



A Comprehensive Survey on Potato Leaf Disease Detection Using Machine Learning, Deep Learning, and Explainable Artificial Intelligence Techniques

Ankush Kumar¹, Deepak Rao khadatkar²

¹Shri Shankaracharya Institute of Professional Management and Technology Raipur India

²Shri Shankaracharya Institute of Professional Management and Technology Raipur India

¹haloanku@gmail.com

²d.khadatkar@ssipmt.com

Corresponding Author: haloanku@gmail.com

Abstract

Potato (*Solanum tuberosum* L.) is considered one of the most significant food crops across the globe. In addition, potatoes are highly valuable for ensuring food security. Nevertheless, potato cultivation is significantly hampered due to different kinds of diseases such as Early Blight, Late Blight, Bacterial Wilt, and Potato Virus Y (PVY). Traditional disease recognition technologies are based on the manual detection of potato diseases performed by agricultural experts. As a result, disease diagnosis requires much time, effort, and is prone to errors. The development of innovative Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Computer Vision technologies allowed developing advanced algorithms that can be used for detecting plant diseases in an automated manner. The research provides an overview of main potato diseases, publicly available data sets, image pre-processing techniques, traditional machine learning algorithms, deep learning models, transfer learning techniques, and Explainable Artificial Intelligence (XAI). Moreover, the paper pays special attention to popular data sets such as Plant Village, PlantDoc, Kaggle Potato Leaf Disease Data Set, and Potato Leaf Dataset (PLD). Besides, the review discusses state-of-the-art models including SVM, RF, CNN, VGG16, ResNet50, DenseNet121, EfficientNet, vision transformers, and hybrid deep learning techniques. The paper captures some of the recent trends on lightweight architecture, real-life disease detection, disease severity prediction, and explainability by Grad-CAM, SHAP, and LIME. A comparison of current work has been done to bring out the problems that exist, such as the variation in the environment, small number of datasets acquired from the field, class imbalance, computational complexity, and model generalization problem. The paper concludes by looking into the future trends such as edge computing, federated learning, multimodal sensing, transformers, and intelligent precision agriculture systems.

Keywords:

Potato Leaf Disease Detection; Machine Learning; Deep Learning; Computer Vision; Explainable Artificial Intelligence (XAI); Precision Agriculture.

Received on 15 July 2026 ; Revised on 26 Aug 2026, Accepted on 19 Sep 2026; Published on 30 Sep 2026

DOI: <https://doi.org/10.67512/jcseam.v1.i3.4>

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are properly cited.

1. Introduction

Potato is one of the most important food crops in the world and is an important source of nutrition and income for many people. It is the fourth largest food crop next to rice, wheat and maize. There is high risk of potato production from several diseases, among which are Early Blight due to *Alternaria solani* and Late Blight due to *Phytophthora infestans*. These diseases affect the growth of plants, reduce the yield and cause significant economic loss to the growers if not recognised and treated in time [1]. The detection of potato diseases is traditionally done via visual analysis conducted by the experts of agriculture; however, such technique takes much time and effort and is quite subjective [2].

The latest developments in the fields of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have made tremendous contributions to the diagnosis of plant diseases in agriculture by making it possible for automated plant disease detection from leaf images. In controlled and uncontrolled environment, convolutional neural networks (CNNs), transfer learning models and hybrid intelligent models have demonstrated outstanding performance for potato leaf disease detection [3]. Moreover, public datasets and explainable artificial intelligence (XAI) technologies have increased the interpretability of these models, thereby making it easier to implement them in practice in precision agriculture. However, there still exist some issues that need to be addressed through research and analysis in this field [4].

Apart from its direct effects on the yield, the economic impact of diseases on potatoes has far-reaching consequences for food security, farmers' incomes, and the agricultural supply chain. Among all diseases affecting the potatoes, Early Blight and Late Blight are the most threatening ones because of the rapid spreading nature and the capacity to affect not only the leaves but also stems and tubers. The Late Blight caused by *Phytophthora infestans*, in particular, becomes a problem especially in the presence of optimal environmental conditions, resulting in yield losses up to 30% – 100%. Similar effects are observed during the case of the Early Blight caused by *Alternaria solani* when the potato productivity can decrease up to 20% – 50% [5].

On a global scale, potato diseases cost millions of dollars each year owing to such negative factors as the reduction in the crop production, increased use of pesticides and decay after harvest. Extra costs related to the disease monitoring, fungicides, labor expenses, and other methods used to control the development of diseases in the field should be taken into account. Also, infected potatoes suffer from the low marketability. In developing countries, the lack of plant pathology specialists and laboratories leads to the delay in disease diagnostics, increasing crop loss and threats to farmers [6].

Climate variability is another factor that has resulted in increased disease occurrence due to the creation of favorable conditions for pathogen growth and infection spread. Thus, intelligent, automatic, and economically viable approaches to disease diagnosis have been made essential. Machine learning and deep learning methods can help make accurate disease diagnosis possible, which will allow farmers to take preventive actions and greatly avoid financial losses caused by potato disease epidemics [7] (For summary see table 1).

