AI-based Intelligent Information System for Plant Leaves Disease Detection

Madan Mohan Mishra
Department of Computer Application & Information Technology and Science
A K S University, Satna MP, India

Pramod Singh, PhD
Department of Computer Application & Information Technology and Science
A K S University, Satna MP, India

ABSTRACT
An Enhanced Intelligent Information System (EIIS) that uses plant leaf images is increasingly applied in agriculture to identify plant leaves diseases. That is processed on feature extraction and classification hybrid model with its techniques that have been developed for this purpose, and their performance is typically evaluated from the developers' point of view. However, this study aims to assess the effectiveness of EIIS in detecting image based plant leaves diseases detection from the perspective of the users. That is an AI-based Intelligent Information System that integrates artificial intelligence techniques, data processing, and user interface components to deliver effective decision support, where information systems are used to guide the health of the crops. These systems have enabled the advancement of smart farming by improving productivity, precision, and resource efficiency through data-driven insights. This study proposes a hybrid approach for plant disease detection that combines Convolutional Neural Networks (CNN) for feature extraction and Weighted K-Nearest Neighbors (KNN) algorithm is used for classification. The CNN model efficiently extracts discriminative features from leaf images, while the Weighted KNN classifier assigns greater importance to closer data points, enhancing classification accuracy. The proposed method is applied to potato plant leaf disease recognition and achieves an accuracy of 98.75%. Furthermore, the system is implemented with a Tkinter-based graphical user interface, ensuring user-friendly interaction and practical usability. The overall framework demonstrates high accuracy, efficiency, and accessibility, making it a reliable intelligent decision support system for real-world agricultural applications.

General Terms
Artificial Intelligence, Plant Disease Detection, Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Weighted K-Nearest Neighbors (WKNN), Image Processing, Potato Leaf Disease Classification, Intelligent Decision Support System, Smart Agriculture, Feature Extraction, Tkinter GUI

Keywords
Computer vision, CNN, Machine Learning, Precision farming, ROI, Weighted KNN.

1. INTRODUCTION
Agriculture is one of the key pillars that is ensuring for food security and economic development worldwide. However, plant diseases significantly affect crop quality and productivity, leading to major losses for farmers and the agricultural industry. Early and accurate detection of plant diseases is essential for improving crop health, increasing yield, and reducing the excessive use of pesticides. Traditional disease identification methods mainly depend on manual observation by agricultural experts, which can be time-consuming, costly, and less effective for large-scale farming with the rapid advancement of artificial intelligence and image processing technologies, Enhanced Intelligent Information Systems (EIIS) have emerged as effective tools for automated plant disease detection. These systems combine artificial intelligence techniques, data processing methods, and user-friendly interfaces to support accurate decision-making in agriculture. By analyzing plant leaf images, EIIS can identify diseases at an early stage and provide reliable guidance for crop management. Such intelligent systems contribute to the development of smart farming by improving productivity, precision, and resource efficiency through data-driven solutions is a part of Agriculture 4.0 [1]. Recent studies have introduced various machine learning and deep learning models for plant disease detection. Among these approaches, Convolutional Neural Networks (CNN) has shown excellent performance in extracting important visual features from plant leaf images [2]. Similarly, classification techniques such as K-Nearest Neighbors (KNN) are widely used due to their simplicity and effectiveness [3]. That is parallel discussed with intelligent information System and classification of Plant disease detection.

2. RELATED WORK ON EFFICIENCY OF INTELLIGENT INFORMATION SYSTEM
Recent advancements in intelligent agriculture have led to the development of various machine learning and deep learning approaches for plant disease detection. Improving classification accuracy and usability from the users' perspective remains an important research area. To address these challenges, this study proposes a hybrid approach that combines CNN for feature extraction and Weighted KNN for classification in vegetable plant leaf disease detection. The CNN model extracts meaningful and discriminative image features, while the Weighted KNN classifier improves prediction accuracy by assigning greater importance to nearby data points. In addition, a Tkinter-based graphical user interface is developed to provide a simple and practical platform for users. The proposed EIIS framework aims to deliver high accuracy, efficiency, and accessibility, making it a reliable intelligent decision support system for modern agricultural applications. Because an AI-based Intelligent Information System is the set of core component of learning, reasoning, perception, heuristic intelligence, and problem solving, discussed in Fig 1.

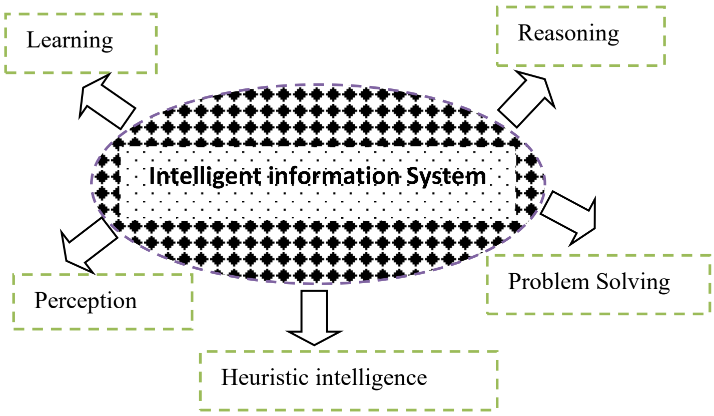


Fig 1: Intelligent information System

It is a system that collects, processes, and uses data with some level of intelligence to support decision-making, but without the advanced feature handling, i.e. real-time learning or highly complex prediction models [4], that is applicable in plant disease detection, while enhanced intelligent Information system is too rich than normal Intelligent information System, because it is an advanced type of information system, that not only process data but also learns, predicts, and makes smart decisions automatically using advancement of artificial intelligence and data analytics. That is a set of components that can run together for receive, process and broadcast the information for decision support system and eager for constant learning that can learn from current or past data processing. These components are the set of hardware, software, data, procedures, and users. That plays a crucial role in transforming raw data into meaningful and useful information. Fig 2 in discussed Enhanced Intelligent Information is the primary purpose of an information system is to support decision-making, coordination, control, and analysis within an intelligent system. By efficiently processing and broadcasting information to the right decision at the right time, it enables famers and decision-makers to make informed, timely, and accurate decisions [4][5][6]. That can represent with the input features are converted with each component into quantifiable features in equation (1), that is proposed for block diagram of Fig 2, such as :

