

1-Pumpkin Leaf disease detection -Adeline.docx

by Adeline Sneha John Chrisastum

Submission date: 19-Apr-2026 12:45AM (UTC+0800)

Submission ID: 2935912018

File name: 58289_Adeline_Sneha_John_Christastum_1-Pumpkin_Leaf_disease_detection_-
Adeline_1175895_141602400.docx (3.28M)

Word count: 6029

Character count: 37852

Deep Learning Based CNN Architecture for Early Diagnosis of Pumpkin Leaf Diseases in Precision Agriculture

Adeline Sneha J

School of Computing
Asia Pacific University of Technology and Innovation
Kuala Lumpur, Malaysia
adeline.sneha@apu.edu.my

Arun Seeralan Balakrishnan

School of Mechatronics Engineering
Asia Pacific University of Technology and Innovation
Kuala Lumpur, Malaysia
arun@apu.edu.my

Maythem K. Abbas

Department of Electrical & Electronics Engineering
Universiti Teknologi PETRONAS
Perak, Malaysia
maythem.abbas@utp.edu.my

Roxanne M. Fabriquel

College of Agriculture and Allied Fields
Mindoro State University
Mindoro, Philippines
roxanne.fabriquel@minsu.edu.ph

Randy A. Joco

Director for Innovation, MinSU-iBIBES
Mindoro State University
Mindoro, Philippines
randy.joco@minsu.edu.ph

Jerrel S. Reyes, PhD

Department of Agricultural and Biosystems Engineering
Mindoro State University
Mindoro, Philippines
jerrel.reyes@minsu.edu.ph

Rina May A. Joco

Project Assistant II, FARMIQ Project
Mindoro State University
Mindoro, Philippines
rina.joco@minsu.edu.ph

Rajasree S

Department of Information Technology
Coimbatore Institute of Technology
Coimbatore, India
rajasree.s@cit.edu.in

Amit Kumar

Department of Electronics and Communication Engineering
Chitkara University
Punjab, India
amit.kumar@chitkara.edu.in

Shivani Malhotra

Department of Electronics and Communication Engineering
Chitkara University
Punjab, India
shivani.malhotra@chitkara.edu.in

Isha Gupta

Department of Electronics and Communication Engineering
Chitkara University
Punjab, India
isha.gupta@chitkara.edu.in

Abstract

Pumpkin leaf diseases pose a serious threat to agricultural productivity, causing significant yield losses. Early detection and effective disease management are therefore crucial to minimizing damage. Traditional methods rely on manual inspection, which is time-consuming, labour intensive, and prone to human error. To address these limitations, this study developed a Convolutional Neural Network (CNN)-based model for efficient diagnosis of pumpkin leaf diseases. By applying preprocessing techniques such as noise reduction, edge detection, and image enhancement, along with data augmentation and normalization, the model achieved an accuracy of 87.99%. The trained CNN was integrated into a Flask-based web application, enabling real-time disease identification through image uploads. This demonstrates the potential of CNNs as powerful tools in early diagnosis and disease management, supporting farmers in timely interventions to protect crop yield. Future improvements aim to broaden the range of detectable diseases, optimize CNN architecture, and enhance prediction accuracy, strengthening the role of such models in sustainable agriculture.

Keywords: Data augmentation, normalization, model integration, agricultural automation, plant disease detection, machine learning, computer vision.

I. INTRODUCTION

In recent decades, the fields of image processing and image analysis have undergone substantial technological progress, particularly in their practical application to agriculture. These advances have been instrumental in domains such as crop yield prediction, early pest detection, and plant health monitoring [1]. Within this broad agricultural spectrum, the pumpkin stands out as a crop of notable global significance. Its importance stems not only from its high nutritional value and economic benefits but also from the crucial role it plays in enhancing food security for many regions. For example, in countries such as Ethiopia, the pumpkin is regarded as an essential agricultural product, and the success or failure of its cultivation can directly influence national objectives aimed at achieving food self-sufficiency [1].

One of the most pressing challenges in pumpkin cultivation is the accurate and timely identification of diseases that can attack the plant. Disease detection is vital because it serves as the first line of defense in protecting both crop yield and quality. If such diseases are not recognized and managed in their early stages, they can lead to significant production losses and long-term economic setbacks for farmers. Unfortunately, in real farming environments, visual inspection of pumpkin leaves is often unreliable; symptoms may be subtle, environmental conditions may obscure visual cues, and the process is highly susceptible to human error, even for experienced farmers or agricultural specialists.

This research is designed to confront these challenges head-on by applying advanced image processing techniques to the task of classifying pumpkin leaf diseases. In particular, the study focuses on five distinct categories of leaf conditions: one category representing healthy leaves and four categories representing diseased leaves. The diseased categories include Bacterial Leaf Spot, Downy Mildew, Mosaic Disease, and Powdery Mildew. By creating a robust, efficient, and practical classification model, the aim is to provide farmers, agronomists, and other agricultural practitioners with a reliable decision-support tool. Such a tool will help them to diagnose diseases accurately and quickly, thereby reducing dependence on subjective manual inspection and enabling faster intervention measures for disease management.

The remainder of this paper is organized as follows. The subsequent sections describe the methodological framework used in designing the proposed model, explain the steps taken in its practical implementation, present the evaluation results in detail, and, finally, offer a discussion and analysis of the overall performance outcomes of the study.

