

Multi-Disease Classification and Severity Estimation of Cotton and Soybean Leaves Using Convolutional Neural Network

Vaishali G. Bhujade^{1*}, Dr V. K. Sambhe²

Submitted: 04/11/2023

Revised: 24/12/2023

Accepted: 03/01/2024

Abstract: Multi-disease classification and severity estimation of cotton and soybean leaves is a crucial task in the field of agriculture. The early detection and management of plant diseases are vital for ensuring crop yield and food security. This study presents a novel approach for the multi-disease classification and severity estimation of cotton and soybean leaves using the hybridization of a Convolutional Neural Network (CNN) with Multi-Resolution Feature Optimization (MRFO). The suggested model takes the benefit of both CNN and MRFO to enhance the classification and severity estimation performance. The dataset used for training and testing consists of images of cotton and soybean leaves affected by multiple diseases. The experimental results demonstrate that the suggested model provides improved classification accuracy and severity estimation compared to the state-of-the-art approaches. The model achieved an overall accuracy of 96.07% for cotton leaves and 95.51% for soybean leaves. Moreover, the proposed model accurately estimated the severity of the diseases in the soybean and cotton leaves to be 99.1% and 91.21% respectively, which is crucial for effective disease management.

Keywords: Convolutional Neural Network, Multi-Resolution Feature Optimization, Classification, Disease Detection, severity estimation

1. Introduction

Agriculture is an important sector because it provides food for people all over the world. The prosperity of a nation's farmers depends on the success of their crops. A major danger to agricultural output and food security, however, comes from crop diseases [Kaur et al. (2019)]. Losses in crop yield, decreased crop quality, and higher production costs are all possible outcomes of crop disease. As a result, proper disease management and control depend on prompt and precise crop disease detection. Therefore, several researchers have turned to image-processing methods to aid in the diagnosis of plant diseases [Bagde et al. (2015), Rastogi et al. (2015)]. The most accurate way for identifying agricultural illnesses has been a visual inspection by professionals, but this process is time-inefficient and prone to error. Additionally, several agricultural diseases have identical symptoms at different phases of plant development, making it difficult to distinguish between them [Prasad et al. (2016), Khirade & Patil (2015)]. Because of this, it is necessary to have procedures that are both accurate and efficient for identifying agricultural illnesses.

There are a variety of diseases that can affect plants; however, the disease that a plant has is determined by its etiology. Additionally, because of this, it is more difficult for computer vision techniques to appropriately recognize them [Barbedo (2016), Prashar et al. (2015)]. Molecular, serological, and DNA analysis are three of the most widely used classic methods. The detection procedure is

automated in novel ways by using volatile organic molecules and imaging and spectroscopic methods [Applanaidu & Kumaravelan (2021)]. These cutting-edge methods are more efficient and require less oversight. The following figure 1 illustrates leaves of cotton and soybean leaves infected with plant diseases.



Fig 1. Disease identification in cotton
[<http://cotton.tamu.edu/Photos/diseasephotos/images/10bacterialblight.jpg>] and soybean
[https://fieldcropnews.com/2016/07/soybean_disease_resources/] leaves

Investment in creating automated systems for disease identification and severity estimation in crops has increased alongside the development of computer vision and machine learning techniques. The detection and categorization of crop diseases using machine learning (ML) methods have shown encouraging results in recent years [Annabel et al. (2019)]. Machine learning algorithms can shift through massive datasets in search of patterns and characteristics that would be invisible to the human eye. The surrounding environment, genetic factors, and management techniques are only a few of the many potential influences on crop health that must be evaluated to detect and classify diseases in multi-variant crops. ML-based algorithms can produce a more precise result by combining data from several different sources and taking some different aspects into account. This study examines deep learning (DL) methods for multi-variant crop disease detection, classification, and severity analysis. DL

*1*Research Scholar, Department of Computer Engineering, Veermata Jijabai Technological Institute, Mumbai

2 Supervisor, Department of Computer Engineering & Information Technology, Veermata Jijabai Technological Institute, Mumbai

* Corresponding Author Email:
vaishali.bhujadesfdc@gmail.com

approaches for identifying agricultural illnesses, how they might enhance agriculture, and its possibility for widespread application will be discussed. This paper proposes a CNN-based multi-disease classification and severity assessment approach for cotton and soybean leaf pictures. The recommended approach may classify leaf diseases and evaluate severity. This goal is to establish a reliable, accurate, and automated crop disease detection and severity assessment system. The suggested method could evaluate large datasets, handle complex characteristics, and produce reproducible and consistent results better than the present methodology.

1.1. Analysis of DL models for plant disease classification

Deep learning (DL) models, especially CNNs, can classify disease of plants including cotton and soybean diseases. These models accurately recognize healthy and sick leaves from leaf pictures [Fan et al. (2022)] [Abade et al. (2021)]. DL models may follow disease progression by processing field sensor time-series data on temperature, humidity, and soil moisture [Saleem et al. (2019)]. Pre-trained DL models from ImageNet may be adapted to cotton or soybean disease datasets to improve model efficiency [Wu et al. (2019)]. This method uses historical data and environmental factors to anticipate and prevent disease outbreaks. CNNs, a fundamental DL architecture, are ideal for automated feature extraction from input images [Amara et al. (2017)]. Their capacity to handle data with minimum variance and low pre-processing helps them identify and diagnose diseases.

The study presents a completely automated technique for diagnosing cotton and soybean plant diseases. This study would target plant disease detection in cotton and soybean, two vital commodities for global food security. Farmers might make educated management decisions and reduce crop losses with the suggested system's accurate plant disease detection and severity estimation. This technology uses image processing and deep learning algorithms to reliably identify and categorize plant illnesses, making it new. The study shows the potential of image processing and deep learning to construct automated and efficient plant disease detection systems, which might benefit plant disease management.

The following paper is organized as follows: Section 2 discusses crop disease diagnostic and severity assessment research. Section 3 introduces crop disease diagnosis study. Section 4 covered crop disease detection and severity analysis problems. Section 5 lists research goals. Research gaps are listed in Section 6. The multi-variant crop disease detection and classification dataset is explained in Section 7. Section 8 describes the study approach, and Section 9 describes the experiments and compares them to existing procedures. Section 10 summarizes the study's findings and discusses crop disease diagnostics' future impact on agriculture.