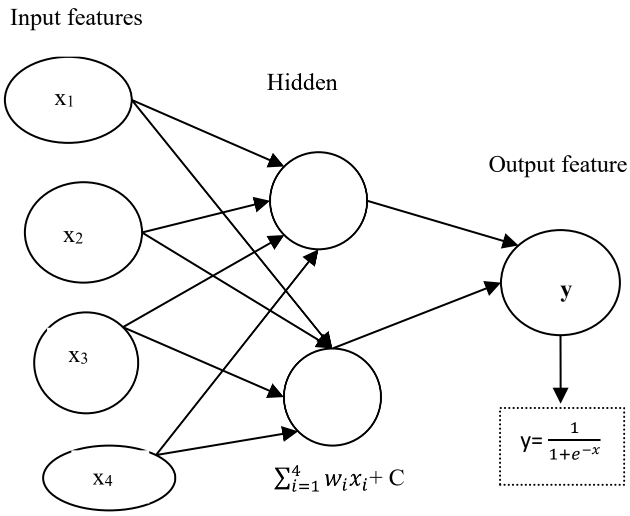


Fig 2: Artificial Neural Network

- User Interface Interaction:
  User Interface Interaction features are count on: number of clicks, response time, and user input frequency, usability score.
- Model-Based Logic:
  Model-Based Logic that are rule activation count, decision confidence score, and model output probability.
- Learning:
  Learning features are based on training accuracy, loss value, number of epochs & adaptation rate.
- Knowledge Base:
  Knowledge Base is direct related to number of stored rules, retrieval accuracy, and knowledge relevance score.
- Response:
  Response is a kind of responsive parameter that is used to count:
  If repeated intentionally, treat as:
  User feedback score or Interaction quality after prediction.

A neural network-based intelligent information system can be developed to improve agricultural productivity and support decision-making in crop management. The proposed model uses four input features: user identification (x1), number of clicks (x2), response time in milliseconds (x3), and input frequency (x4).

These features are processed through an artificial neural network consisting of an input layer with four neurons, one or two hidden layers containing 8–16 neurons, and an output layer that predicts the usability score (y). The sigmoid activation function is used to introduce non-linearity into the model. Where X=[x1, x2, x3, x4] are 4 common input features are there. Such as W = weights.

C = bias feature, that is based on the features of error.

f = activation function (Sigmoid)

And a= Represent the whole application, the mathematical representation of the neural network is given as:

$a = \sum_{i}^{4} w_i x_i \ + C$ ------ (1)

This enhanced intelligent information system supports efficient data analysis, improved communication, and better operational rformance in agriculture. In today's dynamic environment, formation systems play a significant role in increasing agricultural growth, improving economic conditions, and enhancing living standards.

Furthermore, the food processing sector contributes significantly to product exports and economic development, particularly in developing countries such as India, where growth is driven by both export demand and domestic consumption. Crop diseases remain a major challenge in Indian agriculture, causing yield losses ranging from 2% to 74%, and in severe conditions, losses may reach 70–80%. Certain disease ecosystems also report severity index values as high as 52.6%. Therefore, intelligent information systems integrated with machine learning techniques can help in early disease detection, decision support, and sustainable agricultural productivity improvement. [5][6][7].That are the direct impact to Agri production, cutting yields by as much as 50% in rice, potatoes, sugarcane, tomatoes, bananas, and citrus fruits. The K-Nearest-Neighbors (KNN) algorithm, a traditional instance-based lazy learning machine learning approach. That is rich for single sample classification by measuring distances between a test sample and the closest training instances using a chosen metric as Fig 3. KNN depends entirely on manually engineered or pre-computed features, making its effectiveness sensitive to feature quality and prone to poor results with high-dimensional images or massive datasets. It also incurs high inference costs, as it must compute distances from every test sample to the entire training set. By comparison, pre-trained Convolutional neural networks (CNNs) Efficient Net, Dense Net, and Mobile Net trained on vast datasets offer a powerful alternative, although they require huge memory to store its parameters and activations [5]. KNN requires storing and comparing against the entire training set at inference, leading to high computational cost and slow prediction for large datasets. That is best fit when features are continuous and normalized direct distance comparisons between test and training samples. This makes its performance highly sensitive to feature quality and causes it to degrade on high-dimensional image data. In contrast, deep learning models can automatically learn features from raw images and generalize better, while KNN underperforms in scalability, accuracy, and feature learning for real-world plant disease detection [2]. Machine learning methods KNN and Support Vector Machine are classification algorithm now enable precise identification of fungi, bacterial, powdery mildew, early and late blight in phomopsis blight while, These complement traditional structural and design techniques, as well as advanced computational methods such as data analysis and algorithmic modeling, and are called chemically inspired computational approaches because they work directly with the underlying structure and composition of data at a fundamental level[7]. Image processing involves various techniques to improve image quality or extract useful data. Advances in this technology have enabled a wide range of applications in biology. That are encompasses a range of techniques designed to enhance image quality or derive meaningful information. The evolution of such sophisticated technologies has unlocked numerous applications within the biological sciences [8]. While, many deep learning algorithms have emerged to detect patterns across various plant leaves imaging types, tackling issues in classification, object detection, and segmentation. The swift expansion of deep neural networks correlates directly with the growing supply of annotated datasets. Experts agree that bigger training datasets are vital for training deeper networks successfully, boosting performance and robustness while curbing over fitting as data volume rises. For example, the popular ImageNet dataset offers over 14 million mid-resolution images in 1000 categories, enabling pre-trained models for follow-up tasks. At the same time, KNN and SVM excel in classification with modest datasets and aid object detection, proving useful in hybrid algorithms [5]. That are used with color image segmentation is a crucial step in color image analysis.

It involves dividing an image to isolate meaningful regions of interest. Numerous techniques have been developed for color image segmentation; among which k-means clustering is one of the most commonly used methods [6]. Additionally, the study goes further than just creating the model by developing an easy-to-use Tkinter app. This app gives farmers a straightforward, packed-with-features interface to make quick, smart decisions. It predicts plant diseases and delivers customized advice drawn from the analysis. By combining Weighted KNN, ROI extraction and computer vision, the system nails precise disease identification, then farmers can check results on their phones and pick out damaged spots or disease signs in plant photos for instant analysis [7]. That is beginning with simplification of Thresholding technique is a part of feature selection. Its

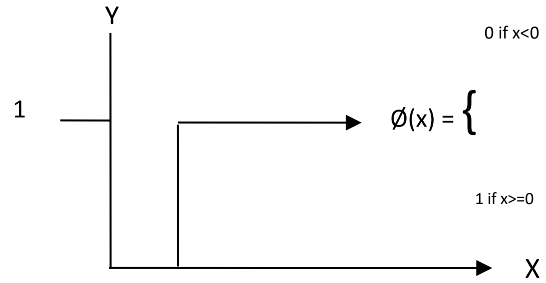


Fig 3: Thresholding function

Mapping of actual vs predicted on accuracy actual range=9.0 & predicted=8.8 that has been processed on level 2 rate on Error = 0.2, that is affected during overfitting occurs when the ANN model memorizes training data instead of learning general patterns. That Fig. 4, presents the information about training and validation loss curves indicates that the proposed ANN model did not suffer from significant overfitting on parameters of Input features in: clicks, response time, input frequency and output features as usability score on following data set is available in Table 1.