II. RELATED WORK

Agriculture was the backbone of global food security. The humanity's future was intrinsically linked to its advancement. However, stagnant cultivation techniques have increased the devastating effects of plant diseases in recent years. Traditional disease diagnosis relies heavily on farmer subjective judgment and experiential knowledge. This method results in inconsistent and frequently inaccurate assessments [2].

Two technological pillars have emerged as essential for modern image recognition which includes digital image processing and machine learning [6]. Digital image processing transforms raw images via contrast enhancement, brightness optimization, noise reduction, and sharpness refinement. These improvements provides an invaluable in applications ranging from medical diagnostics to satellite surveillance and forensic analysis. Meanwhile, machine learning models empowers the computers to autonomously categorize images by detecting learned patterns based on the previous literatures. Convolutional Neural Networks (CNN) exemplify this capability. When an CNN was trained on extensive labeled datasets these algorithms master the recognition of visual elements. Those elements includes an shapes, textures, and colors. Through progressive learning of hierarchical data representations these models effectively differentiate between diverse categories from everyday objects to specific plant pathologies.

Deep learning models has revolutionized plant disease detection in recent years. Unlike traditional methods that demand specialized botanical expertise, deep learning technologies autonomously extract and classify disease characteristics directly from images [3]. This automation helps to identify and eliminates numerous time-consuming manual processes that affects the conventional recognition techniques critically. Added to above, the CNN have emerged as the dominant approach within deep learning due to their superior performance in image analysis and classification tasks.

Recent research demonstrates this technological potential, where Nirmala and Gomathy [7] developed a CNN-based hybrid framework for identifying diseases in pumpkin leaves, meanwhile Soni et al. [1] created a crop disease classification system that achieved an impressive 98.27% accuracy rate across 406 test images.

Agriculture plays a central role in ensuring global food security, but crop diseases cause significant yield losses and remain a serious challenge to productivity. Traditional plant disease diagnosis is mostly visual and depends on the experience of farmers, making it time-consuming, laborintensive, and often inaccurate. Recent reviews emphasize that such limitations highlight the importance of automated, image-based diagnosis systems [9].

Two major technological foundations support automated recognition: digital image processing and machine learning. Digital image processing enhances raw images by reducing noise, adjusting contrast and

brightness, and improving sharpness. These steps are widely applied as preprocessing stages in plant-disease studies and remain essential in preparing datasets for deep learning [10]. At the same time, advances in machine learning, particularly convolutional neural networks (CNNs), have enabled computers to automatically learn features such as shape, texture, and color directly from images. Unlike handcrafted methods, CNNs allow end-to-end training and have shown excellent performance in complex image classification tasks [11].

In recent years, CNN-based models have transformed plant disease detection. When trained or fine-tuned on large, labeled datasets, CNNs consistently outperform traditional approaches while also simplifying deployment on mobile or edge devices. Several studies and surveys report high accuracy levels using transfer learning, data augmentation, and optimized architectures, while also noting challenges such as unbalanced datasets and variability in natural field conditions [12], [13].

Crop-specific research has shown promising results. Agarwal et al. developed ToLeD, a CNN framework that achieved robust accuracy for tomato leaf diseases [14]. For pumpkin (*Cucurbita*), research remains limited but is expanding. Lin et al. applied image processing, principal component analysis (PCA), and classical machine learning to identify pumpkin powdery mildew, providing an early baseline [15]. Nirmala and Gomathy introduced a hybrid CNN model combining feature denoising and CNN layers for *Cucurbita* disease classification, reporting improved detection rates compared to simpler models [16]. More recent work, such as hybrid or ensemble CNNs that fuse ResNet and LeNet features, has shown strong accuracies for pumpkin disease detection, reinforcing CNN effectiveness for this crop [17].

Studies on deeper networks and transfer learning further demonstrate potential. Residual and dense architectures, such as ResNet and EfficientNet, have achieved state-of-the-art performance across multi-crop datasets. Pandian et al. highlighted that deep backbones trained with strong augmentation strategies can significantly improve accuracy, although they also increase computational requirements [18].

Overall, prior studies illustrate a clear trend from classical methods toward deep and hybrid CNN approaches, with increasing focus on explainability and robustness. However, pumpkin and *Cucurbita* species remain underrepresented compared to major crops like rice, maize, and tomato. Open challenges include dataset diversity (lab versus field images), early-stage symptom detection, lightweight deployment models, and interpretability. These gaps motivate the present study, which applies CNNs to pumpkin leaf disease detection with preprocessing, augmentation, and real-time deployment to support precision agriculture.

III. MODEL DESIGN

A. System Overview

The development of this pumpkin leaf disease classification application follows a systematic pipeline contains the four core phases which includes the image acquisition and preprocessing, model architecture design, training implementation, and application deployment using the finalized model. Figure 1 illustrates the complete project workflow.

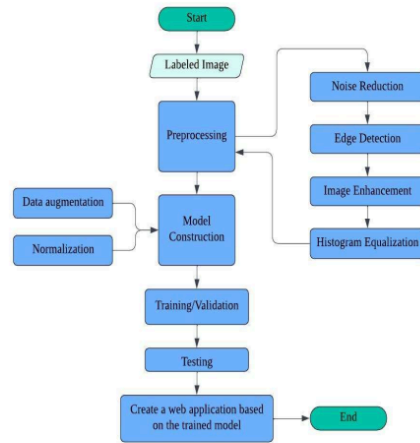


Fig.1. Flow Chart of the Model Design Preprocessing

Image preprocessing serves as the foundation for successful computer vision and image analysis applications. It helps to transform the raw images into optimized formats which was suitable for downstream processing. This critical phase employs various techniques such as noise suppression, normalization, and edge enhancement. The above techniques used to eliminate objects, amplify relevant features, and standardize input data. Such preprocessing data enables algorithms and models to extract more and more meaningful patterns efficiently and reliably than conventional approach. This eventually results in the improvements in model accuracy, reduced computational demands, and enhanced feature extraction make preprocessing indispensable for classification, detection, and recognition systems.