1.1 Types of Potato Diseases

Fungi and oomycetes, as well as bacteria, viruses, and nematodes have been known to infect potatoes. Both Early Blight and Late Blight are the most investigated diseases since they have significant effect on potato production. *Alternaria solani* is the cause of Early Blight. The symptoms of such disease are dark brown spots on older leaves with concentric circles which cause the leaves to drop and disrupt photosynthesis. The disease can occur in humid and warm

climate and can result in great losses of production. *P. infestans* is the cause of Late Blight. It is considered to be one of the most destructive diseases of potatoes all over the world.

1.1.1 Early Blight (See Fig.1), *Alternaria solani* causes a fungal disease of potato called blight, characterized by black spots with rings and yellow halos on old leaves. It causes low photosynthesis, leading to early leaf fall, with yield losses of 20 to 50 percent and poor quality of tubers[8].



Fig. 1. Early blight

Causative Agent: *Alternaria solani*

Symptoms:

- Presence of small dark brown lesions on older leaves
- Ring formation around the lesions (Target Board Appearance)
- Formation of yellow halo around the lesions
- Premature defoliation
- Decreased photosynthesis

Impact:

Yield loss of 20–50%

Reduced tuber quality

1.1.2 Late blight (See Fig.2), One of the most harmful diseases that can occur to potato plants in the world is caused by *Phytophthora infestans*. The symptoms are watery black spots on leaves, fungal growth under the leaves, rapid development of the disease in humid weather conditions, browning of the leaves, and tuber rots. This disease may cause crop failure, with losses reaching up to 100 percent [9].



Fig. 2. Late blight

Causative Agent: *Phytophthora infestans*

Symptoms:

Water-soaked lesions on leaves

Rapid spread under humid conditions
White fungal growth on leaf undersides
Leaf blackening and collapse
Tuber rot

Impact:

Yield losses up to 100% in severe outbreaks
Most destructive potato disease worldwide

1.1.3 Bacterial Wilt (See Fig.3)

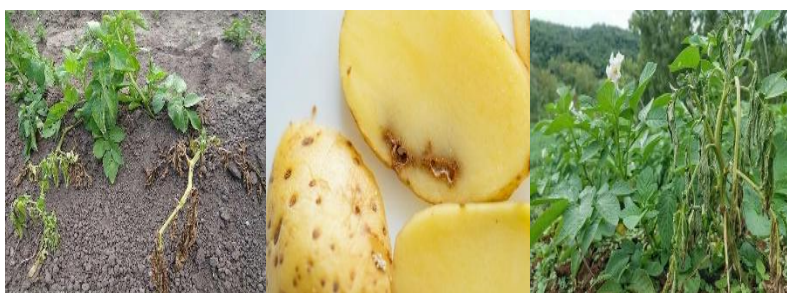


Fig. 3. Bacterial wilt

Causative Agent: *Ralstonia solanacearum*

Symptoms:

Sudden wilting of leaves
Stem discoloration
Brown vascular tissues
Plant death without yellowing

Impact:

Severe yield reduction
Soil-borne disease difficult to eradicate

1.1.4 The Potato Virus Y (PVY) (See Fig.4) is a virus disease that leads to the development of mosaic symptoms and leaf curling, as well as stunted plant growth. The impact of the disease on the plant growth and tuber formation is negative and results in small size of potatoes and yield reduction up to 80% [10].



Fig. 4. Potato virus

Causative Agent: Potato Virus Y

Symptoms:

- Reduced tuber size

Impact:

- Yield reduction of 10–80%
-

Table 1. Major Potato Diseases and Their Characteristics

Disease	Causal Agent	Disease Type	Major Symptoms
Early Blight	Alternaria solani	Fungal	Brown concentric lesions, yellow halos
Late Blight	Phytophthora infestans	Oomycete	Water-soaked lesions, leaf collapse
Bacterial Wilt	Ralstonia solanacearum	Bacterial	Wilting, vascular browning
Common Scab	Streptomyces scabies	Bacterial	Corky lesions on tubers
Black Scurf	Rhizoctonia solani	Fungal	Black sclerotia on tubers
Potato Virus Y (PVY)	Potato Virus Y	Viral	Mosaic patterns, leaf curling
Potato Leafroll Virus (PLRV)	PLRV	Viral	Leaf rolling, stunted growth

1.2 Need for early disease detection

Due to the harmful effects of diseases in potato farming, such as Early Blight, Late Blight, and Potato Virus Y, early disease detection has become an important factor in current agriculture practice. Diseases often occur with visible changes in the appearance of leaves; however, if the diseases are not detected promptly, the pathogens can rapidly spread and result in serious crop loss. In particular, Late Blight may damage entire potato fields in a short period of time with appropriate environmental conditions [9].

The conventional technique of disease detection involves visual inspection performed by farmers and other agriculture professionals. While this approach works to some extent, it is quite subjective and time-consuming and does not allow

for disease detection at the early stage of development. Insufficient number of specialists in the area of plant pathology in many agricultural regions prevents prompt disease detection and diagnosis.