Table 1: ANN Training Data set

State	Clicks	Response_time_ms	Input_frequency	Usability_score
0	12	340	5	8.2
1	20	220	7	9.1
2	5	500	2	6.5
3	15	300	6	8.7
4	18	260	8	9.3
5	7	450	3	7.0
6	25	200	9	9.8
7	10	390	4	7.5
8	30	180	10	10.0
9	14	320	5	8.4

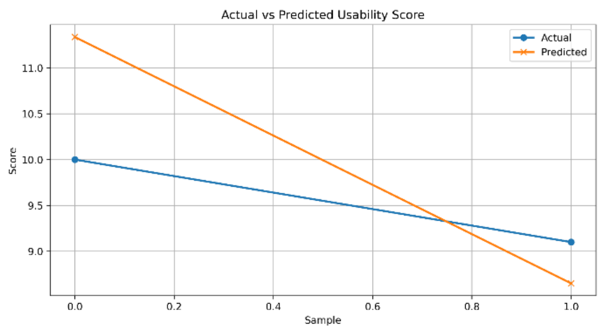


Fig 4: Actual vs predicted Score

The ANN Training and Validation Loss, Fig 5 explains about the model error. That has been changed on every training epochs. The horizontal axis shows the epoch number, and the vertical axis shows the loss value where the blue line labeled "Training Loss" shows the loss computed on the training data at each epoch & the orange line labeled "Validation Loss" shows the loss computed on the held-out validation data at each epoch. That curves tell us in following points:

a) If training and validation loss decrease together and eventually stabilize at a low value, it indicates that the model is learning useful patterns and is likely well-fit.

b) If training loss keeps decreasing but validation loss starts increasing, that is a sign of overfitting: the model is memorizing the training data and performing worse on unseen data.

c) If both losses stay high and do not drop much, the model may be under fitting.

Fig 5: The ANN Training and Validation Loss

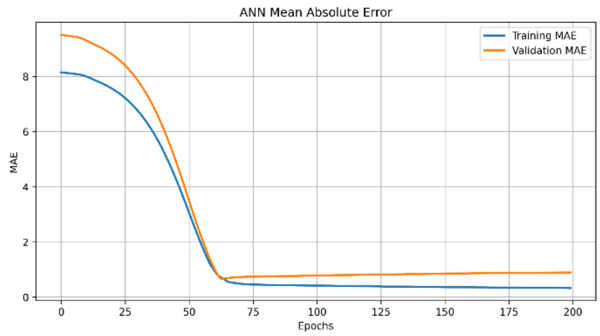


Fig 6: ANN Mean Absolute Error

Fig, 6 ANN Mean Absolute Error plot shows about that Mean Absolute Error starts at a relatively high value in early epochs (for example, around 1–2 or more, depending on random initialization) and gradually decreases as the model learns. And, by the end of 200 epochs, the training MAE typically settles at a low value, often somewhere in the range of 0.3–0.8 or even below, depending on your exact run and data scaling.

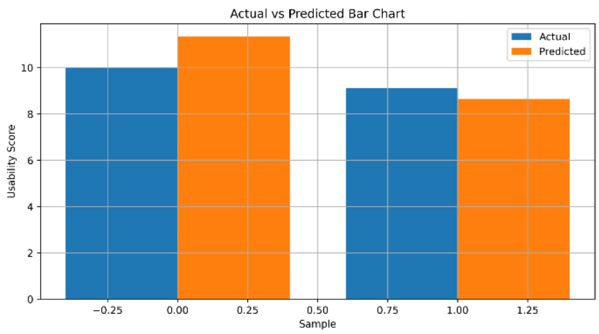


Fig 7: Bar Chart: Actual vs Predicted

Based on user's click, application response time and input comparison between the actual usability vs the predicted scores generated by the ANN model. that can indicating that the model has learned the data patterns effectively and can make reasonably accurate predictions are under Fig 7, that shows a small between 0.15 to 0.20 due to limited dataset.

3. RELATED WORK ON CLASSIFICATION OF PLANT DISEASE DETECTION USING CNN

Plant leaf disease detection is a type of data mining that can be categorized into two main tool classes: model building and pattern discovery [8]. Symptoms of plant diseases include leaf spots, color changes, mold, fungal infections, and pest invasion, all of which are becoming more prevalent due to nutritional deficiencies and changing climatic conditions [9]. By integrating information technology, Precision Agriculture helps farmers manage and access the latest farm information. It is a modern, technology-driven approach that supports various methods for farm improvement [10].

Advanced technologies allow for precise evaluation of crop disease severity. In the context of machine learning, precision farming utilizes predictive models to optimize farming processes. A Convolutional Neural Network (CNN) is a specialized neural network tailored for computer vision tasks that is leveraged for excellent work on classification and pattern detection, because it can automatically identify disease patterns from plant leaf images with high accuracy. Its key strength is the automatic discovery and learning of image features during training, by passing the manual feature engineering required in older methods. CNNs consist of core building blocks: Convolutional layers, pooling layers, and fully connected layers. At the heart are the Convolutional layers, which use compact numerical filters known as kernels to spot features in images. These kernels slide over the input to generate feature maps, each highlighting distinct patterns. The quantity of Convolutional layers typically scales with the input image's dimensions [11].

3.1 Convolution process

The Convolutional layer serves a vital role in Deep Convolutional Neural Networks (DCNNs) by detecting pixel relationships and extracting features straight from input images. That is a key technique for computer vision challenges. As illustrated in Figure 8 (this convolution process relies on a mathematical operation involving three primary elements: the image matrix, the filter or kernel and the stride. This study emphasizes the detection of crop diseases through image analysis through the Fig, 8 Pixels of Leaf Image processing in CNN, that can identified by the diseased areas on plants typically display patterns in circular, oval, square, or triangular forms.

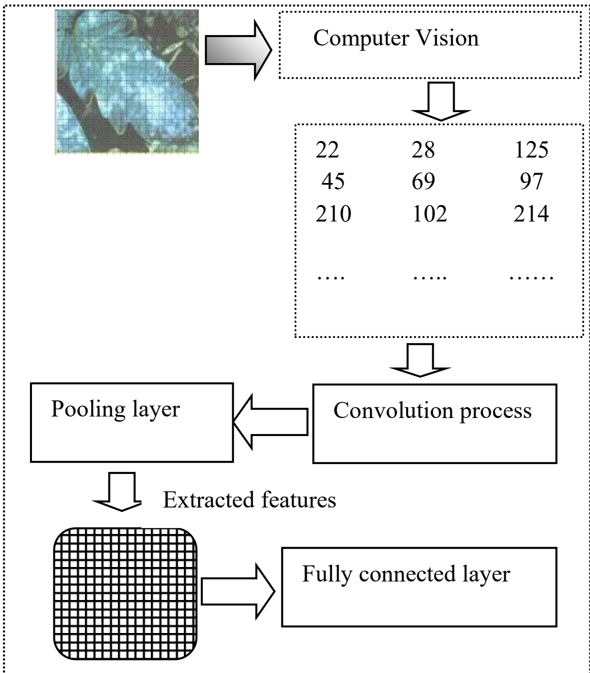


Fig 8: Pixels of Leaf, Image processing in CNN

3.2 Filters

Filters are the basic units of CNNs that is a small filter matrix used to recognize the discriminative patterns by using the learned weights that slides across the input image during the convolution operation. The filter detects specific local patterns like edges, textures, or more complex patterns allowing the network to extract discriminative information for subsequent layers. i.e. 3x3, 5x5, 7x7 or other dimensions. A filter or kernel that can be represented as a 3D tensor:

$K \in R^{UX\ RX\ C}$

Where U = height of the filter (e.g., 3, 5, 7…)
V = width of the filter (e.g., 3, 5, 7…)
C = number of channels in the input image (e.g., 3 for BGR, 1 for grayscale)., that is process with nn.conv2d()[12].
nn.Conv2d(in_channels=3,out_channels=32,kernel_size=3), that is categorized with input channels= 3 (color set)

3.3 Stride:

The stride is the number of pixels the filter (kernel) moves each time it slides across the input image during the convolution operation.
Stride = 1: the filter moves one pixel at a time → maximum overlap, larger output feature map.
Stride = 2: the filter moves two pixels at a time → less overlap, smaller output feature map.
Stride = S: moves S pixels at a time in both height and width directions.

3.4 Pooling Layer:

A Pooling Layer is a downsampling technique that is leveraged to feature maps produced by Convolutional layers. Its main purposes are:

- Reduce spatial dimensions (height × width) to lower computation and memory requirements.
- Extract dominant features and make the network more translation-invariant (less sensitive to small shifts or distortions).

3.5 Extracted features:

An extracted feature is the meaningful information captured from the input image by the network's layers
(Convolution + pooling).
That is represents with a feature encodes patterns or characteristics of the input image that are relevant for the task, such as:
Edges (vertical, horizontal, diagonal) in early layers
Textures or shapes in middle layers
Complex objects or semantic patterns in deeper layers.

That is the set of Convolution layer and Activation function (e.g., ReLU, Softmax) that keeps only strong signals (non-linear transformation).

3.6 Fully Connected Layer:

That is the combination set for all extracted features from the previous layers.
That can make final predictions (like class scores in classification).
That will apply in weighted KNN, where Leaf and fruit images were collected to assess the infection levels. Morphological characteristics of healthy plants include vibrant green leaves with abundant chlorophyll, which can now be recorded using mobile cameras with resolutions from 1 to 200 megapixels. Image segmentation serves as an important tool in image processing and preliminary visualization [12].
Healthy leaves are characterized by deep green color and freshness, showing that the plant is free from diseases and genetically healthy. Image segmentation is a method used in image processing for initial analysis and visualization of plant images [13].

4. SAMPLED DATA

Fig 9: Potato Leaf image

Fig 10: Eggplant Leaf

Fig 11: Tomato plant leaf

Fig 12: Eggplant leaf image

5. APPLIED METHODOLOGY

Weighted K-Nearest Neighbors (KNN) is a case-based learning approach that stores the entire training dataset for making classifications. As a lazy learner, it becomes inefficient for large-scale applications like dynamic web mining. To improve efficiency, one effective strategy involves choosing key representative points that condense the training data, turning it into an inductive learning model for classification. This representative model can handle predictions. Algorithms like decision trees and neural networks have been created for this task, and their effectiveness is typically assessed through performance metrics [14]. This unique work has been initiated with the participation of users. Who can examine whether any specific spot contains a disease by selecting a particular part of the leaf. The leaf is visualized based on its morphological characteristics. Plant diseases are identified through visual features such as blurred regions, dead leaf cells, or distinct patterns. These visual characteristics help in determining the type of disease present. Diseases such as bacterial spot, powdery mildew, and early blight are discussed here. That is here discussed with following steps, on:

Step 1: First process the image

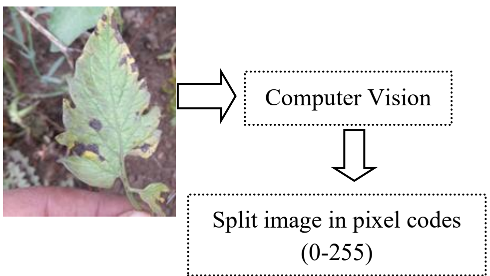


Fig 13: Image process in Computer Vision [28]

Image, is the set of pixels. That image information is present with of height, width and three color channels BGR (OpenCV uses BGR, blue green and red), let's image height=100 pixels, width=200 pixels and Color image Channels = 3 (BGR).

Then Total pixels = Height ×Width

Total pixels = 100×200

= 20,000 pixels

Consumed memory = Height×Width×3 Channels×1 byte.

= 20,000 pixels x 3 channel x1

= 60,000 bytes.

Step 2: Then pass the BGR code Image data to Grayscale mode:

Grayscale mode is a binary color model that is contained with black and white color channel. RGB primaries used in BT.601, i.e. discussed in Fig 14 Grayscale equation data have been processed on Grayscale mode because Grayscale (Y=0.299 + 0.587 +0.114) has only 1 intensity color channel. That is lower storage, faster process and consumes less data.

$$\text{Gray}(Y) = 0.299R + 0.587G + 0.114B$$

Fig 14: Grayscale equation

Its weighted Grayscale sum formula is the sum of Red, Green and Blue with luminance (Y) equation for much computer vision applications. That is specifies to pattern recognitions, such as edges, contours and intensity patterns [16] [17]. Are processed with sample image like Fig, 13 Potato Leaf in first step to pass the image for dataset Computer vision–based digital imaging technology represents a modern approach to plant disease detection and plays a crucial role in minimizing crop yield losses while improving crop quality. Pattern- and color-based disease detection methods offer higher precision and are especially effective in single-sample–based analysis. Since each disease exhibits distinct patterns and color characteristics, sample-based colors can be visualized within a specified range using Hue–Saturation–Value (HSV) bounds. These lower and upper HSV range limits are controlled through a tracker scale, enabling dynamic optimization of disease detection based on end-user requirements. The core RGB colors—red, green, and blue—enable vibrant visual displays, specified through hexadecimal codes where color intensity ranges from 00 to FF. Combining these values produces virtually any color. Manual inspection of infected crop leaves is unreliable, as the human eye overlooks subtle pixel-level variations. Each color has a unique code, with disease symptoms emerging as changes in leaf coloration—from small spots that grow larger as the infection spreads. A crucial first step is to segment affected regions by converting the image to grayscale in the Hue-Saturation-Value (HSV) color space, applying weighted K-Nearest Neighbors (KNN) for smoother and more precise results through compute distance to all training samples usually classify a new leaf X based on a labeled dataset of $y^1, y^2, y^2, \dots\dots\dots, y^n$.