1) **Noise Reduction:**

Noise reduction techniques help to enhance the image quality by filtering out unwanted variations introduced by inadequate lighting conditions, sensor limitations, or data transmission errors. These methods are ideally applied by sophisticated smoothing algorithms which suppress the random fluctuations by maintaining essential structural information and fine details. This balance between noise elimination and detail preservation proves particularly valuable in demanding applications such as medical diagnostics, remote sensing analysis, and professional photography where image clarity was more important and directly impacts decision-making and interpretation accuracy.

2) **Edge Detection:**

Edge detection is used to extract boundaries and structural outlines of objects in the image. This technique highlights areas with significant intensity change, which is vital for segmenting diseased regions on leaves. Accurate edge information enhances feature extraction and helps the CNN model to focus on patterns that differentiate one disease from another.

3) **Image Enhancement:**

Image enhancement techniques improve visual clarity by adjusting parameters such as brightness, contrast, and sharpness. Enhanced images allow better visibility of disease symptoms, aiding both manual and automated diagnosis. This step is especially useful in conditions where disease symptoms are subtle or obscured by natural variations in leaf appearance.

4) **Histogram Equalization:**

Histogram equalization is used to improve image contrast by redistributing the intensity values more evenly across the image. This adjustment allows finer details in underexposed or low-contrast regions to be more visible. It is particularly helpful in scenarios such as detecting early-stage symptoms or distinguishing between diseases with similar color features.

5) **Normalization:**

Normalization ensures that all pixel values are scaled to a common range, typically between 0 and 1. This standardization facilitates faster and more stable training of deep learning models by preventing one feature from disproportionately influencing the model due to scale differences. Normalization also aids in generalization, especially when input images vary in lighting or intensity.

6) **Data Augmentation**

Collecting and labeling a large dataset of pumpkin leaf images is resource-intensive. Additionally, deep learning models are prone to overfitting when trained on small or unbalanced datasets. To

address this, data augmentation techniques such as image rotation, horizontal flipping, and vertical flipping are employed. These augmentations simulate real-world variations in leaf orientation and appearance, allowing the model to become invariant to such changes and improving its generalization capabilities. As highlighted by Li et al. [3], these techniques are critical in improving recognition accuracy in plant disease classification tasks.

B. Model Construction

Deep learning has seen remarkable advancements in the field of agricultural research, primarily due to its ability to automatically extract complex features from data [5]. When compared to traditional image processing methods, deep learning approaches demonstrate significantly higher accuracy and generalization capability [4]. Conventional machine learning techniques such as SVM and KNN rely heavily on manual feature extraction, making them dependent on prior domain knowledge and limiting their ability to capture complex visual patterns. In contrast, Convolutional Neural Networks (CNNs) automatically learn hierarchical features—such as edges, textures, and shapes—directly from image data, similar to how humans visually interpret objects. While advanced architectures like ResNet and EfficientNet offer improved performance through deeper or more optimized designs, they require higher computational resources and larger datasets. Therefore, CNN provides a balanced approach by delivering effective feature learning with relatively lower complexity, making it a suitable and practical choice for image-based applications such as fruit quality assessment.

Among these, Convolutional Neural Networks (CNNs) have emerged as a particularly effective model for image classification tasks, including the identification of plant diseases.

CNNs are capable of learning hierarchical feature representations directly from images, eliminating the need for manual feature engineering. This property makes them especially suitable for detecting subtle symptoms such as leaf discoloration, lesions, and spot patterns, which are often indicative of specific plant diseases. Given these advantages, this study adopts a CNN-based approach to classify various pumpkin leaf conditions and effectively distinguish between healthy and diseased leaves.

C. Model Integration

To facilitate user interaction with the trained classification model, this study incorporates Flask , a lightweight, flexible web framework written in Python. Flask is well-suited for rapid development and deployment of web applications, providing a minimal yet extensible core that supports routing, request handling, session management, and integration with databases.

In this project, Flask serves as the interface between users and the CNN model. It enables the deployment of a web-based application that allows users to upload images of pumpkin leaves and receive real-time

disease classification results. The simplicity and extensibility of Flask make it an ideal choice for integrating the trained deep learning model into a functional and user-friendly web application.

IV. IMPLEMENTATION

A. Image Acquisition

Kaggle serves as a comprehensive repository for diverse datasets which contains multiple research domains from plant pathology to sentiment analysis and social media analytics. The platform datasets benefit from rigorous data curation from the prestigious organizations including the WHO, World Bank, NASA, and leading research institutions, making them ideal for academic and scientific investigations. This study utilizes the Pumpkin Leaf Diseases Dataset from Kaggle, comprising 2,000 high-resolution images (approximately 600×800 pixels) of pumpkin foliage. The dataset helps to make an robust analysis on the disease identification and classification, distinguishing between pathological conditions and healthy specimens. The collection maintains perfect class balance with 400 images across five categories such as bacterial Leaf Spot, downy mildew, healthy leaves, mosaic disease, and powdery mildew. Professor Dr. Mohammad Rifat Ahmmad Rashid, Joy Biswas, and Md Miskat Hossain created and retain rights to this 2024 dataset.