However, due to the emergence of machine learning and computer vision techniques, scientists have started developing automated algorithms for disease detection by analyzing pictures of the leaf. Prompt detection of diseases will allow farmers to apply efficient disease treatment techniques and reduce chemical use.

1.3 Limitations of traditional manual inspection

Even though it is extensively applied in the practice of agriculture, the manual inspection has some drawbacks that influence its effectiveness in diagnosing diseases of potatoes. The basis for the process of diagnosing is done through visual recognition of symptoms as leaf signs conducted by farmers or experts of phytopathology. However, there are some diseases, like Early Blight and Late Blight, which have similar symptoms at the first stage of development, which makes them difficult to be distinguished [13]. The process of diagnosing highly depends on the level of skills of the observer, which leads to incorrect disease diagnosis [13].

The second limitation of the method is its high labour intensity, making it ineffective for the inspection of large fields with potatoes. An ordinary commercial field contains tens of thousands of potato plants growing in different parts of the field, occupying hectares of territory. Furthermore, the symptoms of a disease may occur only in several leaves that would be easily overlooked in the process of manual inspection of the field. As far as the Late Blight is concerned, its quick spread occurs under moist conditions, leading to crop loss in a couple of weeks [14].

The external environmental factors further complicate visual recognition. Variation in the level of lighting, shading, dust, interwoven leaves, and background can pose a problem in identifying the disease symptoms, reducing the accuracy level. In addition, there is a shortage of agronomists in rural areas where agricultural practices take place, where the occurrence of such diseases is high. For instance, the disparity in the number of plant pathologists and farmers in developing countries results in poor disease management practices [15].

1.4 Role of Artificial Intelligence (AI) in Agriculture and Potato Leaf Disease Detection

Modern agriculture has introduced a new technology, Artificial Intelligence (AI), that is one of the solutions to these disadvantages of manual testing. With the help of AI, various applications can be developed for the analysis of agricultural data, including crop monitoring, yield prediction, weed identification, irrigation management, and disease identification. AI uses computer vision, machine learning (ML), and deep learning (DL) techniques to process a huge amount of images and recognize symptoms of the disease faster and more accurately than conventional methods [16].

There is an increased use of the Artificial Intelligence Technology in potato cultivation for disease identification because of its ability to detect the disease at an early stage from the images of the potato leaves. Classification of healthy and unhealthy leave using Algorithms of ML, SVM, RF and KNN based on color, texture and shape. Several modern deep learning architectures such as CNN, VGG16, ResNet50, DenseNet121, and EfficientNet, achieved classification accuracy higher than 95% on various datasets including PlantVillage and potato leaf collected datasets. Specifically, transfer learning architectures have proven to be very effective at the detection of Early Blight, Late Blight and Healthy potatoes [17].

Furthermore, techniques such as Grad-CAM and SHAP from Explainable Artificial Intelligence (XAI) increase the transparency of the models and identify the parts of leaves affected by diseases, thus increasing users' trust and implementing the models. Therefore, artificial intelligence-based systems for disease identification have become critical for precision agriculture [18].

2. Literature Survey

Table 2. Comparative Analysis of Recent Potato Leaf Disease Detection Studies

Author & Year	Dataset Used	Methodology	Description of Work	Summary / Findings
Too et al. (2019) [28]	PlantVillage	Transfer Learning (VGG16, ResNet50, DenseNet121, InceptionV3)	Compared multiple pre-trained CNN architectures for plant disease classification.	DenseNet121 achieved superior classification performance and demonstrated the effectiveness of transfer learning on agricultural datasets.
Agarwal et al. (2020) [29]	PlantVillage Potato Dataset	CNN with Data Augmentation	Developed a CNN-based framework for automatic potato disease classification using leaf images.	Achieved high classification accuracy but performance was limited to controlled laboratory conditions.
Ferentinos (2021) [30]	PlantVillage and Field Images	Deep CNN Architecture	Evaluated deep learning models for plant disease recognition across multiple crop species.	Demonstrated that CNN models significantly outperform traditional machine learning techniques.
Tm et al. (2021) [31]	Potato Leaf Dataset	CNN + Image Processing	Applied preprocessing and CNN-based classification for Early Blight and Late Blight detection.	Improved disease recognition accuracy through image enhancement and feature extraction.
Noyan et al. (2023) [32]	PlantVillage	Fine-Tuned CNN	Proposed a fine-tuned CNN architecture for potato blight detection with balanced training samples.	Improved classification accuracy and reduced overfitting through augmentation.
Khan et al. (2024) [33]	PlantVillage + PLD	Attention-Based Lightweight Deep Learning (DenseNet169, MobileNet, XceptionNet, VGG16)	Developed lightweight models suitable for mobile deployment and resource-constrained environments.	Attention mechanisms enhanced disease localization while maintaining computational efficiency.
Natarajan et al. (2024) [34]	PlantDoc	Explainable Deep CNN	Introduced explainable AI for disease classification under real agricultural conditions.	Improved interpretability and robustness for field-acquired images.
Alshammari et al. (2025) [35]	Kaggle Potato Dataset + PlantVillage	CNN + Zero-Shot Learning	Investigated disease recognition on unseen disease	Enhanced model generalization and