Unknown Test leaf feature vectors of $x = [x_1, x_2, x_3, \dots\dots, x_m]^T$

And unknown training leaf feature vectors of

$$y^i = [y_1^i, y_2^i, y_3^i, \dots\dots, y_m^i]^T$$

(1) Euclidean distance:

$$d = \sqrt{\sum (x_i - y_i)^2}$$

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

That representation of x= [mean of H, contrast, area]

So, $x_1 = \text{Mean value}$, $x_2 = \text{Textrure contrast}$ and

$x3 = \text{Lesion area}$ then Select the nearest Neighbors samples

(2) Assign weights:

Common weighting: $w_i = \frac{1}{d_i}$

(3) Compute weighted voting per class

Let C_k be a class k, that is disease type

Sum the weights of all neighbors belonging to C_k:

$$W_k = \sum_{\text{neighbors of class } C_k} w^i$$

Finally, assign X to the class with maximum total weight:

Most supported disease label $\hat{y} = \arg\max_k W_k$

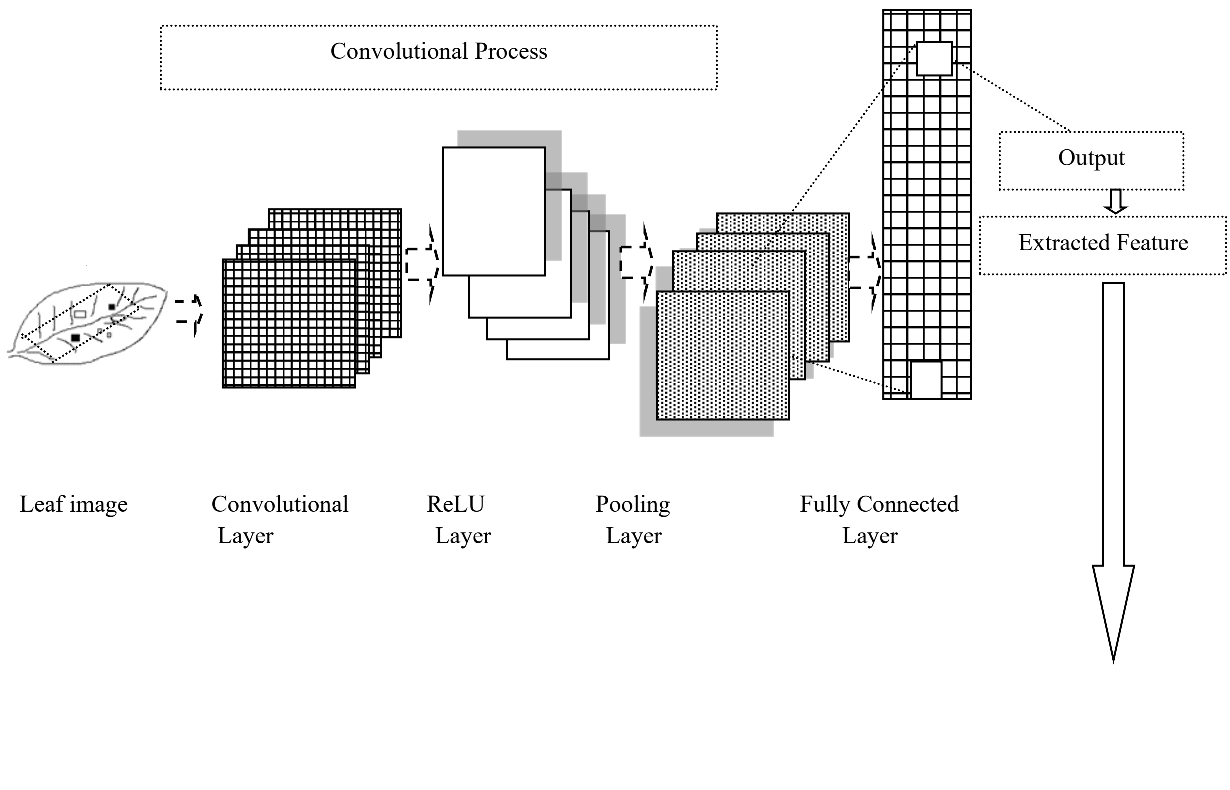
Then Closer samples contribute more to final decision are obtaining for:

“Early Blight": ((10, 100, 100), (20, 255, 255)),

"Powdery Mildew": ((0, 0, 200), (255, 255, 255)),

"Bacterial Leaf Spot": ((30, 100, 100), (90, 255, 255))))

These samples are feasible for the same diseases are examined and found that to process in next levels [19][20][21][22]. Computer vision–based digital imaging technology represents a modern approach to plant disease detection and plays a crucial role in minimizing crop yield losses while improving crop quality. By capturing and analyzing visual features of leaves, stems, and fruits, these systems enable early and non-invasive identification of pathogens, supporting timely intervention and more efficient use of agrochemicals. Patterns and color-based disease detection methods offer higher precision and are especially effective in single-sample–based analysis, where localized symptoms such as spots, discoloration, and necrotic regions can be reliably segmented and classified. i.e. processed on the algorithm of SSCVRA (Single sample computer vision recognition algorithm) [18]. That is filtered by weighted KNN [19] is a recent advanced in machine learning algorithm, that have further enhanced the accuracy and robustness of image- based models, making computer vision a key enabler of smart and sustainable agriculture. That is the major application have been distributed for a single sample leaf disease detection.