B. Image Preprocessing (Dataset-specific)

Following the methodology outlined in Section III, comprehensive preprocessing precedes model development. The Pumpkin Leaf Diseases Dataset undergoes systematic transformation before training. Figure 2 displays representative samples from the healthy leaf category in their original form.

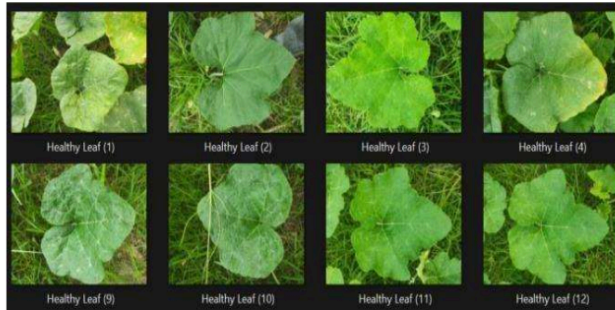


Fig.2. Original images of healthy leaf

The preprocessing pipeline encompasses the noise suppression, edge detection, image enhancement, and histogram equalization. The workflow organizes original data in the "Pumpkin_Leaf_Diseases_Dataset"

directory. It processed outputs reside in "Preprocessed_Dataset". Figure 3 demonstrates the transformation results.



Fig.3. Examples of pre-processed images

Additional preprocessing operations includes an data augmentation, normalization, and dimensional standardization. These are integrated directly into the classification model architecture.

C. Model Construction (TensorFlow specifics)

The classification system employs a CNN implemented through TensorFlow. CNN excel at visual data processing by autonomously identifying critical features. It also uses the edges, textures, and complex patterns through hierarchical learning. TensorFlow comprehensive framework facilitates

CNN development through modular components convolutional, pooling, and fully connected layers. Meanwhile native GPU acceleration enhances training efficiency for large-scale datasets, proving invaluable for applications spanning medical imaging, autonomous navigation, and agricultural disease detection.

```
Found 1950 files belonging to 5 classes.  
  
class_names = dataset.class_names  
print(class_names)  
  
['Bacterial Leaf Spot', 'Downy Mildew', 'Healthy Leaf', 'Mosaic Disease', 'Powdery Mildew']
```

Fig.4. Examples of pre-processed images

The TensorFlow implementation processes 1,950 pumpkin leaf images distributed across training, validation, and testing sets. The model distinguishes between five conditions based on the Bacterial Leaf Spot, Downy Mildew, Healthy Leaves, Mosaic Disease, and Powdery Mildew (Fig. 4).

The CNN architecture captures disease-specific visual signatures including lesions, discoloration patterns, and morphological abnormalities. Convolutional layers detect these features progressively, pooling operations retain salient information while eliminating redundancy, and dense layers synthesize learned representations for final classification. Non-linear activation functions enable the network to model complex manifestations including irregular disease boundaries and subtle color variations.

Figure 5 details the complete model architecture:

- 1) Input Layer: Processes images with dimensions (600, 800, 3)
 - 2) Preprocessing Pipeline:
 - Resize and Rescale: Normalizes input dimensions and pixel values
 - Data Augmentation: Implements random transformations (flips, rotations) to enhance model generalization
 - 3) Feature Extraction: Five Conv2D layers employ ReLU activation with 3×3 kernels. Each convolution pairs with MaxPooling2D (2×2 window) for progressive spatial reduction and computational efficiency
 - 4) Classification Head: Flatten operation transforms multidimensional features into vectors, feeding a 64-unit Dense layer with ReLU activation
 - 5) Output Layer: Five-unit Dense layer with softmax activation generates class probability distributions
 - 6) Training Configuration: Adam optimizer minimizes sparse categorical crossentropy loss while tracking accuracy metrics
- 1) *Input Layer*: Accepts images with a shape of (600, 800, 3).
 - 2) *Preprocessing Layers*:
 - a) *Resize and Rescale*: Standardizes input images for further processing.
 - b) *Data Augmentation*: Applies random flips and rotations to increase dataset diversity and improve robustness.
 - 3) *Convolution and Pooling Layers*: Five Conv2D layers with ReLU activation and a kernel size of (3×3) are applied. Each convolutional layer was followed by a MaxPooling2D layer with a pool size of (2×2) to reduce spatial dimensions and computational complexity.
 - 4) *Flattening and Dense Layers*: The Flatten layer converts feature maps into a 1D vector, followed by a Dense layer with 64 units and ReLU activation.
 - 5) *Output Layer*: A Dense layer with 5 units (corresponding to the output classes) and a softmax activation produces a probability distribution over the classes.

6) *Compilation:* The model is compiled using the Adam optimizer. It used with sparse categorical crossentropy as the loss function and accuracy as the evaluation metric.