			samples and cross-dataset validation.	transferability across datasets.
Dame et al. (2025) [36]	Smartphone Potato Dataset (4,200 Images)	CNN-Based Severity Assessment Framework	Developed a disease classification and severity estimation system using field-acquired images.	Enabled simultaneous disease diagnosis and severity prediction for precision agriculture.
Sinamenye et al. (2025) [37]	Real-Field Potato Dataset	EfficientNetV2B3 + Vision Transformer	Combined CNN and Transformer architectures to improve disease detection performance.	Achieved superior classification accuracy and robustness under environmental variability.
Zhang et al. (2025) [38]	Potato Disease Dataset	LDL-MobileNetV3S	Proposed a lightweight deep learning model with dynamic dilated convolution and attention modules.	Suitable for mobile and edge-based disease diagnosis with low computational cost.
Bhavani and Chalapathi (2026) [39]	PlantVillage	PotatoLeafNet (Two-Stage CNN)	Developed a two-stage CNN framework for potato disease identification and classification.	Achieved 98.52% accuracy with reduced model complexity.
Nawaz et al. (2026) [40]	Real-Field Potato Dataset	PotatoGuardNet (Faster R-CNN + InceptionResNet-V2)	Integrated disease localization and classification into a single framework.	Improved disease detection under complex field conditions and enabled precise localization.

It is obvious from the analysis of the researched works (See Table 2) that there is a variety of developments in the detection of leaf diseases of potatoes through deep learning approaches, which are superior to the traditional machine learning approaches in terms of accuracy. The first set of researches analyzes transfer learning based on models, such as VGG16, ResNet, DenseNet, and InceptionV3, which can classify PlantVillage images with more than 95% of accuracy [19-22]. The latest studies analyze the light models, explainable artificial intelligence and hybrid models CNN-Transformer [23-27].

However, it is necessary to note that nowadays the scientists tend to switch from controlled datasets to real-field datasets and smartphone datasets. One of the most efficient models in the analyzed literature is the PotatoLeafNet model with the accuracy of 98.52% [28]. Nevertheless, there are still some problems connected with the lack of field dataset, diverse environment conditions and the generalization capability of the models. Therefore, the future researches should be focused on the real-time implementation, multimodal approach, disease severity assessment and Explainable Artificial Intelligence.

Scope of the Survey

- Analyzes commonly used datasets, preprocessing methods, and classification approaches.
- Evaluates recent advancements in transfer learning and Explainable Artificial Intelligence (XAI).

- Identifies challenges, research gaps, and future directions in potato leaf disease detection.

Objectives of the Survey

- To provide a comprehensive review of potato leaf disease detection methods reported from 2017–2026.
- To analyze datasets, feature extraction techniques, and evaluation metrics used in existing studies.
- To investigate the role of XAI techniques in improving model interpretability
- To identify limitations and suggest future research opportunities for intelligent disease diagnosis systems.
-

3. Dataset used in the different research work

3.1 PlantVillage Dataset

Data set PlantVillage is one of the most popular datasets used in potato leaf disease detection (See Table 3). Images in PlantVillage dataset are captured in the lab with the same background and lighting. This dataset has a sub-dataset called "potato" which contains 2,152 images, categorized into three classes such as Early Blight (1,000 images), Late Blight (1,000 images) and Healthy leaves (152 images). This is one of the most widely-used datasets for Deep Learning algorithms because of the quality and labeling of the images [29].

Table 3. Implementation of PlantVillage Dataset in Recent Research Studies

Reference	Model	Preprocessing	Performance	Key Contribution
Alhammad et al. [19] (2025)	VGG16 + Grad-CAM	Image resizing (224×224), normalization, rotation, flipping, and zooming augmentation	Accuracy $\approx$ 98%	Improved interpretability through Grad-CAM-based disease localization and visualization.
Bhavani and Chalapathi [20] (2026) (PotatoLeafNet)	Two-Stage CNN	Sequential augmentation, CNN feature extraction, Softmax classification	Accuracy: 98.52%, Precision: 98.67%, Recall: 99.67%, F1-Score: 99.16%	Achieved high classification accuracy with reduced computational complexity.

Advantages

- High-quality and noise-free images.
- Well-labeled disease categories.
- Suitable for benchmarking deep learning models.
- Balanced representation of major potato diseases.

Limitations

- Captured under controlled conditions.
- Uniform background and lighting.
- Limited representation of real-field environments.
- Models may exhibit poor generalization on field-acquired images.

3.2 PlantDoc Dataset

(See Table 4) PlantDoc database is made to counteract limitations associated with laboratory environment databases such as PlantVillage. It consists of 2,598 images captured in practical agricultural settings and includes 13 crop types and 27 disease classes. The images include practical challenges such as differences in light, complex backgrounds, leaf overlap, occlusion, and view angle differences[30].