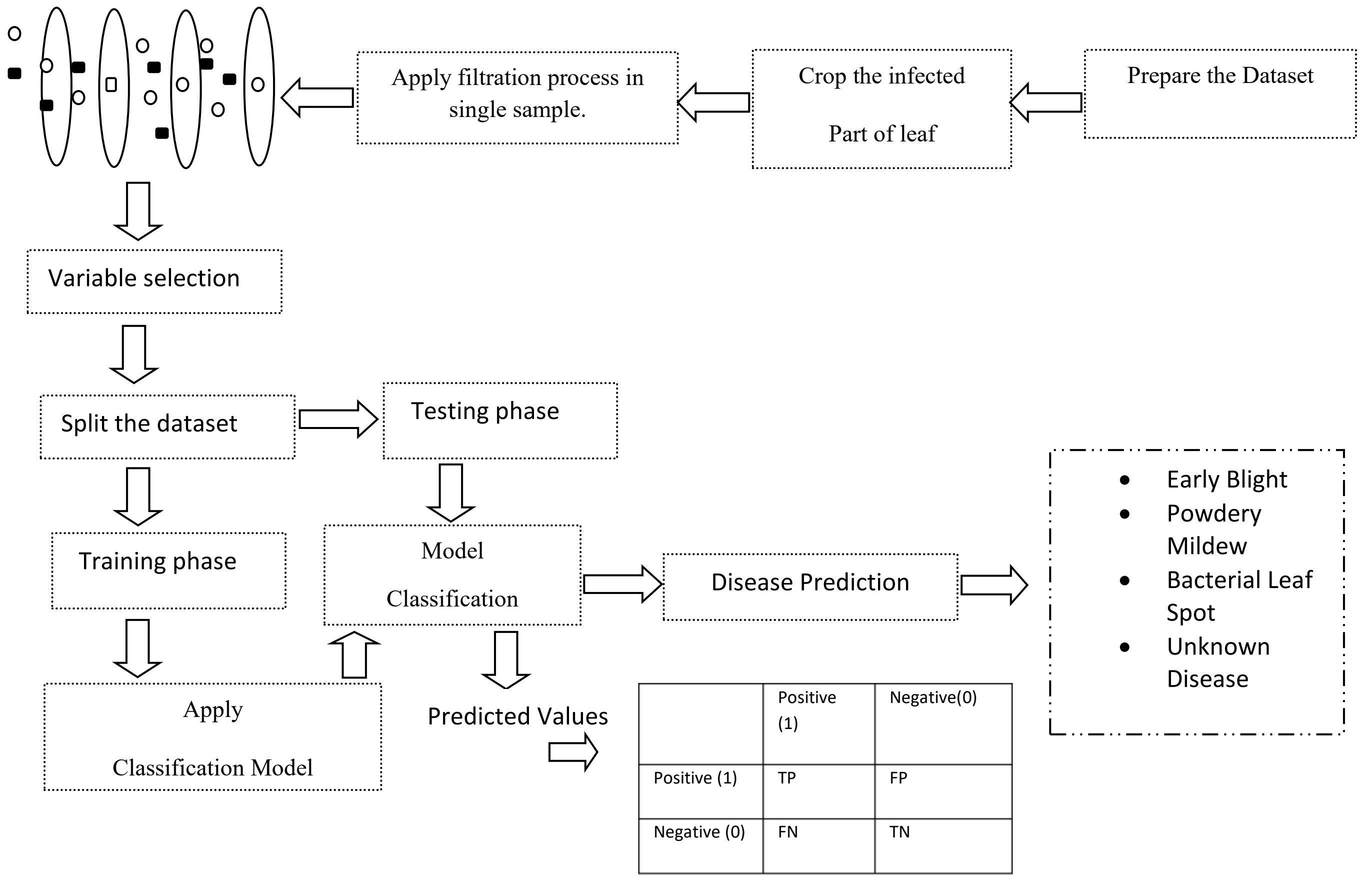


Fig. 15 Flow chart of Weighted KNN based leaf disease classification.

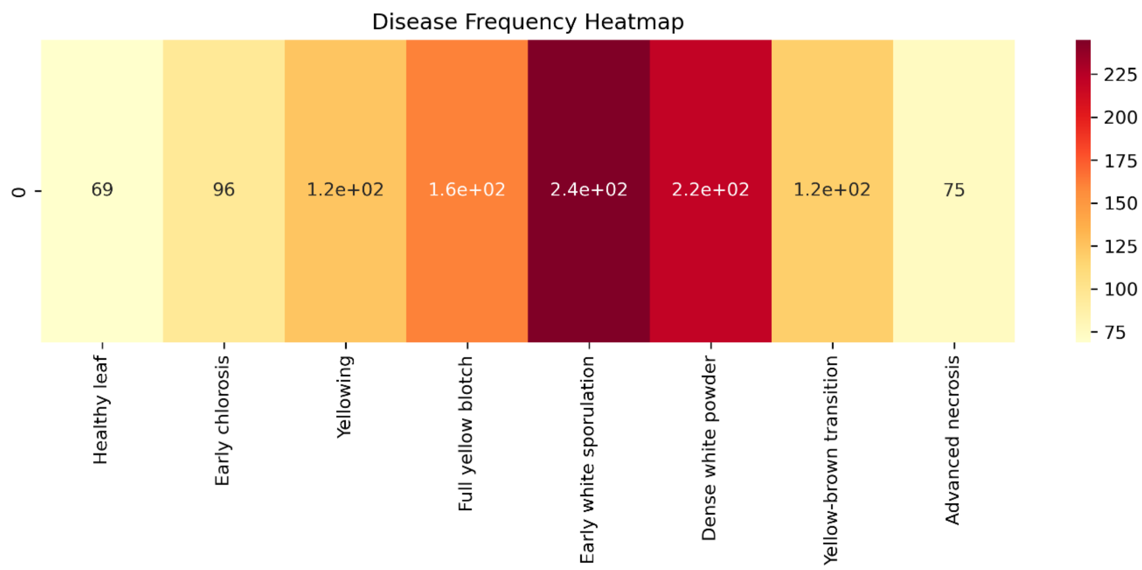


Fig. 16: Disease Frequency Heat map

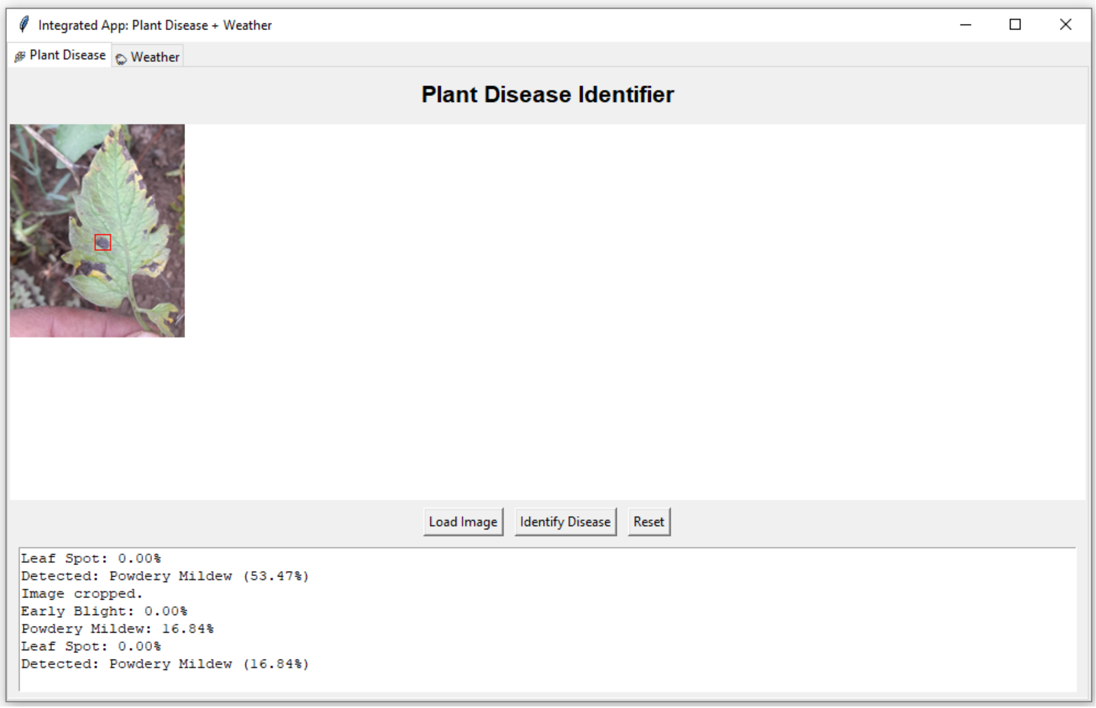


Fig. 17: Application of Plant disease detection

As illustrated in the Disease frequency heat map presented in Fig. 16, the disease frequency distribution in the present dataset represents the progressive symptom development associated with powdery mildew infection. The collected samples were categorized into eight distinct classes: Healthy Leaf (69), Early Chlorosis (96), Yellowing (124), Full Yellow Blotch (161), Early White Sporulation (245), Dense White Powder (220), Yellow-Brown Transition (119), and Advanced Necrosis (75). The highest sample frequencies were observed in the Early White Sporulation and Dense White Powder stages, indicating a greater representation of actively infected samples with visible fungal colonization. In comparison, the Healthy Leaf and Advanced Necrosis categories contained fewer samples.