2. Literature of review

Several different approaches to the diagnosis of plant diseases are

Authors	Techniques	Outcomes
Zhang et al., (2021)	MF3 R-CNN	The suggested MF3 R-CNN achieved mean average precision of 83.34% in disease identification.
Tetila et al., (2020)	Hybrid fine-tuning Transfer learning	DLThe experimental results show that the fine-tuned DL architectures achieved better classification with an accuracy of up to 93.82%.
Jadhav et al., (2019)	GLCM, SVM, KNN	The proposed classifiers system classified samples infected with blight, brown spot
Ghatol et al., (2019)	SVM, KNN	Accuracy levels of 93.45% and 93.58% were achieved

discussed, and a comprehensive review of both conventional and cutting-edge methods is provided in the relevant literature.

Zhang K. et al., (2021) explained that the quality of the soybean crop and the agricultural economy depend on early and precise diagnosis of leaf diseases in the soybean field. This study develops a multi-feature fusion Faster Region-CNN (MF3 R-CNN) to deal with the aforementioned difficult issue with a mean average precision of 83.34% in a real-world test dataset. Tesla et al., (2020) utilized five different DL architectures to classify images of soybean. The experimental results show that the fine-tuned DL architectures may achieve greater classification rates than alternative methods, with an accuracy of up to 93.82%. Jadhav et al., (2019) identified soybean leaf infections and grouped them using color images of infected leaves. Samples infected with blight, brown spots, and frog eye leaf spots may be correctly classified at an 87.3% rate, whereas healthy samples are classified at an 83.6% rate using the proposed classifiers system. Ghatol et al., (2019) observed that the main problem is ensuring the plants are healthy and identifying any diseases that may be affecting them. This study used variety of ML algorithms. Accuracy levels of 93.45% and 93.58% were achieved. Kumari et al., (2019) suggested a system to detect the leaf spot with the use of image processing methods. The detection rates of bacterial leaf spots and targeted spots in cotton leaf diseases are found to be 90% and 80%, respectively. Sarangdhar et al., (2017) present SVM-based regression approach for cotton leaf diseases detection and categorization. The proposed approach achieves an overall accuracy of 83.26% in its classifications. Parikh et al., (2016) observed that cotton leaf reflects most disease signs. The novelty of the idea is in the processing of images obtained in the field by an inexperienced person using a standard or mobile phone camera under uncontrolled conditions. The proposed work utilizes a cascade of KNN classifiers and various training sets to successfully detect this disease along with its stage from unconstrained images. Gharge et al., (2016) offers an algorithm for identifying the presence and severity of soybean diseases. The k-means segmentation algorithm is used to distinguish between healthy and diseased clusters. To categorize Various diseases of Soybeans, a neural network is employed. After processing 30 images, a success rate of 93.3% was achieved. Rothe et al., (2015) introduces a pattern recognition approach for identifying and classifying Alternaria, mesothelium, and bacterial blight leaf diseases in cotton. This study utilized an active contour model for picture segmentation and extract Hu's moments as features. When applied to images of diseased cotton leaves, back propagation neural networks had an average accuracy of 85.52%. Ranjan et al., (2015) suggested a method that makes use of a variety of image processing techniques and ANN to detect cotton leaf diseases early and reliably. The experiments revealed that using an ANN with a feature set improves classification performance by about 80%. The following table 1 presents the comparative analysis of reviewed literature.

Kumari et al., (2019)	Neural network	The detection rates for bacterial leaf spots and targeted spots in cotton leaves were 90% and 80%, using the NN classifier.
Sarangdhar et al., (2017)	SVM	The proposed SVM-based regression approach achieves an overall accuracy of 83.26% in its classifications.
Parikh et al., (2016)	KNN	This study successfully detected disease along with its stage from unconstrained images.
Gharge et al., (2016)	K-means segmentation Neural Network	Findings revealed that the proposed neural network achieved a success rate of 93.3% after processing 30 images.
Rothe et al., (2015)	Backpropagation neural networks	Results revealed that back propagation neural networks had an average accuracy of 85.52%
Ranjan et al., (2015)	ANN	The experiments revealed that using an ANN with a feature set improves classification performance by about 80%,

Table 1. depicts the Summary of the Reviewed Literature

3. Background study

Infectious diseases of plants are a major cause of economic losses in agriculture because they reduce crop yields and reduce food quality. The use of deep learning, and in particular deep CNN with transfer learning, is becoming increasingly popular in the agriculture industry as a means of dealing with this problem. Recent research has improved upon previous approaches for identifying plant leaf diseases by using models such as VGGNet and Inception. The approach was shown to be reliable and successful in identifying plant diseases, with an average accuracy of 92.00% on a difficult dataset. [Chen et al. (2020)].

4. Problem formulation

Plant diseases are the main cause of crop yield and quality losses worldwide. Plant diseases must be found early and accurately diagnosed to minimize crop losses and ensure food security. Cotton and soybean production need an accurate, efficient approach for detecting, categorizing, and assessing plant disease severity. Current illness diagnosis and severity assessment methods are time-consuming and manual. To solve this difficulty, image processing has been suggested. The suggested method would employ a CNN hybridized with the MRFO algorithm to improve cotton and soybean disease detection, classification, and severity assessment.

5. Research objectives

- To collect the several damaged cotton and soybean leaves for accurate disease detection.
- To preprocess the images with a Gaussian filter to improve segmentation and feature extraction.
- To create an excellent color and texture image segmentation system utilizing PWFCM clustering.
- To enhance illness categorization by extracting key characteristics from segmented regions utilizing GLCM, LDP, and Global Filter.
- To improve illness prediction and severity evaluation using a hybrid CNN-MRFO technique and construct a visual symptom severity estimate system.

6. Research gaps

- Insufficient cotton and soybean leaf disease databases reduce system accuracy.
- Pre-processing challenges related to noise and blur reduction,

segmentation, and feature extraction.

- Limitations in image segmentation approaches for reliable area separation limit diseases categorization.
- Limited feature extraction strategies can affect disease prognosis and severity estimate for plant diseases classified by visual symptoms.