Table 4. Implementation of PlantDoc Dataset in Recent Research Studies

Reference	Model	Preprocessing	Performance	Key Contribution
Natarajan et al. [21] (2024)	Deep CNN with Explainable AI	Image normalization, data augmentation, and background adaptation techniques	Improved field-level disease recognition	Enhanced model generalization under real agricultural conditions and provided interpretable disease predictions.
Attention-CNN Cross-Dataset Study [22] (2025)	Attention-based CNN	Training on PlantVillage and PlantDoc datasets; testing on PlantDoc potato images	Early Blight: 83.33%, Late Blight: 56.25%	Demonstrated the challenges of real-world disease detection and highlighted performance degradation due to field variability.

Advantages

- Suitable for evaluating model robustness and practical deployment.
- Supports cross-dataset generalization studies.

Limitations

- Limited number of potato leaf images.
- Class imbalance across disease categories.
- Higher intra-class variability compared to laboratory datasets.
- More challenging for model training and classification.

3.3 Kaggle Potato Leaf Disease Dataset

There are many open-source datasets on Kaggle that can be employed for classification of the diseases affecting potato leaves (See Table 5). Majority of them consist of 2,000 to 5,000 images pertaining to Healthy, Early Blight and Late Blight categories. These images are acquired under semi-controlled and field conditions and provide more variation compared to lab-based data sets.[31]

Table 5. Implementation of Kaggle Potato Leaf Disease Dataset in Recent Research Studies

Reference	Model	Preprocessing	Performance/Objective	Key Contribution
Zero-Shot Early Blight Detection Study [23] (2025)	CNN with Zero-Shot Learning Framework	Image resizing, data balancing, normalization	External validation performed on PlantVillage dataset (1,200 images)	Improved disease recognition on unseen data and enhanced model generalization across datasets.

Dataset Details Used in the Study

Class	Number of Images
Healthy	1,000
Early Blight	1,151
Total	2,151

Advantages

- Publicly accessible and easy to obtain.
- Multiple dataset versions are available.
- Suitable for transfer learning and cross-dataset validation.
- Contains images from diverse acquisition conditions.

Limitations

- Dataset quality varies across different Kaggle sources.
- Annotation inconsistencies may exist.
- Image resolution and class distribution are not standardized.

3.4 Potato Leaf Dataset (PLD)

PLD (See Table 6) stands for the Potato Leaf Dataset and it is a large dataset designed to add variety to the images compared to the PlantVillage. PLD dataset contains 4,062 images of potato leaves divided into three categories – Healthy, Early Blight, and Late Blight. Due to its size and realism of images, PLD is usually used alongside the PlantVillage.[32]

Table 6. Implementation of PLD in Recent Research Studies

Reference	Dataset	Model	Preprocessing	Hardware	Objective
Attention-Based Lightweight Deep Learning Model [24] (2024)	PlantVillage + PLD	DenseNet169, MobileNet, XceptionNet, VGG16	Image augmentation, CLAHE enhancement, image resizing	NVIDIA Tesla T4 GPU, 32 GB RAM	Development of lightweight and accurate disease diagnosis models suitable for mobile deployment

PLD Dataset Distribution

Class	Number of Images
Early Blight	1,628
Late Blight	1,414
Healthy	1,020
Total	4,062

Advantages

- Larger sample size than many potato disease datasets.
- Greater diversity in image characteristics.
- Improves model generalization when combined with PlantVillage.
- Suitable for training deep learning architectures.

Limitations

- Moderate class imbalance among categories.
- Some images contain environmental variations that increase classification complexity.
- Requires additional preprocessing for optimal performance.

Various datasets have been employed for detecting potato leaf diseases, such as artificial or actual datasets. PlantVillage is the most popularly used dataset because of the high-quality images and balance between disease classes in it. However, more researchers nowadays use the PlantDoc, Kaggle Potato Leaf Disease Datasets, and Potato Leaf Dataset (PLD) for their analysis. Table 7 shows information regarding the characteristics, advantages, and disadvantages of the primary datasets that have been used in studies.

Table 7. Comparative Summary of Potato Leaf Disease Datasets

Dataset	Total Images	Classes	Image Environment	Major Applications	Advantages	Limitations
PlantVillage	2,152	Healthy, Early Blight, Late Blight	Controlled laboratory conditions with uniform background	CNN, Transfer Learning, Explainable AI, Disease Classification	High-quality images, balanced disease classes, standard benchmark dataset	Uniform background, controlled illumination, poor field generalization
PlantDoc	2,598	27 disease classes across 13 crops	Real agricultural field conditions	Robustness evaluation, Explainable AI, Cross-dataset learning	Real-world images, varying illumination and backgrounds, better deployment evaluation	Small potato subset, class imbalance, higher variability
Kaggle Potato Leaf Disease Dataset	2,000–5,000 (varies by version)	Healthy, Early Blight, Late Blight	Semi-controlled and field conditions	Transfer Learning, Zero-Shot Learning, CNN-based classification	Easily accessible, multiple versions available, suitable for transfer learning	Dataset quality varies, annotation inconsistencies, non-standardized collection
Potato Leaf Dataset (PLD)	4,062	Healthy, Early Blight, Late Blight	Diverse field and semi-controlled conditions	Deep Learning, Mobile-based Disease Detection,	Large sample size, better diversity, improved generalization	Moderate class imbalance, requires extensive preprocessing