The observed progression of symptoms, beginning with chlorosis and yellowing and advancing toward white fungal growth and necrotic tissue development, agrees with previously reported disease progression patterns in powdery mildew path systems affecting grapevine and cucurbit crops [34], [35]. The frequency distribution presented in this work reflects the composition of the experimental dataset and was utilized for data analysis, visualization, and machine learning model development. The implemented workflow and corresponding processing stages are presented in Fig. 15.

6. CONCLUSIONS

The proposed Enhanced Intelligent Information System (EIIS) integrates Convolutional Neural Networks (CNN) for automated feature extraction, Weighted K-Nearest Neighbors (WKNN) for disease classification, and an Artificial Neural Network (ANN) for usability score prediction, forming an efficient and intelligent framework for plant leaf disease detection. The developed CNN-WKNN model successfully analyzed potato leaf images and achieved a classification accuracy of 98.75% for the identification of major diseases, including Early Blight, Powdery Mildew, and Bacterial Leaf Spot. The incorporation of image preprocessing, optimized feature extraction, and weighted classification strategies significantly enhanced the accuracy, robustness, and reliability of the proposed disease identification system. Furthermore, environmental factors such as temperature, humidity, and weather conditions serve as important parameters that influence pathogen activity and create favorable conditions for disease growth and progression; therefore, their integration can further improve future disease forecasting capabilities.

The ANN-based usability prediction model demonstrated effective performance, where both training and validation losses decreased significantly and stabilized after approximately 60 epochs, indicating efficient learning and model convergence. The obtained validation Mean Absolute Error (MAE) of approximately 0.90 and Root Mean Square Error (RMSE) of nearly 1.00 demonstrate the model's capability to accurately estimate usability scores with satisfactory prediction performance and generalization ability. The developed Tkinter-based graphical user interface further improved system accessibility by providing a user-friendly platform for farmers, researchers, and agricultural experts.

Overall, the proposed EIIS provides a reliable decision-support framework for early plant disease identification, reducing potential crop losses and supporting effective disease management practices. The integration of intelligent image-based diagnosis with environmental parameters highlights its potential application in precision agriculture, smart farming, and sustainable crop production.

7. FUTURE SCOPE

The proposed EIIS can be extended to support multiple crops and diseases, integrate advanced deep learning models, enable real-time mobile and web-based diagnosis, incorporate IoT sensors and drone-based monitoring, assess disease severity, provide automated treatment recommendations, utilize explainable AI for transparent decision-making, leverage cloud computing for large-scale deployment, adopt reinforcement learning for adaptive disease management, combine multimodal agricultural data, and contribute to precision agriculture, run on weather logger and smart farming systems for improved productivity and sustainability.

8. ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my research supervisor, Dr. Pramod Singh, Department of Computer Science, AKS University, Satna, Madhya Pradesh, for his continuous guidance, valuable suggestions, and unwavering support throughout the course of my research work. His insightful advice, constructive feedback, patience, and encouragement have significantly contributed to the successful completion and improvement of this research.

I am deeply thankful to AKS University and the Department of Computer Science for providing the required facilities, resources, and a supportive academic environment for carrying out this research. I also extend my sincere appreciation to all faculty members and staff for their assistance, cooperation, and valuable contributions. I would like to acknowledge my family and friends for their constant motivation, encouragement, and moral support throughout my research journey. I am also grateful to my colleagues and peers for their helpful discussions and constructive suggestions. Finally, I sincerely acknowledge all individuals and organizations whose direct or indirect support contributed to the successful completion of this research work.

9. REFERENCES

[1] Zhaoyu Zhai, José Fernán Martínez, Victoria Beltran, and Néstor Lucas Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Computers and Electronics in Agriculture*, vol. 170, p. 105256, 2020, doi: 10.1016/j.compag.2020.105256.

[2] Utkarsh Yashwant Tambe, A. Shobanadevi, A. Shanthini, and Hsiu-Chun Hsu, "Potato Leaf Disease Classification Using Deep Learning: A Convolutional Neural Network Approach," arXiv preprint arXiv:2311.02338, 2023.

[3] S. Patil and M. Sasikala, "Segmentation and identification of medicinal plant through weighted KNN," Multimedia Tools and Applications, vol. 82, no. 2, pp. 2805–2819, 2023, doi: 10.1007/s11042-022-13201-7.

[4] A. Iglesias, K. Arai, and S. Kapoor, Eds., Information Systems for Intelligent Systems: Proceedings of ISBM 2025, vol. 10. Cham, Switzerland: Springer, 2026.

[5] A. S. A. Osman, S. H. Zahra, and A. Alotaibi, "An intelligent information system for plant disease detection using machine learning and image processing," Indian Journal of Science and Technology, vol. 18, no. 45, pp. 3582–3594, 2025, doi: 10.17485/IJST/v18i45.1777.

[6] G. S. Krishna, Z. Gulzar, A. Baronia, and J. Srinivas, "AI-Enabled Intelligent System for Automatic Detection and Classification of Plant Diseases towards Precision Agriculture," Informatics, vol. 12, no. 4, p. 138, 2025. doi: 10.3390/informatics12040138.

[7] A. González-Briones et al., "Enhancing Plant Disease Detection: Incorporating Advanced CNN Architectures for Better Accuracy and Interpretability," International Journal of Computational Intelligence Systems, vol. 18, p. 120, 2025. doi: 10.1007/s44196-025-00835-2.

[8] N. V. L. Rajarajeswari and K. Muralidharan, "Assessments of farm yield and district production loss from bacterial leaf blight epidemics in rice," Crop Protection, vol. 25, no. 3, pp. 244–252, 2006, doi: 10.1016/j.cropro.2005.04.013.

[9] A. Bijlwan, R. Ranjan, M. Singh, et al., "Predicting crop disease severity using real time weather variability through machine learning algorithms," Scientific Reports, vol. 15, p. 34767, 2025. doi: 10.1038/s41598-025-18613-7.

[10] A. R., G. Sunkad, R. B. T., M. K. Yadav, Y. S. T., R. H. N., B. P. R., and P. D., "Severity and distribution of bacterial leaf blight of rice in different rice growing ecosystems of Karnataka state of India," ORYZA – An International Journal of Rice, vol. 60, no. 2, pp. 297–303, 2023. Available: https://epubs.icar.org.in/index.php/OIJR/article/view/138492 .