7. Dataset description

The dataset utilized in this study consists of real field images of Soybean and Cotton obtained from the Indian Council of Agriculture Research (ICAR) - Central Institute for Cotton Research (CICR), located in Nagpur. A phone camera recorded these photographs under varied lighting and environmental situations, revealing plant features. The collection comprises photos from top and bottom leaf sides, healthy and sick, and was classified by ICAR-CICR scientists. The soybean dataset for disease prediction includes 990 images of 9 soybean leaf illnesses and 110 healthy soybean leaf images. Disease types and healthy controls were studied using 110 images each. Some diseases of plants include yellow mosaic disease, frogeye leaf spot, healthy leaves, sudden death syndrome, target leaf spot, bacterial pustule, rust, and a combination of the three. This uniformly distributed sample of images allows researchers to study these diseases on soybean and cotton plants in their natural habitats. Figure 2 shows soybean leaves. The dataset contains soybean leaf samples, and the purpose is to determine whether disease is impacting them based on color, texture, etc.

The cotton dataset for disease prediction consists of a total of 990 images, which are classified into 9 diseases that occurred in cotton leaves including 110 images of healthy cotton leaves. The number of images in each category is recorded in the cotton plant disease dataset. Among the total, 39 portray healthy cotton leaves, 825 exhibit diseased cotton plants, and the remaining 40 show bacterial blight, 38 show powdery mildew, 40 show armyworm infestations, 41 show target spot, and 39 show aphids. In addition to the 421 photographs of healthy cotton plants, there are also 288 images of cotton leaves that have been affected by disease. This extensive dataset represents a wide range of cotton plant conditions, making it useful for testing new methods of disease detection and classification. Figure 3 given below illustrates the image of cotton leaves.



Fig 1. Soybean leaf image



Fig 2. Cotton leaf image

8. Research Methodology

9.1. Technique used

The various techniques used in the proposed methodology are discussed below:

8.1.1. Piece-wise Fuzzy C-means clustering

To segment an image is to divide it into smaller parts that each represent a different object or element of the background. Applying clustering methods, such as the PWFCM clustering algorithm, to the task of image segmentation is one possible strategy. In the realm of unsupervised machine learning, PWFCM is a spinoff of the traditional Fuzzy C-means (FCM) clustering algorithm used to solve problems of clustering and classification. PFCM clustering is a method for picture segmentation in which each pixel is assigned a fuzzy membership to one or more clusters [Sathish et al. (2020)]. The image is first partitioned into several sub-regions in PFCM, and then each of those is subjected to fuzzy clustering analysis. The mathematical formula for PFCM is as follows:

Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of n pixels in an image, and let k be the number of clusters. Divide the image into sub-regions $R_1, R_2, \dots,$ and R_m . For each sub-region R_i , initialize the cluster centroids $c_{i1}, c_{i2}, \dots, c_{ik}$. For each pixel x_i in R_i , calculate the fuzzy membership of x_i to each cluster centroid using the following formula:

$$\mu_{ij} = \left[1 + \left(\left(\frac{d(x_i, c_{ij})}{d(x_i, c_{k+1})} \right)^2 \right)^{-1} \right] \quad (1)$$

where $d(x_i, c_{ij})$ is the Euclidean distance between pixel x_i and cluster centroid c_{ij} , and $d(x_i, c_{k+1})$ is the Euclidean distance between pixel x_i and the farthest cluster centroid c_{k+1} in sub-region R_i . Update the cluster centroids using the following formula:

$$C_{ij} = \left(\frac{\sum \mu_{ij}^m x_i}{\sum \mu_{ij}^m} \right) \quad (2)$$

where m is a fuzzifier parameter that controls the degree of fuzziness of the clustering. Repeat the above two steps until convergence. Merge the sub-region clusters to obtain the final segmentation result. PFCM is effective in segmenting images with non-uniform lighting and complex textures [Nandhini & Ashokkumar (2022)].

8.1.2. GLCM

Digital image processing analyzes texture using GLCM. Pixel pair frequency and spatial connection are represented by GLCM. It extracts plant disease picture attributes to identify healthy and ill plants. Pixel intensity statistics are used to calculate second-order gray-level degrees for texture analysis using GLCM [Surya et al. (2017)]. It produces orientation patterns and metrics including comparison, dissimilarity, entropy, energy, IDM, and correlation. Pictures are classified and clustered using GLCM matrix contrast, energy, and correlation [Sari et al. (2021)].

8.1.3. Local Directional Pattern (LDP)

LDP feature extraction can be utilized for object identification, segmentation, and classification. It creates texture descriptors for each pixel by calculating the local direction of edges [Chahi et al. (2018)]. The general mathematical formula for LDP is as follows:

$$LDP(x, y) = \sum_{i=0}^{m-1} [g(I(x, y) - I(x + r_i, y + c_i)) \times 2^i] \quad (3)$$

where: (x, y) = pixel position in the image, $I(x, y)$ = gray level of the pixel at (x, y) ,

r_i and c_i = displacement vectors used to calculate the direction of edges at pixel (x, y) , m = several bits used to represent the LDP code, $g()$ = function that maps the intensity difference to a binary code (0 or 1) based on a threshold value T .

Each pixel's binary LDP code encodes edge local direction [Ahmed et al. (2014)]. The gray level difference between the pixel and its neighbors along the displacement vectors r_i and c_i is used to calculate it. Using a threshold value T , the $g()$ method converts this difference to a binary code that specifies the descriptor's noise sensitivity. An image analysis feature vector may be created from each pixel's LDP code. Compact and resilient, it collects local texture information and may be utilized with other feature extraction algorithms for more complicated jobs.

8.1.4. Gabor Filter

The following equation describes the band-pass nature of the Gabor filter, which provides its two-dimensional impulse response in the spatial domain:

$$Gab(x, y) = \frac{1}{2\pi\sigma^2} e^{\left[-\frac{1}{2} \left(\frac{x^2 + y^2}{\sigma^2} \right) + j\omega(x\cos\phi + y\sin\phi) \right]} \quad (4)$$

where, σ = spatial width, ϕ = orientation angle and ω = frequency. Based on its parameters, the Gaussian kernel function may estimate Gabor filter forms. It can enhance images, smooth them,

and extract texture features. This Gabor filter is for texture extraction. Best results for edge-shaped texture recognition use an odd filter. Input picture and filter are convolved [Madiwalar & Wyawahare (2017)].