				Cross-dataset training		
--	--	--	--	------------------------	--	--

From the above analysis of the existing datasets in Table 7, it is evident that the PlantVillage datasets are good for benchmarking and training the models, whereas the PlantDoc and PLD datasets are more realistic when it comes to testing the models. At the same time, Kaggle-based datasets are very flexible and easily accessible but are not consistent regarding the image quality and annotations. This has resulted in many researchers using more than one datasets during the training phase of the models.

4. Methodology Framework for Potato Leaf Disease Detection

Typically, a universal methodology for development of automated detection systems of potato leaf diseases is used which includes image processing, machine learning, deep learning and explainable AI methods. The purpose of the methodology is to diagnose the disease symptoms through images of the potato leaf and help farmers to take necessary actions regarding the disease management. Based on the surveys made in the articles mentioned above, the methodology can be divided into six main stages: image acquisition, image pre-processing, data augmentation, feature learning, classification and explainability (See Fig.5).

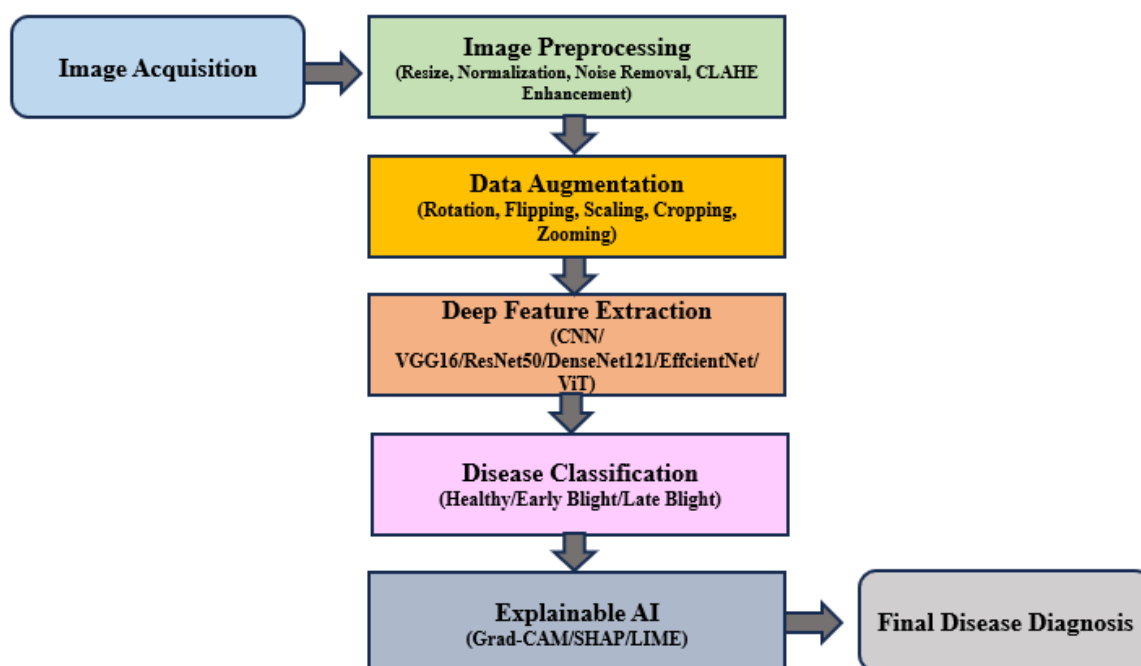


Fig. 5. General Methodology Framework for AI-Based Potato Leaf Disease Detection and Classification

4.1 Data Acquisition

The first stage of the methodology is acquiring the images of potato leaves from public datasets or field environment. The most commonly used datasets for disease detection include PlantVillage, PlantDoc, Potato Leaf Disease Datasets of Kaggle and PLD (Potato Leaf Datasets). All of these datasets have both healthy leaf images and diseased leaf images of early blight and late blight diseases. While lab-based datasets include annotated high-quality images, field-based

datasets can have issues like varying illumination, occlusion, shadowing and background. Images taken from different sources ensure greater reliability of the disease detection models.

4.2 Image Preprocessing

The inconsistency and noise in the agricultural images may lead to poor classification results during processing. To enhance the quality of images, image preprocessing methods are used prior to the process of training for the detection of a specific disease. The most common techniques used are image resizing, normalization, contrast improvement, noise reduction and color correction. The CLAHE (Contrast Limited Adaptive Histogram Equalization) method is used in many studies to improve symptoms' visibility.