Layer (type)	Output Shape	Param #
sequential (Sequential)	(None, 300, 400, 3)	0
sequential_1 (Sequential)	(None, 300, 400, 3)	0
conv2d (Conv2D)	(None, 298, 398, 32)	896
max_pooling2d (MaxPooling2D)	(None, 149, 199, 32)	0
conv2d_1 (Conv2D)	(None, 147, 197, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 73, 98, 64)	0
conv2d_2 (Conv2D)	(None, 71, 96, 64)	36,928
max_pooling2d_2 (MaxPooling2D)	(None, 35, 48, 64)	0
conv2d_3 (Conv2D)	(None, 33, 46, 64)	36,928
max_pooling2d_3 (MaxPooling2D)	(None, 16, 23, 64)	0
conv2d_4 (Conv2D)	(None, 14, 21, 64)	36,928
max_pooling2d_4 (MaxPooling2D)	(None, 7, 10, 64)	0
flatten (Flatten)	(None, 4480)	0
dense (Dense)	(None, 64)	286,784
dense_1 (Dense)	(None, 5)	325

Fig.5. The structure of the model

V. TRAINING, VALIDATION, AND EVALUATION

The dataset was divided into training, validation, and test sets using a 70:20:10 split. A batch size of 30 was used during training, and the model was trained over 100 epochs. The training process yielded an average accuracy of 86%, with a peak accuracy of 90.03%.

```

1 scores = OM_model.evaluate(test_ds)
[19]
... 8/8 ----- 6s 191ms/step - accuracy: 0.9003 - loss: 0.2616

```

Fig.6. Evaluate the model using test dataset.

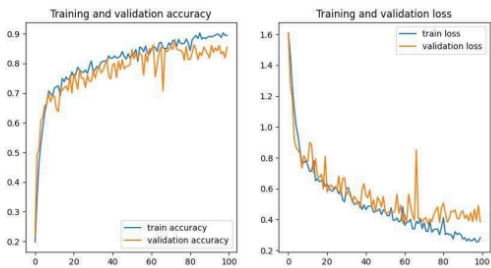


Fig.7. Accuracy and loss during training

Cross-validation was not employed due to the timeconsuming nature of retraining deep models with varying architectures. Instead, 10% of the data was set aside for evaluation purposes. The evaluation results confirmed a highest classification accuracy of 90.03%.



Fig. 8. Testing Result Visualization

VI. TESTING AND EVALUATION

To evaluate the model accuracy with test images the simple web interface was designed. It was allowed to feed input images to the model. The Flask library opted for its simplicity in hosting web applications via Python. Note that in real-world applications, this model likely be integrated with real-time camera feeds used by agricultural workers to scan plants for disease detection.

Using Flask, a straightforward web application was created that allows images to be uploaded and classified by the model. The main interface is shown in the following screenshot.

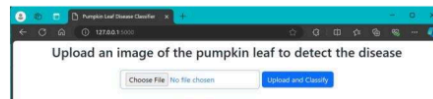


Fig.9. Interface of the web application

For testing purposes, the researcher selected 5 images from each category and removed them from the test dataset. These images serve as validation data to test the model's accuracy on previously unseen samples.

As a part of demonstration the pumpkin leaf with Bacterial Leaf Spot was first tested. The image is selected by clicking the "Choose File". It was found in the web application. Once selected, the image will be sent to the model for classification though "Upload and Classify."

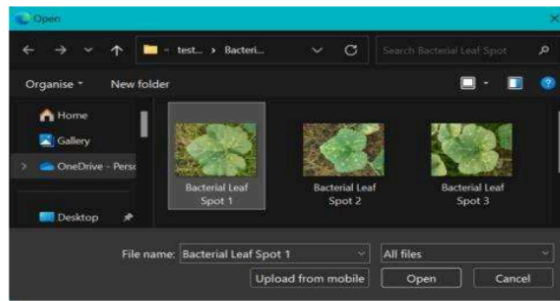


Fig.10. Choose a Bacterial Leaf Spot leaf image

The classification process was quick. The results are displayed under "Predicted Class" along with the confidence rate. As shown in the figure below, the application correctly identifies the disease type as Bacterial Leaf Spot with 58.51% confidence.

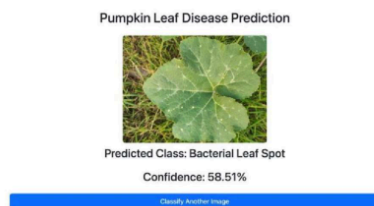


Fig.11. Result of testing 1

To test the next image quickly clicking "Classify Another Image" returns to the main interface. The second test case involves an image with Downy Mildew. The prediction results are shown in Figure 12. Here the model identifies the pumpkin leaf disease as Downy Mildew with 98.23% confidence.

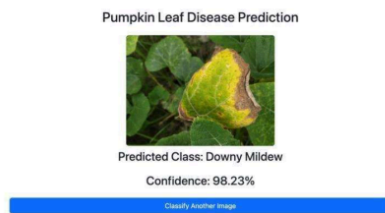


Fig.12. Result of testing 2

When a healthy leaf was selected for disease analysis (Figure 13), the model correctly identifies it as healthy with 94.92% confidence. Throughout testing, the model demonstrated relatively accurate identification of disease types on pumpkin leaves.



Fig.13. Result of testing 3

All 25 test images were processed by the model to determine its accuracy. Furthermore, on instances not included in the training set. The following table presents the results for each test case, which will be critically analyzed in the subsequent section.