[11] A. R. Deepa, "An Efficient Deep Learning Method Using a Hybrid Model for Plant Leaf Disease Classification," in *Proceedings of the Advances in Intelligent Systems and Computing*, Springer, 2026, doi:10.1007/978-981-96-8107-5_27.

[12] M. H. Saad and A. E. Salman, "A plant disease classification using one-shot learning technique with field images," *Multimedia Tools and Applications*, vol. 83, pp. 58935–58960, 2024, doi: 10.1007/s11042-023-17830-4.

[13] M. A. O. Balendres, "Diaporthe blight in eggplant (Solanum melongena L.): etiology, epidemiology, and current management," *International Journal of Vegetable Science*, vol. 31, no. 5, pp. 768–793, 2025.

[14] KALE, M. R., & SHITOLE, M. S. (2021). Analysis of crop disease detection with SVM, KNN and random forest classification. Information Technology in Industry, 9(1), 364-372.

[15] D. Wang et al., "How Much Data are Enough? Investigating Dataset Requirements for Patch-Based Brain MRI Segmentation Tasks," Medical Image Anal., 2024. [Online]. Available: https://arxiv.org/abs/2404.03451.

[16] Yadav, M. K., Manohar, D. D., Mukherjee, G., & Chakraborty, C. (2013). Segmentation of chronic wound areas by clustering techniques using selected color space. Journal of Medical Imaging and Health Informatics, 3(1), 22–29.

[17] Paneru, B., Paneru, B., & Shah, K. B. (2024). Analysis of convolutional neural network-based image classifications: A multi-featured application for rice leaf disease prediction and recommendations for farmers. arXiv. https://doi.org/10.48550/arXiv.2410.01827.

[18] Hand, D.J. Principles of Data Mining. Drug-Safety 30, 621–622 (2007). https://doi.org/10.2165/00002018-200730070-00010.

[19] U. Shruthi, V. Nagaveni and B. K. Raghavendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 281-284, doi: 10.1109/ICACCS.2019.8728415.

[20] Emmi, L., Gonzalez-de-Soto, M., Gonzalez-de-Santos, P. (2014). Configuring a Fleet of Ground Robots for

Agricultural Tasks. In: Armada, M., Sanfeliu, A., Ferre, M. (eds) ROBOT2013: First Iberian Robotics Conference. Advances in Intelligent Systems and Computing, vol 252. Springer, Cham. https://doi.org/10.1007/978-3-319-03413-3_37.

[21] Pena, J. M., Torres-Sánchez, J., de Castro, A. I., Kelly, M., & López-Granados, F. (2015). Quantifying efficacy and limits of unmanned aerial vehicle (UAV) technology for weed seedling detection as affected by sensor resolution. Sensors, 15(3), 5609–5626. https://doi.org/10.3390/s150305609.

[22] Samsung. (n.d.). Samsung Galaxy M12 – 64 GB, Green (SM-M127FZGGMEA). Samsung Business Gulf. Retrieved January 31, 2026, from https://www.samsung.com/ae/business/smartphones/galaxy-m/galaxy-m12-green-64gb-sm-m127fzggmea/

[23] Shen Weizheng, Wu Yachun, and colleagues (2008) proposed a method for grading leaf spot disease using image processing techniques, presented in Wuhan, China (pp. 491–494).

[24] Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN model-based approach in classification. In R. Meersman, Z. Tari, & D. C. Schmidt (Eds.), On the Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM 2003 (Lecture Notes in Computer Science, Vol. 2888, pp. 986–996). Springer. https://doi.org/10.1007/978-3-540-39964-3_62.

[25] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," *Computers and Electronics in Agriculture*, vol. 72, no. 1, pp. 1–13, 2010. doi: 10.1016/j.compag.2010.02.007

[26] M. Sugawara, S.-Y. Choi, and D. Wood, “Ultra-High-Definition Television (Rec. ITU-R BT.2020): A Generational Leap in the Evolution of Television [Standards in a Nutshell],” IEEE Signal Processing Magazine, vol. 31, no. 3, pp. 170–174, May 2014, doi:10.1109/MSP.2014.2302331.

[27] Wei, Q., Li, D., Liu, C., Liu, B., Liu, Z. (2011). Converting Digital Image to SVG for User Interaction. In: Zeng, D. (eds) Applied Informatics and Communication. ICAIC 2011. Communications in Computer and Information Science, vol 224. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23214-5_50.

[28] M. M. Mishra and P. Singh, “Pattern Based Leaves Disease Classification Using AI,” International Journal of Latest Technology in Engineering, Management & Applied Science, vol. 14, no. 6, pp. 908–914, 2025, doi: 10.51583/IJLTEMAS.2025.1406000100.

[29] E. Hossain, M. F. Hossain, and M. A. Rahaman, “A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier,” in Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox’s Bazar, Bangladesh, 2019, pp. 1–6, doi: 10.1109/ECACE.2019.8679247.

[30] Hayit, T., Endes, A. & Hayit, F. KNN-based approach for the classification of fusarium wilt disease in chickpea based on color and texture features. Eur J Plant Pathol 168, 665–681 (2024). https://doi.org/10.1007/s10658-023-02791-z

[31] M. Li *et al*., “Wheat Powdery Mildew Severity Classification Based on an Improved ResNet34 Model,” *Agriculture*, vol. 15, no. 15, p. 1580, Jul. 23, 2025, doi: 10.3390/agriculture15151580.

[32] D. K. D. Bourzig, A. Mansour, and M. Mus, “Powdery Mildew Disease Classification in Laboratory and Real-Field Images using Convolutional Neural Networks for Precision Agriculture,” in *2024 1st International Conference on Innovative and Intelligent Information Technologies (IC3IT)*, 2024, pp. -----, doi: 10.1109/IC3IT63743.2024.10869354.

[33] B. Sowmya and S. Guruprasad, “Deep learning based plant health disease detection in tomatoes using Inception v4 Convolutional neural network and YOLO v8,” *Artificial Intelligence in Agriculture*, vol. 5, art. 278, Oct. 22, 2025, doi: **10.1007/s44163-025-00540-1*

[34] PÉREZ-GARCÍA, A., ROMERO, D., FERNÁNDEZ-ORTUÑO, D., LÓPEZ-RUIZ, F., DE VICENTE, A. and TORÉS, J.A. (2009), The powdery mildew fungus Podosphaera fusca (synonym Podosphaera xanthii), a constant threat to cucurbits. Molecular Plant Pathology, 10: 153-160. https://doi.org/10.1111/j.1364-3703.2008.00527.x

[35] Gadoury, D. M., Cadle-Davidson, L. A. N. C. E., Wilcox, W. F., Dry, I. B., Seem, R. C., & Milgroom, M. G. (2012). Grapevine powdery mildew (Erysiphe necator): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. Molecular plant pathology, 13(1), 1-16.