4.3 Data Augmentation

Large and diversified datasets are required for effective functioning of deep learning algorithms. Since agricultural datasets are often small, the augmentation of the dataset is carried out to artificially enlarge the dataset size by creating more training examples. Data augmentation includes the rotation, flipping, scaling, translation, cropping, zooming and change of image brightness. This way, the variety of the images increases and the chance of overfitting decreases.

4.4 Feature learning and representation

After pre-processing step, image enters into machine learning or deep learning algorithm in order to get features. Machine learning uses predefined features like colors, textures, shapes, edges. Deep learning algorithms are able to learn the features of the image. CNN, transfer learning, and transformer models are capable of learning the discriminative features of disease including lesion boundaries, discoloration, texture abnormalities, and infected regions.

4.5 Disease classification

The features learned are employed for classification of the image. Depending on the goal, classification algorithm divides the images into categories, such as healthy, early blight, late blight. Some of the most popular classification models are SVM, RF and KNN. Nowadays deep learning algorithms like CNN, VGG16, ResNet50, DenseNet121, EfficientNet, Vision Transformers are more often used. These metrics such as accuracy, precision, recall, F1-score, specificity, or the AUC can be used for model evaluation.

4.6 Explainable artificial intelligence

Although deep learning algorithms provide good classification, the explanation of their decision-making process is difficult. In order to overcome this problem, approaches of XAI are used in the disease detection process. Methods such as Grad-CAM, SHAP, LIME allow seeing what parts of the picture affect the algorithm's decision-making process. It increases transparency of the process, improves user's trust and makes it possible to implement AI systems for disease detection in the agriculture.

5. Conclusion

The potato leaf diseases are the most important ones affecting potato production and food security in the world and the wellbeing of the farmers. If not detected and controlled early diseases like Early Blight, Late Blight, Bacterial Wilt, and Potato Virus Y (PVY) can cause considerable crop and economic losses. Visual identification of disease is not always a reliable and suitable method to identify the disease for agricultural application. It has inspired the development of Disease Identification System with the technologies of Artificial Intelligence, Machine Learning and Deep Learning.

This survey performed a detailed analysis of the different potato leaf disease detection models available during the period 2019 to 2026. The main potato diseases, the datasets, image processing techniques, machine learning algorithms, deep learning models, transfer learning models, and Explainable Artificial Intelligence (XAI) techniques were considered in this survey. The results indicated that deep learning models like CNN and transfer learning models

like VGG16, ResNet50, DenseNet121, EfficientNet and Hybrid CNN-Transformer models outperformed traditional machine learning models. Different models obtained a classification accuracy of more than 95% in PlantVillage and Potato Leaf Dataset (PLD).

Furthermore, this survey found the value of the datasets to the disease detection studies. The PlantVillage dataset is the most used because of its high quality and class balance, however, the images are only taken in a controlled environment condition, restricting its use. Recent studies have thus explored data gathered in the field, like PlantDoc, PLD and smartphone data, to improve their generalizability. Furthermore, the applications of Explainable AI (XAI) methods like Grad-CAM, SHAP, and LIME have played a crucial role in enhancing the transparency of the models and building trust among the farmers for machine learning tools in disease detection.

Despite great progress that has been made, there are still some problems that exist. These problems are lack of large field datasets, environmental variation, class imbalance, severity level prediction, computational challenge, and limitations of deploying the models on low resource hardware among others. Potential future research directions can include the design of lightweight and efficient models that can be used in edge devices, use of multimodal data sources, the application of Vision Transformers and foundation models, federated learning solutions for private agricultural use, and the creation of XAI-based decision-making systems. The integration of AI, IoT, cloud computing, and precision agriculture can bring a revolution in disease detection and treatment.

In conclusion, AI-powered potato leaf disease detection systems seem to be a feasible option for implementing sustainable and precise farming methods. With the ongoing advancements in deep learning, Explainable AI, and real-time deployment, disease detection accuracy will keep on improving in the future to come.

Conflict of Interest Statement:

The authors declare that there is no conflict of interest regarding the publication of this work.

Funding Statement:

This research received no external funding.