Table I. Results of all test cases

Test Case	Screenshot of Result	Actual Category	Predicted Category	Result
A1		Healthy Leaf	Bacterial Leaf Spot (55.73%)	Incorrect
A2		Healthy Leaf	Bacterial Leaf Spot (65.62%)	Incorrect

A5		Healthy Leaf	Powdery Mildew (61.41%)	Incorrect
B1		Bacterial Leaf Spot	Bacterial Leaf Spot (58.51%)	Correct
B2		Bacterial Leaf Spot	Bacterial Leaf Spot	Correct

			(70.72%)	
B3		Bacterial Leaf Spot	Bacterial Leaf Spot (86.95%)	Correct
B4		Bacterial Leaf Spot	Powdery Mildew (63.53%)	Incorrect
B5		Bacterial Leaf Spot	Bacterial Leaf Spot (99.25%)	Correct
C1		Downy Mildew	Downy Mildew (98.23%)	Correct
C2		Downy Mildew	Downy Mildew (49.83%)	Correct
C3		Downy Mildew	Downy Mildew (77.67%)	Correct
C4		Downy Mildew	Downy Mildew	Correct

			(99.49%)	
C5		Downy Mildew	Downy Mildew (78.45%)	Correct
D1		Mosaic Disease	Mosaic Disease (100%)	Correct
D2		Mosaic Disease	Mosaic Disease (99.86%)	Correct
D3		Mosaic Disease	Mosaic Disease (95.85%)	Correct
D4		Mosaic Disease	Mosaic Disease (99.83%)	Correct
D5		Mosaic Disease	Powdery Mildew (87.05%)	Incorrect
E1		Powdery Mildew	Powdery Mildew (94.57%)	Correct
E2		Powdery Mildew	Powdery Mildew (82.71%)	Correct
E3		Powdery Mildew	Powdery Mildew (95.17%)	Correct
E4		Powdery Mildew	Mosaic Disease (100%)	Incorrect
E5		Powdery	Powdery	Correct

		Mildew	Mildew (99.91%)	
--	---	--------	--------------------	--

VII. CRITICAL EVALUATION

Despite an overall success rate of 72%, the model underperformed in correctly identifying healthy leaves, with only one correct classification out of five. Several misclassifications were categorized as Bacterial Leaf Spot, suggesting potential overfitting or feature overlap between diseased and healthy leaves.



Fig. 14. Comparison between healthy leaf (left) and bacterial spot leaf (right).

On the other hand, diseases with distinct color features, such as Downy Mildew, Powdery Mildew, and Mosaic Disease, were accurately classified, especially Downy Mildew which achieved 100% test accuracy.

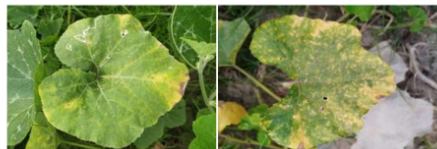


Fig. 15. Downy Mildew (left) and Powdery Mildew (right).

VIII. CONCLUSION

The current study successfully employed CNN to develop a deep learning model for pumpkin leaf disease classification. The results recorded a strong performance with 90.03% accuracy on the test dataset. The project was executed through several systematic phases. Initially, comprehensive preprocessing was performed on images from the selected Pumpkin Leaf Diseases Dataset. These preprocessing steps contain noise reduction, edge detection, image enhancement, and histogram equalization. During the subsequent model development phase, the additional data augmentation and normalization techniques were implemented cautiously. The primary objective of these preprocessing steps was to optimize the final accuracy of the leaf disease classification model.

Following model training completion, a web application was developed using the Flask framework. It was done to integrate with the trained model. During practical testing, this application demonstrated a reliable detection capabilities for various pumpkin leaf diseases with reasonable accuracy. However, the model showed limitations. Which includes correctly identifying healthy leaves, primarily due to their visual similarity to bacterial leaf spot, which shares comparable morphological features.

Several opportunities exist for enhancing this model performance and practical applicability. These include expanding the range of detectable pumpkin leaf diseases and fine-tuning the model architecture to achieve higher prediction accuracy. Moreover, addressing the current limitation in healthy leaf identification could significantly improve overall model reliability. With these enhancements, the model would become highly robust and practical for realworld agricultural applications. This provides farmers and agricultural professionals with a valuable tool for early disease detection and crop management.

Acknowledgement

This collaborative research under RDIG/2025/011 , composed of the Asia Pacific University of Technology and Innovation in Kuala Lumpur, Malaysia; Mindoro State University in Alcate, Victoria, Oriental Mindoro, Philippines; Chitkara University Institute of Engineering and Technology in Punjab, India; and at Coimbatore Institute of Technology, India. is supported by the respective research and innovation units of each university.

Biographical notes

DR. ADELINE SNEHA J

Dr. Adeline Sneha J is a Senior Lecturer in the School of Computing, Asia Pacific University of Technology and Innovation (APU), Technology Park Malaysia, Kuala Lumpur, Malaysia. She received her Ph. D from Sathyabama Institute of Science and Technology, Chennai, India. Her research spans several areas of IoT data analytics and machine learning, including Precision agriculture, Signal Processing, Image Processing, Data Science and Control Engineering. She started her teaching career in 2011 and since then, she has produced many highly qualified data scientists and engineers who have gone on to work for prestigious organizations like IBM, Amazon Web Services and Google. She is also a Life member of the International Society of Research and Development. She has delivered several invited or plenary lectures at conferences throughout the world. She has published more than 30 papers in various National/International Journals. She has received several important recognitions for her research career. She has also been awarded “Dr.

APJ Abdul Kalam Award for Innovative Research” awarded by Society for Engineering Education Enrichment (SEEE), 2021. She received “Women Leadership Award 2022” from GlantorX.