8.1.5. CNN

CNN is a popular deep learning technique for image categorization. CNNs can automatically assess plant leaf photos to classify cotton and soybean plant diseases. CNN has pooling, convolution, and totally linked layers [Yamashita et al. (2018)].

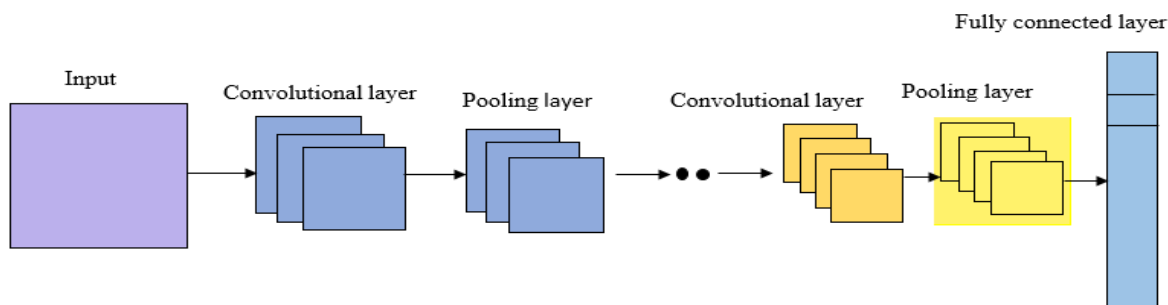


Fig 3. Structure of CNN [Gu et al. (2019)].

As seen in Fig 4., the fundamental components of CNN architecture could be separated into five categories.

The input layer of the neural network gets the primary data set in the form of the pixel values of individual pictures. Unique characteristics are extracted from the input data using kernels in the convolutional layers, with different kernels capturing various qualities. The fully connected layer generates probability for different outcomes by combining and normalizing the characteristics collected from earlier convolutions. The number of neurons in the output layer is tailored to individual needs and is dictated by the nature of the layer itself.

CNNs are trained on a massive dataset of annotated photos to identify disease and healthy plant traits. The network then classifies fresh pictures using this knowledge [Li et al. (2020)].

CNNs are good at picture categorization because they can recognize patterns at different abstraction levels [Picon et al. (2019)]. They use a hierarchical, multi-layered system that learns more complex information from incoming photos to detect healthy and ill plants. CNNs can swiftly and correctly classify cotton and soybean plant diseases, helping farmers and researchers control the illness and limit crop loss. This can boost agricultural yields and make agriculture more sustainable.

8.1.6. MRFO

Image processing and computer vision employ MRFO deep learning. Through feature extraction at different scales, MRFO improves picture classification. Conventional deep-learning algorithms extract features from images at a fixed resolution. Images may include useful information at different resolutions, improving categorization accuracy. The MRFO method extracts features at multiple sizes or resolutions and combines them into a feature vector. A deep learning model is updated to incorporate numerous branches that process the input picture at different scales to perform MRFO. The output of each node is merged before categorization by a fully connected layer. Combining CNN and MRFO to forecast cotton and soybean plant diseases improves

CNN construction is shown in Figure 4.

accuracy. The model better captures modest plant disease symptom changes across scales by extracting information at numerous resolutions. This improves crop management and treatment forecasts. Overall, CNN-MRFO hybridization for cotton and soybean plant disease prediction can boost crop yields and minimize plant disease impacts on agricultural output.

9.2. Proposed methodology

The suggested approach uses image processing and deep learning algorithms to create a fully automated system for identifying plant diseases. Specifically, the system focuses on detecting diseases in cotton and soybean plants, two crops that are essential for global food security. The system involves collecting two separate datasets of images of cotton and soybean plant leaves affected by different diseases. The collected image datasets undergo pre-processing using a Gaussian filter to remove noise and blur the images. Then, the images are segmented into separate parts according to color and texture using PWFCM clustering. Features are extracted from each segmented region using GLCM, LDP, and Gabor Filter. Next, the extracted features are used to train a CNN classifier to classify the diseases present in the image. The CNN is hybridized with the MRFO algorithm to get a higher degree of precision in disease prediction. The MRFO algorithm optimizes the features at different scales to capture more information about the diseases present in the image. Finally, an algorithm is developed to estimate the severity of the diseases present in the image based on the size and number of lesions or other visual symptoms present in the image. The proposed system is expected to provide an accurate and efficient means of plant disease detection and severity estimation, which could help farmers make informed management decisions and minimize crop losses. The flowchart given below in Figure 5 illustrates the working of the proposed methodology.

