. Acharya, "RLMDPA: A reinforcement learning-based myocarditis diagnosis combined with a population-based algorithm for pretraining weights," *Contrast Media Mol. Imag.*, vol. 2022, no. 1, Jan. 2022, Art. no. 8733632.
- [12] E. Amri, Y. Gulzar, A. Yeafi, S. Jendoubi, F. Dhawi, and M. S. Mir, "Advancing automatic plant classification system in Saudi Arabia: Introducing a novel dataset and ensemble deep learning approach," *Model. Earth Syst. Environ.*, vol. 10, no. 2, pp. 2693–2709, Apr. 2024.
- [13] Y. Gulzar, "Enhancing soybean classification with modified inception model: A transfer learning approach," *Emirates J. Food Agricult.*, vol. 36, pp. 1–9, Apr. 2024.
- [14] Y. Gulzar, Z. Ünal, H. Aktas, and M. S. Mir, "Harnessing the power of transfer learning in sunflower disease detection: A comparative study," *Agriculture*, vol. 13, no. 8, p. 1479, Jul. 2023.

- [15] L. Lei, Q. Yang, L. Yang, T. Shen, R. Wang, and C. Fu, "Deep learning implementation of image segmentation in agricultural applications: A comprehensive review," *Artif. Intell. Rev.*, vol. 57, no. 6, p. 149, May 2024.
- [16] A. A. Salamai, "Enhancing mango disease diagnosis through eco-informatics: A deep learning approach," *Ecological Informat.*, vol. 77, Nov. 2023, Art. no. 102216.
- [17] S. I. Ahmed, M. Ibrahim, M. Nadim, M. M. Rahman, M. M. Shejunti, T. Jabid, and M. S. Ali, "MangoLeafBD: A comprehensive image dataset to classify diseased and healthy mango leaves," *Data Brief*, vol. 47, Apr. 2023, Art. no. 108941.
- [18] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers Plant Sci.*, vol. 7, 2016, Art. no. 215232.
- [19] R. M. Mirjat, S. A. Mahar, M. I. M. Siddiqui, J. A. M. Mahar, and A. M. Magsi, "A framework for guava wilt disease segmentation using K-means clustering and neural network techniques," *VAWKUM Trans. Comput. Sci.*, vol. 12, no. 1, pp. 76–93, May 2024.
- [20] K. Sharma, J. Jayanthi, A. Kumar, S. Jain, G. Jain, and H. Vishwakarma, "Computer vision based smart leaf disease detection system," in *Proc. IEEE Int. Students' Conf. Electr., Electron. Comput. Sci. (SCEECS)*, Feb. 2024, pp. 1–9.
- [21] M. O. Pallavi and A. Raj, "Tomato disease prediction model using machine learning algorithms and image processing techniques," in *Proc. 6th Int. Conf. Comput. Intell. Netw. (CINE)*, Dec. 2024, pp. 1–6.
- [22] A. Elhanashi, S. Saponara, and Q. Zheng, "Classification and localization of multi-type abnormalities on chest X-rays images," *IEEE Access*, vol. 11, pp. 83264–83277, 2023.
- [23] M. Bhagat, D. Kumar, and S. Kumar, "Optimized transfer learning approach for leaf disease classification in smart agriculture," *Multimedia Tools Appl.*, vol. 83, no. 20, pp. 58103–58123, Dec. 2023.
- [24] S. Tammina, "Transfer learning using VGG-16 with deep convolutional neural network for classifying images," *Int. J. Sci. Res. Publications*, vol. 9, no. 10, p. p9420, Oct. 2019.
- [25] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [27] Z. Jiang, Z. Dong, W. Jiang, and Y. Yang, "Recognition of Rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning," *Comput. Electron. Agricult.*, vol. 186, Jul. 2021, Art. no. 106184.
- [28] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2019, pp. 6105–6114.
- [29] R. Kaur, R. Kumar, and M. Gupta, "Deep neural network for food image classification and nutrient identification: A systematic review," *Rev. Endocrine Metabolic Disorders*, vol. 24, no. 4, pp. 633–653, Aug. 2023.
- [30] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of remote sensing images using EfficientNet-B3 CNN model with attention," *IEEE Access*, vol. 9, pp. 14078–14094, 2021.
- [31] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324.
- [32] W. Glegoła, A. Karpus, and A. Przybyłek, "MobileNet family tailored for Raspberry Pi," *Procedia Comput. Sci.*, vol. 192, pp. 2249–2258, 2021.
- [33] D. A. Pramudhita, F. Azzahra, I. K. Arfat, R. Magdalena, and S. Saidah, "Strawberry plant diseases classification using CNN based on MobileNetV3-large and EfficientNet-B0 architecture," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 3, pp. 522–534, Jul. 2023.
- [34] Q. Hou, Z. Jiang, L. Yuan, M.-M. Cheng, S. Yan, and J. Feng, "Vision permutator: A permutable MLP-like architecture for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 1328–1334, Jan. 2023.
- [35] A. Hassani, S. Walton, N. Shah, A. Abduweili, J. Li, and H. Shi, "Escaping the big data paradigm with compact transformers," 2021, *arXiv:2104.05704*.
- [36] H.-T. Thai, K.-H. Le, and N. L.-T. Nguyen, "FormerLeaf: An efficient vision transformer for cassava leaf disease detection," *Comput. Electron. Agricult.*, vol. 204, Jan. 2023, Art. no. 107518.
- [37] Y. Guo, Y. Lan, and X. Chen, "CST: Convolutional Swin transformer for detecting the degree and types of plant diseases," *Comput. Electron. Agricult.*, vol. 202, Nov. 2022, Art. no. 107407.



T. OZCAN received the B.S. degree in computer engineering from İstanbul Kültür University, İstanbul, Türkiye, in 2010, and the M.Sc. and Ph.D. degrees from the Department of Computer Engineering, Erciyes University, in 2016 and 2020, respectively. He joined the Department of Computer Engineering, Erciyes University, as a Research Assistant, in 2013. His research interests include intelligent optimization algorithms, image processing, machine learning, deep learning, and Kalman filtering. His application areas focus on human action recognition, human-computer interaction, and artificial intelligence in healthcare and agriculture.



E. POLAT received the B.S. degree in computer engineering from Sivas Cumhuriyet University, in 2020. He is currently pursuing the master's degree with the Computer Engineering Department, Erciyes University. He started his first work experience as a Research and Development Engineer with Gitek Vision. Since 2024, he has been continuing his career as a Research Assistant in the field of computer science with Niğde Ömer Halisdemir University. His research interests include artificial intelligence in agriculture, image processing, deep learning, and robotic systems.