TS. DR. ARUN SEERALAN BALAKRISHNAN MIMECHE, C.ENG., (UK)

Ts. Dr. Arun Seeralan Balakrishnan is an academician in Mechatronics Engineering at the Asia Pacific University of Technology & Innovation (APU), Kuala Lumpur. With foundational training in Electrical, Electronics, and Mechanical Engineering, he gained seven years of industry experience in automation before completing his Master of Engineering in 2010 and later earning his PhD from Universiti Tenaga Nasional. Since 2010, he has guided numerous robotics projects, contributed to building APU’s robotics research center, and mentored teams for Malaysia’s ROBOCON competition. A Chartered Engineer of IMechE (UK) and BEM member, he conducts workshops, technical talks, and research in IR 4.0. He has published books, journal papers, and led notable industry-university collaboration projects, including a tram system prototype in 2025.

MAYTHEM K. ABBAS

Dr. Maythem K. Abbas is an Assistant Professor and seasoned university academician with more than 20 years of teaching experience (Academic and Industry), specializing in Artificial Intelligence (AI), Machine Learning (ML), programming, and the Internet of Things (IoT). He has published over 30 research papers featuring implemented systems and prototypes, contributing extensively to applied AI and IoT research. Dr. Maythem holds a BSc in Control and System Engineering from the University of Technology in Baghdad, a Higher Diploma in IT from Middlesex University in London, and both an MSc in IT and a PhD in Electrical and Electronics Engineering from Universiti Teknologi PETRONAS. Renowned for his dedication to innovation in teaching and research, he continues to advance impactful work at the intersection of intelligent systems and emerging technologies.

ROXANNE M. FABRIQUEL, MSc.

Ms. Roxanne M. FabriqueL is a licensed agriculturist and dedicated faculty member of the College of Agriculture and Allied Fields at Mindoro State University (MinSU). She earned her Master of Science in Agriculture, major in Crop Science, from Mindoro State University, and obtained her Bachelor of Science in Agriculture, major in Plant Pathology, from the University of the Philippines Los Baños (UPLB). As an academic and extension professional, she has been actively engaged in instruction, research, and community outreach promoting sustainable crop production, soil health management, and integrated pest and disease control. Her research interests include sustainable vegetable production, crop protection, organic agriculture, and plant disease diagnostics. She is an active member of Philippine Phytopathological Society,

Inc. and actively engaged in the Pest Management Council of the Philippines and Philippine Agriculturist Association, Inc.

ENGR. RANDY A. JOCO, MSc.

Engr. Randy A. Joco is the Director for Innovation at Mindoro State University and Project Lead of MinSU–iBIBES, a DOST–PCIEERD–funded incubator for biosystems-based technologies. He leads major R&D initiatives in agricultural mechanization, postharvest engineering, and biosystems innovation, focusing on transforming research into practical, scalable solutions. Under his guidance, MinSU–iBIBES advances smart and sustainable agricultural technologies and urban farming systems, promoting startup development and technology commercialization. With several Scopus-indexed publications, Engr. Joco emphasizes converting research outputs into viable technologies that address local agricultural challenges. His work positions Mindoro State University as a regional hub for innovation, supporting startups, building partnerships, and strengthening the ecosystem for biosystems-based solutions in MIMAROPA.

ENGR. JERREL S. REYES, PhD

Engr. Jerrel S. Reyes, Ph.D., is a licensed Professional Agricultural and Biosystems Engineer, with a Doctorate Degree in Crop Science, a Master's Degree in Agriculture, a Bachelor's Degree in Agricultural Engineering in the Philippines, and a Certificate in Leadership and Management Certificate Program from the University of Pennsylvania – The Wharton School. As an academician, he is currently an Associate Professor at Mindoro State University, Philippines, teaching in the bachelor's and graduate programs. As a researcher and innovator, he has authored and co-authored numerous publications in agriculture, agricultural and biosystems engineering, and related disciplines. He is also a co-inventor of a patented invention and several registered utility models. In the area of leadership and management, he currently serves as Director of External Linkages and International Affairs, Chief of Presidential Management Staff of the University, and Regional Quality Assurance Evaluator for the Commission on Higher Education – MIMAROPA.

RINAMAY A. JOCO

Ms. Rina May Aclan Joco is a registered agriculturist with a Bachelor of Science in Agriculture, major in Crop Science, from Mindoro State University. She currently serves as Project Assistant II for the research project “FARMIQ: AI-Powered Intelligent Farm Management and Resource Allocation—A Comparison Between Malaysia, India, and the Philippines.” Her work involves coordinating field activities, supporting data collection, and assisting in research on AI-driven solutions for sustainable and efficient farm

management. With her expertise in crop science, she contributes to advancing agricultural innovations that enhance productivity and promote smart farming systems.

DR. RAJASREE S.

Dr. Rajasree is an Assistant Professor in the Department of Information Technology at Coimbatore Institute of Technology, India. She holds a B.Tech in Information Technology from Amrita Vishwa Vidyapeetham, Coimbatore, and an M.Tech in Information Technology from Anna University, Chennai. She earned her Ph.D. in Information and Communication Engineering from Anna University, Chennai.

Her research interests span across interdisciplinary domains, with a focus on health care data analytics, gene expression analysis, and smart farming. Her work integrates computational intelligence with real-world applications, contributing to data-driven solutions in health informatics and precision agriculture.

DR AMIT KUMAR

Dr Amit Kumar is working as an Assistant Professor in the Department of Electronics and Communication Engineering, Chitkara University Punjab. He has completed his PhD with specialization in AR/VR technology for Embedded systems and a Masters in ECE with specialization in computer vision. He has published 9 SCI/Scopus journals and 21 conference papers. His research areas are IoT and Embedded Systems, AR/VR, HCI, ML, Device integration and multidisciplinary approach for product development.