8.2.1. Proposed diagram

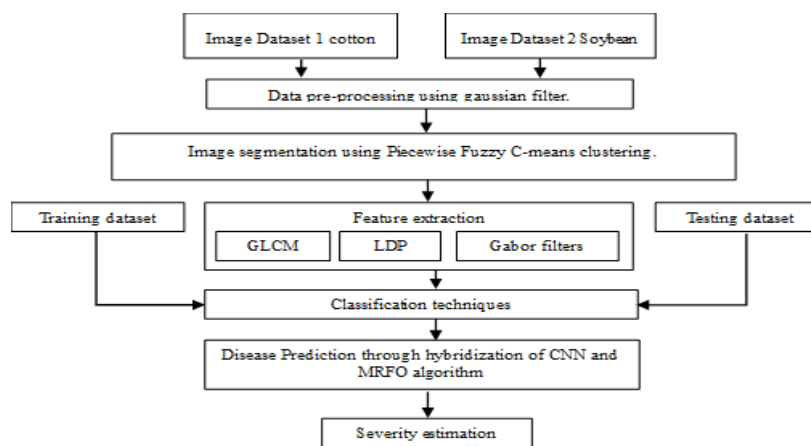


Fig 5 proposed methodology

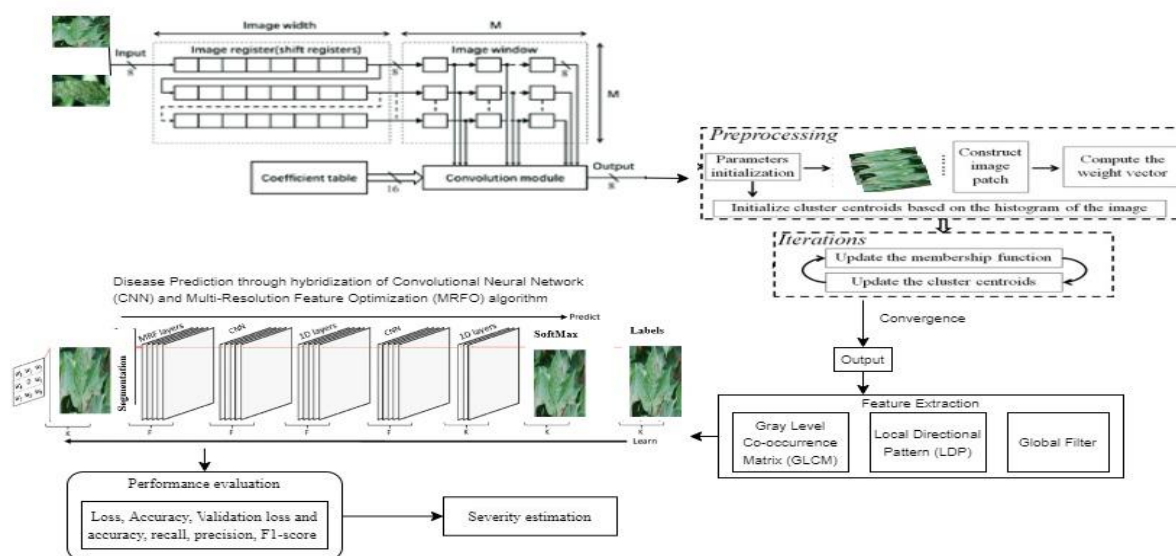


Fig 6. Multi-disease classification and severity estimation model

8.2.2. Proposed algorithm

ALGORITHM 1: Multi-disease classification and severity estimation

Input assumptions:

- 1: Let D_c be the dataset of images of cotton leaves affected by different diseases, and D_s be the dataset of images of soybean plant leaves affected by different diseases.
- 2: Let I_i be the i th image in dataset D_c or D_s .
- 3: Let K and σ be the kernel size and standard deviation of the Gaussian filter, respectively.
- 4: Let C_i be the set of clusters obtained by applying Piecewise Fuzzy C-means clustering on image I_i , with the number of clusters denoted by N .
- 5: Let $F_{i,j}$ be the set of features extracted from the j th cluster of image I_i using GLCM, LDP, and Global Filter.
- 6: Let X_i be the set of features extracted from image I_i , i.e., $X_i = \{F_{i,1}, F_{i,2}, \dots, F_{i,N}\}$.
- 7: Let Y be the set of labels corresponding to the diseases present in the images.
- 8: Let M be the number of images in dataset D_c or D_s .
- 9: Let $P(X_i)$ be the predicted label of image I_i using the trained CNN and MRFO algorithm.
- 10: Let $S(I_i)$ be the severity level of the disease present in image I_i , estimated using an algorithm based on the size and number of lesions or other visual symptoms present in the image.

Start

Phase I: Dataset collection :

Step 1: Given datasets D_c and D_s , collect M images of cotton leaves and soybean plant leaves affected by different diseases.

Phase II: Data pre-processing :

Step 2: Apply a Gaussian filter to each image I_i in D_c and D_s to obtain filtered image G_i , with kernel size K and standard deviation σ .

Phase III: Image segmentation :

Step 3: Apply Piecewise Fuzzy C-means clustering on each filtered image G_i to segment it into N clusters C_i .

Phase IV: Feature extraction :

Step 4: Step 4: Extract features from each cluster of each image I_i using GLCM, LDP, and Global Filter to obtain feature set $F_{i,j}$.

Phase V: Classification :

Step 5: Train a CNN on the set of features $X = \{X_i \mid i=1 \text{ to } M\}$ and corresponding labels Y to classify the diseases present in the images and then further perform testing.

Phase VII: Model performance enhancement:

Step 6: Hybridize the trained CNN with the MRFO algorithm to improve the accuracy of disease prediction.

Phase IX: Disease Prediction:

Step 7: Given a new image I , apply the trained CNN-MRFO model to extract features X_i and predict the label $P(X)$ corresponding to the disease present in the image.

Phase X: Severity estimation using algorithm 1.1.

Step 8: Estimate the severity level $S(I)$ of the disease present in the image using an algorithm based on the size and number of lesions or other visual symptoms present in the image.

End

ALGORITHM 1.1: Multi-disease classification and severity estimation

Start

Let the input be a result array R of size $(n \times m)$, where n is the number of plant leaves and m is the number of disease types.

Step 1: Let num be the maximum value of R :

$$num = \max(R)$$

Step 2: Let Val be the severity score of the disease in the plant leaf, which is obtained by multiplying num by 100

$$Val = num * 100$$

Step 3: Let $output$ be the string representation of Val with 9 decimal places:

$$output = "{:.9f}".format(Val)$$

Step 4: Let a be the index of the maximum value in R , which represents the disease type:

$$a = \operatorname{argmax}(R, \text{axis}=1)$$

Step 5: Let $Type$ be the type of plant disease based on the value of a , where:

$$Type = \text{'Disease name' if } a = 0 \text{ and so on till } a = 9$$

Step 6: Let V be the severity of the disease based on the value of Val , where:

$$V = \text{'Low' if } Val \leq 80$$

$$V = \text{'Average' if } 50 \leq Val \leq 80$$

$$V = \text{'High' if } Val > 80$$

Step 7: Let St be the string that represents the severity of the disease in the plant leaf, where:

$$St = \text{"Severity of \{ \} in this leaf is \{ \}"}.\text{format}(Type, V)$$

Step 8: Print St , which is the final output indicating the severity of the disease in the plant leaf identified by the algorithm.

End

9. Result and discussion

9.1. Soybean disease classification

The model was trained for a total of ten epochs; for each epoch, the loss and accuracy were recorded for both the training set and the validation set. The loss decreased significantly over the epochs, indicating that the model was improving in performance. At the end of the training, there was no overfitting because the model scored 97.50% accuracy on the training data and 95.51% accuracy on the validation data. The given table 4 shows the results of training a hybridized CNN and Multi-Resolution Feature Optimization model for predicting cotton disease.