DR. SHIVANI MALHOTRA

Dr. Shivani Malhotra, Ph.D. (Chitkara University), is Professor and Dean, Department of Electronics and Communication Engineering, CUIET, and leads Software & Mechatronics Engineering at Chitkara University, Punjab. With 20+ years of teaching and over a decade of academic leadership, she specializes in nanomaterials for microwave applications, IoT antennas, and solar cell design. Dr. Malhotra has published 30+ research articles, held 12+ patents (3 granted), and managed government projects worth INR 2.5 crore. An IIECP-certified educator, she received IGIP accreditation (IUCEE), the 2021 ISTE Best Faculty award, and is an MHRD Innovation Ambassador.

DR. ISHA GUPTA

Dr. Isha Gupta is an Assistant Professor in the Department of Electronics and Communication Engineering at Chitkara University, Punjab, specializing in Artificial Intelligence and

Machine Learning with emphasis on Deep Learning applications. Her research focuses on medical image analysis, disease detection, and computer vision. Dr. Gupta has authored eleven research articles in reputed international journals, with significant contributions to healthcare diagnostics, cancer detection using CNN architectures, and FPGA-based hardware optimization.

IX. REFERENCES

- [1] J. K. Soni and S. Hota, "Detection of multiple crop diseases using image processing techniques," in *Proc. IEEE High Performance Extreme Computing Conf. (HPEC)*, Waltham, MA, USA, 2021, pp. 1–6, doi: 10.1109/HPEC49654.2021.9622812.
- [2] H. Lin, H. Sheng, G. Sun, *et al.*, "Identification of pumpkin powdery mildew based on image processing PCA and machine learning," *Multimedia Tools Appl.*, vol. 80, pp. 21085–21099, 2021, doi: 10.1007/s11042-020-10419-1.
- [3] L. Li, S. Zhang, and B. Wang, "Plant disease detection and classification by deep learning—A review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3076359.
- [4] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 8, p. 707, 2021, doi: 10.3390/agriculture11080707.
- [5] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers," *Plants*, vol. 9, no. 10, p. 1319, 2020, doi: 10.3390/plants9101319.
- [6] S. K. Upadhyay and A. Kumar, "A novel approach for rice plant diseases classification with deep convolutional neural network," *Int. J. Inf. Technol.*, vol. 14, pp. 185–199, 2022, doi: 10.1007/s41870021-00817-5.
- [7] V. Nirmala and B. Gomathy, "Hybrid deep learning method for diagnosis of cucurbita leaf diseases," *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. xxx–xxx, 2022, doi: 10.32604/csse.2023.027512.
- [8] Y. A. Bezabh, B. M. Abuhayi, A. M. Ayalew, and H. A. Asegie, "Classification of pumpkin disease by using a hybrid approach," *Smart Agric. Technol.*, vol. 7, p. 100398, 2024, doi: 10.1016/j.atech.2024.100398.
- [9] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311–318, Feb. 2018, doi: 10.1016/j.compag.2018.01.009.
- [10] P. Patidar, M. Gupta, S. Srivastava, and A. K. Nagawat, "Image de-noising by various filters for different noise," *Int. J. Comput. Appl.*, vol. 9, no. 4, pp. 45–50, Nov. 2010, doi: 10.5120/1370-1846.

- [11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [12] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 8, p. 707, Aug. 2021, doi: 10.3390/agriculture11080707.
- [13] M. Shoaib et al., "An advanced deep learning models-based plant disease detection: A review of recent research," *Front. Plant Sci.*, vol. 14, p. 1158933, Apr. 2023, doi: 10.3389/fpls.2023.1158933.
- [14] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: Tomato leaf disease detection using convolution neural network," *Procedia Comput. Sci.*, vol. 167, pp. 293–301, 2020, doi: 10.1016/j.procs.2020.03.225.
- [15] H. Lin, H. Sheng, G. Sun, Y. Li, M. Xiao, and X. Wang, "Identification of pumpkin powdery mildew based on image processing, PCA and machine learning," *Multimedia Tools Appl.*, vol. 80, pp. 21085–21099, 2021, doi: 10.1007/s11042-02010419-1.
- [16] V. Nirmala and B. Gomathy, "Hybrid deep learning method for diagnosis of Cucurbita leaf diseases," *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. 2585–2601, 2023, doi: 10.32604/csse.2023.027512.
- [17] Y. A. Bezabih, B. M. Abuhayi, A. M. Ayalew, and H. A. Asegie, "Classification of pumpkin disease by using a hybrid approach," *Smart Agric. Technol.*, vol. 7, p. 100398, Feb. 2024, doi: 10.1016/j.atech.2024.100398.
- [18] A. J. Pandian, K. Kanchanadevi, N. R. Rajalakshmi, and M. R. Kumar, "An improved deep residual convolutional neural network for plant leaf disease detection," *Comput. Intell. Neurosci.*, vol. 2022, p. 5102290, Feb. 2022, doi: 10.1155/2022/5102290.

1-Pumpkin Leaf disease detection -Adeline.docx

ORIGINALITY REPORT

13%

SIMILARITY INDEX

8%

INTERNET SOURCES

7%

PUBLICATIONS

5%

STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

1%

★ dokumen.pub

Internet Source

Exclude quotes On

Exclude matches Off

Exclude bibliography On

1-Pumpkin Leaf disease detection -Adeline.docx

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

/100

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26