Table 4. Training and validation loss and accuracy for each epoch

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
Epoch 1	1.7342	0.4464	0.9813	0.7013
Epoch 2	0.8017	0.7255	0.8047	0.6872
Epoch 3	0.5930	0.7891	0.5591	0.8513
Epoch 4	0.4573	0.8400	0.4766	0.8541
Epoch 5	0.3658	0.8809	0.4028	0.9004
Epoch 6	0.3351	0.8945	0.3054	0.9341
Epoch 7	0.2905	0.9082	0.2818	0.9299
Epoch 8	0.2513	0.9236	0.2412	0.9551
Epoch 9	0.2256	0.9382	0.1774	0.9734
Epoch 10	0.1841	0.9536	0.1748	0.9607

Figure 7 is a graph depicting the model's training loss and accuracy throughout its 10 epochs of training.

The classification of soybean plant disease using the proposed model yielded impressive results. The model obtained 0.96 accuracy, which means it successfully recognized 96% of occurrences of diseases in soybean plants. The precision and recall values for each class of disease are also high, indicating that the model has good accuracy in identifying the different types of soybean plant diseases as shown in table 5.

Table 5. Performance metric values for different soybean plant diseases.

Disease	Precision	Recall	F1-score
Yellow Mosaic	1.00	1.00	1.00
Frogeye Leaf Spot	1.00	1.00	0.69
Healthy	0.75	1.00	1.00
Sudden Death Syndrome	1.00	0.89	0.97
Yellow Mosaic + Target Leaf Spot	1.00	1.00	1.00
Target Leaf Spot	1.00	1.00	1.00
Frogeye LS + Sudden Death Syndrome	1.00	1.00	1.00
Bacterial Pustule	0.83	1.00	1.00

Rust	1.00	1.00	1.00
Accuracy			0.96
Macro avg	0.96	0.95	0.94
Weighted avg	0.97	0.95	0.95

representation for each epoch is illustrated in Figure 8.

The highest precision and recall values were achieved for the Yellow Mosaic disease and Yellow Mosaic + Target Leaf Spot disease, both achieving perfect scores of 1.00 for precision, recall, and F1-score. The model also performed well for the other types of diseases, such as Rust and Frogeye Leaf Spot, achieving perfect precision and recall scores of 1.00. The only exception was the Frogeye Leaf Spot disease, which had a lower F1-score of 0.69, indicating that the model may have had difficulty identifying this disease. Precision, recall, and F1-score all averaged out to 0.96, 0.95, and 0.94 across all classes on a global and weighted basis, respectively, demonstrating that the approach performs consistently across all classes.

9.2. Cotton disease classification

The given table 6 shows the results of training a hybridized CNN and Multi-Resolution Feature Optimization model for predicting cotton disease. The model was trained for 10 epochs, the training and validation loss and accuracy are recorded, and the graphical

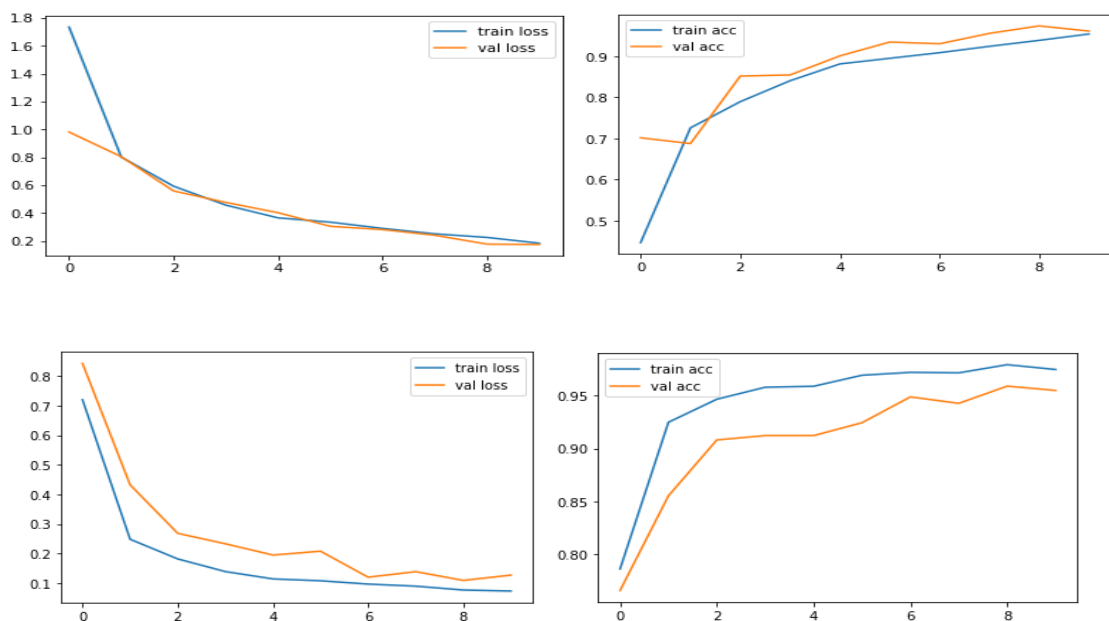


Fig 8. Graphical plot of training and validation loss and accuracy

The results indicated that the model's accuracy in estimating cotton disease greatly improved across epochs, suggesting that it performed very well. After finishing training, the model was 97.50% accurate on the training data and 95.51% accurate on the validation data. This means the model is not overfitting and is successful on both data.

Table 6. Training and validation loss and accuracy for each epoch

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
Epoch 1	0.7201	0.7857	0.8432	0.7653
Epoch 2	0.2476	0.9249	0.4321	0.8551
Epoch 3	0.1810	0.9468	0.2677	0.9082
Epoch 4	0.1382	0.9581	0.2323	0.9122
Epoch 5	0.1134	0.9591	0.1942	0.9122
Epoch 6	0.1070	0.9695	0.2072	0.9245
Epoch 7	0.0958	0.9722	0.1191	0.9490
Epoch 8	0.0888	0.9718	0.1379	0.9429
Epoch 9	0.0758	0.9795	0.1083	0.9592
Epoch 10	0.0722	0.9750	0.1262	0.9551

The classification result of cotton disease prediction obtained by hybridized CNN and MRFO illustrate that the overall accuracy achieved by the model is 0.95, which is a good indication of its performance. For each class, the recall, precision, and F1-score values are provided, which indicate the model's ability to correctly identify each class. The Aphids, Armyworm, Bacterial Blight, Powdery Mildew, and Target spot classes have a precision, recall, and F1-score of 1.00 given in Table 7 indicating that the model has correctly identified these classes without any false positives or false negatives.

Table 7. Performance metric values for different cotton plant diseases

Disease	Precision	Recall	F1-score
Aphids	1.00	1.00	1.00
Armyworm	1.00	1.00	1.00
Bacterial Blight	1.00	1.00	1.00
Healthy leaf	1.00	0.54	0.70
Powdery mildew	1.00	1.00	1.00
Target spot	1.00	1.00	1.00
Diseased cotton leaf	1.00	1.00	1.00
Diseased cotton plant	1.00	1.00	1.00
Fresh cotton leaf	0.59	0.59	0.74
Fresh cotton plant	1.00	1.00	1.00
Accuracy			0.95
Macro avg	0.96	0.95	0.94
Weighted avg	0.97	0.95	0.94

The affected cotton leaf, diseased cotton plant, and fresh cotton plant classes all have scores of 1.00 for precision, recall, and F1-score, which indicates that the model accurately identified each of these classes without producing any false positives or false negatives. The classes "healthy leaf" and "fresh cotton leaf" had a lower precision, recall, and F1 score, which indicates that the model had some difficulty detecting these classes. This is mostly true for the healthy leaf class. Particularly, the Healthy Leaf has a recall of only 0.54, which suggests that the model is not very good at recognizing this class when it is present. This is because the model is not very effective at distinguishing between different types of leaves.

Overall, the model performs well with a macro-avgF1-score of 0.94 and a weighted average F1-score of 0.95, indicating that it correctly identifies most classes with considerable precision and recall.

9.3. Severity analysis of cotton and soybean plant disease

The severity estimation of soybean leaves was calculated using the hybridized CNN and MRFO model. The predicted class was obtained by finding the index of the maximum value in the output, which corresponded to Bacterial Pustule with a severity score of 99.127590656. Based on this score, the severity level was determined to be "High" as it was greater than 80%. Therefore, the severity estimation of the soybean leaf is "Severity of Bacterial Pustule in this leaf is High". Similarly, the severity estimation of cotton leaves was determined using the hybridized CNN and Multi-Resolution Feature Optimization model, which resulted in a severity score of 91.213607788. The disease type was identified as "Bacterial Blight" and the severity level was again determined to be "High" based on the severity score. Therefore, the final output is "Severity of Bacterial Blight in this leaf is High".

10. Conclusion and Future Scope

The hybrid CNN-MRFO model had an accuracy rate of 95% and 96%, respectively, when it came to diagnosing cotton and soybean plant diseases. This indicates that the model did pretty well overall. This hybrid approach has the potential to help farmers in the early identification and management of plant diseases in cotton and soybean crops. This, in turn, will lead to increased crop yields and economic advantages. As evidenced by its good recall, accuracy, and F1 scores, this hybrid technique has the ability to assist farmers in this process. The model makes use of a combination of CNN and MRFO techniques, which is demonstrated by the exceptionally high accuracy with which it can forecast illnesses, particularly those affecting cotton.

The hybrid CNN-MRFO model may be improved in the future for larger crop disease applications, integrated into agricultural monitoring systems, and the dataset can be expanded to better understand its potential in a variety of settings. The use of this approach has the potential to improve the monitoring and control of crop diseases.

11. Funding

None

12. Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Abade, A., Ferreira, P. A., & de Barros Vidal, F. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Computers and Electronics in Agriculture*, 185, 106125. (<https://doi.org/10.1016/j.compag.2021.106125>)
- [2] Ahmed, F., Kabir, M. H., Bhuyan, S., Bari, H., & Hossain, E. (2014). Automated weed classification with local pattern-based texture descriptors. *Int. Arab J. Inf. Technol.*, 11(1), 87-94.
- [3] Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf disease classification. *Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband*.
- [4] Annabel, L. S. P., Annapoorani, T., & Deepalakshmi, P. (2019, April). Machine learning for plant leaf disease detection and classification—a review. In 2019 international conference on Communication and signal processing (ICCSP) (pp. 0538-0542). IEEE. (<https://doi.org/10.1109/ICCSP.2019.8698004>)
- [5] Appalanaidu, M. V., & Kumaravelan, G. (2021, February). A review of machine learning approaches in plant leaf disease detection and classification. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 716-724). IEEE. (<https://doi.org/10.1109/ICICV50876.2021.9388488>)
- [6] Bagde, S., Patil, S., Patil, S., & Patil, P. (2015). Artificial neural network based plant leaf disease detection. *Int J Comput Sci Mob Comput*, 4(4), 900-905.
- [7] Barbedo, J. G. A. (2016). A review on the main challenges in

- automatic plant disease identification based on visible range images. *Biosystems engineering*, 144, 52-60. (<https://doi.org/10.1016/j.biosystemseng.2016.01.017>)
- [8] Chahi, A., Ruichek, Y., & Touahni, R. (2018). Local directional ternary pattern: A new texture descriptor for texture classification. *Computer vision and image understanding*, 169, 14-27. (<https://doi.org/10.1016/j.cviu.2018.01.004>)
- [9] Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanekaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, 173, 105393. (<https://doi.org/10.1016/j.compag.2020.105393>)
- [10] Fan, X., Luo, P., Mu, Y., Zhou, R., Tjahjadi, T., & Ren, Y. (2022). Leaf image based plant disease identification using transfer learning and feature fusion. *Computers and Electronics in Agriculture*, 196, 106892. (<https://doi.org/10.1016/j.compag.2022.106892>)
- [11] Gharge, S., & Singh, P. (2016). Image processing for soybean disease classification and severity estimation. In *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2015, Volume 2* (pp. 493-500). Springer India. (https://doi.org/10.1007/978-81-322-2553-9_44)
- [12] Ghatol, N., & Dhok, G. P. (2019). Texture Based Features Approach for Crop Diseases Classification and Diagnosis-a Research. no. May, 7469-7474.
- [13] Gu, H., Wang, Y., Hong, S., & Gui, G. (2019). Blind channel identification aided generalized automatic modulation recognition based on deep learning. *IEEE Access*, 7, 110722-110729. (<https://doi.org/10.1109/ACCESS.2019.2934354>)
- [14] <http://cotton.tamu.edu/Photos/diseasephotos/images/10bacterialblight.jpg>
- [15] https://fieldcropnews.com/2016/07/soybean_disease_resources/
- [16] Jadhav, S. B., Udup, V. R., & Patil, S. B. (2019). Soybean leaf disease detection and severity measurement using multiclass SVM and KNN classifier. *International Journal of Electrical and Computer Engineering*, 9(5), 4092.
- [17] Kaur, S., Pandey, S., & Goel, S. (2019). Plants disease identification and classification through leaf images: A survey. *Archives of Computational Methods in Engineering*, 26, 507-530. (<https://doi.org/10.1007/s11831-018-9255-6>)
- [18] Khirade, S. D., & Patil, A. B. (2015, February). Plant disease detection using image processing. In 2015 International conference on computing communication control and automation (pp. 768-771). IEEE. (<https://doi.org/10.1109/ICCUBEA.2015.153>)
- [19] Kumari, C. U., Prasad, S. J., & Mounika, G. (2019, March). Leaf disease detection: feature extraction with K-means clustering and classification with ANN. In 2019 3rd international conference on computing methodologies and communication (ICCMC) (pp. 1095-1098). IEEE. (<https://doi.org/10.1109/ICCMC.2019.8819750>)
- [20] Li, Y., Nie, J., & Chao, X. (2020). Do we really need deep CNN for plant diseases identification?. *Computers and Electronics in Agriculture*, 178, 105803. (<https://doi.org/10.1016/j.compag.2020.105803>)
- [21] Madiwalar, S. C., & Wyawahare, M. V. (2017, February). Plant disease identification: A comparative study. In 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI) (pp. 13-18). IEEE.
- [22] Nandhini, S., & Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Computing and Applications*, 1-22. (<https://doi.org/10.1007/s00521-021-06714-z>)
- [23] Parikh, A., Raval, M. S., Parmar, C., & Chaudhary, S. (2016, October). Disease detection and severity estimation in cotton plant from unconstrained images. In 2016 IEEE international conference on data science and advanced analytics (DSAA) (pp. 594-601). IEEE. (<https://doi.org/10.1109/DSAA.2016.81>)
- [24] Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, 161, 280-290. (<https://doi.org/10.1016/j.compag.2018.04.002>)
- [25] Prasad, S., Peddoju, S. K., & Ghosh, D. (2016). Multi-resolution mobile vision system for plant leaf disease diagnosis. *Signal, image and video processing*, 10(2), 379-388. (<https://doi.org/10.1007/s11760-015-0751-y>)
- [26] Prashar, K., Talwar, R., & Kant, C. (2015). A review on efficient identification of american cotton leaf diseases through training set. *Int J Comput Appl*, 132(7), 32-39.
- [27] Ranjan, M., Weginwar, M. R., Joshi, N., & Ingole, A. B. (2015). Detection and classification of leaf disease using artificial neural network. *International Journal of Technical Research and Applications*, 3(3), 331-333.
- [28] Rastogi, A., Arora, R., & Sharma, S. (2015, February). Leaf disease detection and grading using computer vision technology & fuzzy logic. In 2015 2nd international conference on signal processing and integrated networks (SPIN) (pp. 500-505). IEEE. (<https://doi.org/10.1109/SPIN.2015.7095350>)
- [29] Rothe, P. R., & Kshirsagar, R. V. (2015, January). Cotton leaf disease identification using pattern recognition techniques. In 2015 International conference on pervasive computing (ICPC) (pp. 1-6). IEEE. (<https://doi.org/10.1109/PERVASIVE.2015.7086983>)
- [30] Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, 8(11), 468. (<https://doi.org/10.3390/plants8110468>)
- [31] Sarangdhar, A. A., & Pawar, V. R. (2017, April). Machine learning regression technique for cotton leaf disease detection and controlling using IoT. In 2017 international conference of electronics, communication and aerospace technology (ICECA) (Vol. 2, pp. 449-454). IEEE. (<https://doi.org/10.1109/ICECA.2017.8212855>)
- [32] Sari, Y., Baskara, A. R., & Wahyuni, R. (2021, November). Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods. In 2021 Sixth International Conference on Informatics and Computing (ICIC) (pp. 1-4). IEEE. (<https://doi.org/10.1109/ICIC54025.2021.9632920>)
- [33] Sathish, P., Elango, N. M., & Thirunavukkarasu, V. (2020). Piecewise fuzzy C-means clustering and deep convolutional neural network for automatic brain tumour classification using MRI images. *Test Eng. Manage*, 83, 3729-3736.

- [34] Surya, R. A., Fadlil, A., & Yudhana, A. (2017). Ekstraksi ciri metode Gray Level Co-Occurrence Matrix (GLCM) dan Filter Gabor untuk klasifikasi citra batik pekalongan. *Jurnal Informatika: Jurnal Pengembangan IT*, 2(2), 23-26.
- [35] Tetila, E. C., Machado, B. B., Astolfi, G., de Souza Belete, N. A., Amorim, W. P., Roel, A. R., & Pistori, H. (2020). Detection and classification of soybean pests using deep learning with UAV images. *Computers and Electronics in Agriculture*, 179, 105836. (<https://doi.org/10.1016/j.compag.2020.105836>)
- [36] Wu, Q., Zhang, K., & Meng, J. (2019). Identification of soybean leaf diseases via deep learning. *Journal of The Institution of Engineers (India): Series A*, 100, 659-666. (<https://doi.org/10.1007/s40030-019-00390-y>)
- [37] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9, 611-629. (<https://doi.org/10.1007/s13244-018-0639-9>)
- [38] Zhang, K., Wu, Q., & Chen, Y. (2021). Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN. *Computers and Electronics in Agriculture*, 183, 106064. (<https://doi.org/10.1016/j.compag.2021.106064>)