References

- [1] Pacal, I. Kunduracioglu, M. H. Alma, M. Deveci, S. Kadry, and R. Martinek, "A Systematic Review of Deep Learning Techniques for Plant Diseases," *Artificial Intelligence Review*, vol. 58, no. 2, pp. 1–48, 2025.
- [2] M. H. Shah et al., "A Review on the Detection of Plant Disease Using Machine Learning and Deep Learning Approaches," *Journal of Imaging*, vol. 11, no. 10, pp. 326–348, 2025.
- [3] Kumar, R. Singh, and P. Sharma, "Potato Leaf Disease Detection Using Machine Learning," *Current Agriculture Research Journal*, vol. 11, no. 3, pp. 1–12, 2023.
- [4] M. A. Khan et al., "Cross-Comparative Review of Machine Learning for Plant Disease Detection: Apple, Cassava, Cotton and Potato Plants," *Smart Agricultural Technology*, vol. 8, pp. 100–118, 2024.
- [5] J. H. Gold, D. M. Andersson, and R. K. Singh, "Advances in Potato Disease Monitoring and Management," *Plant Pathology Review*, vol. 72, no. 4, pp. 401–420, 2020.
- [6] N. Dolničar, "Potato Diseases and Their Impact on Global Food Security," *Agricultural Systems*, vol. 190, pp. 103–118, 2021.
- [7] Y. Zhang, H. Wang, and X. Li, "Artificial Intelligence for Physiological and Molecular Monitoring, Detection, and Forecasting of Potato Late Blight," *Computers and Electronics in Agriculture*, vol. 235, pp. 109–128, 2026.
- [8] G. N. Agrios, *Plant Pathology*, 5th ed. Burlington, MA, USA: Elsevier Academic Press, 2005; and recent discussions in potato disease management literature including S. K. Pandey and A. Kumar, "Disease Management in Seed Potato Production," *Indian Farming*, vol. 74, no. 1, pp. 45–52, 2024.
- [9] Y. Zhang, H. Wang, and X. Li, "Large-Scale Monitoring of Potato Late Blight Using Multi-Source Time-Series Data," *Remote Sensing*, vol. 17, no. 6, pp. 978–995, 2025.

- [10] M. A. Khan et al., "Cross-Comparative Review of Machine Learning for Plant Disease Detection: Apple, Cassava, Cotton and Potato Plants," *Smart Agricultural Technology*, vol. 8, pp. 100–118, 2024.
- [11] M. H. Shah et al., "A Review on the Detection of Plant Disease Using Machine Learning and Deep Learning Approaches," *Journal of Imaging*, vol. 11, no. 10, pp. 326–348, 2025.
- [12] Kumar, R. Singh, and P. Sharma, "Potato Leaf Disease Detection Using Machine Learning," *Current Agriculture Research Journal*, vol. 11, no. 3, pp. 1–12, 2023.
- [13] M. H. Shah et al., "A Review on the Detection of Plant Disease Using Machine Learning and Deep Learning Approaches," *Journal of Imaging*, vol. 11, no. 10, pp. 326–348, 2025.
- [14] Y. Zhang, H. Wang, and X. Li, "Large-Scale Monitoring of Potato Late Blight Using Multi-Source Time-Series Data," *Remote Sensing*, vol. 17, no. 6, pp. 978–995, 2025.
- [15] M. A. Khan et al., "Cross-Comparative Review of Machine Learning for Plant Disease Detection: Apple, Cassava, Cotton and Potato Plants," *Smart Agricultural Technology*, vol. 8, pp. 100–118, 2024.
- [16] M. H. Shah et al., "A Review on the Detection of Plant Disease Using Machine Learning and Deep Learning Approaches," *Journal of Imaging*, vol. 11, no. 10, pp. 326–348, 2025.
- [17] Kumar, R. Singh, and P. Sharma, "Potato Leaf Disease Detection Using Machine Learning," *Current Agriculture Research Journal*, vol. 11, no. 3, pp. 1–12, 2023.
- [18] M. A. Khan et al., "Cross-Comparative Review of Machine Learning for Plant Disease Detection: Apple, Cassava, Cotton and Potato Plants," *Smart Agricultural Technology*, vol. 8, pp. 100–118, 2024.
- [19] E. C. Too, L. Yujian, S. Njuki, and Y. Yingchun, "A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.
- [20] Z. Saeed et al., "Deep Learning-Based Classification of Potato Leaf Diseases Using ResNet152 and InceptionV3," 2021.
- [21] R. Mahum et al., "Enhanced Deep Learning Technique for Diagnosis and Categorization of Potato Leaf Diseases," 2022.
- [22] M. A. Noyan et al., "Potato Blight Detection Using Fine-Tuned CNN Architecture," *Mathematics*, vol. 11, no. 6, 2023.
- [23] M. S. Khan et al., "Attention-Based Light Weight Deep Learning Models for Early Potato Disease Detection," *Applied Sciences*, vol. 14, no. 17, 2024.
- [24] T. Astuti et al., "Enhancing Potato Leaf Disease Detection Using Convolutional Vision Transformer," *Journal of Imaging and Vision Computing*, 2024.
- [25] S. M. Alhammad et al., "Deep Learning and Explainable AI for Classification of Potato Leaf Diseases," *Frontiers in Artificial Intelligence*, vol. 7, 2025.
- [26] T. A. Dame et al., "Deep Learning-Based Potato Leaf Disease Classification and Severity Assessment," *Discover Applied Sciences*, 2025.
- [27] J. H. Sinamenye, A. Chatterjee, and R. Shrestha, "Potato Plant Disease Detection: Leveraging Hybrid Deep Learning Models," *BMC Plant Biology*, vol. 25, 2025.
- [28] G. D. Bhavani and M. M. V. Chalapathi, "PotatoLeafNet: Two-Stage Convolutional Neural Networks for Effective Potato Leaf Disease Identification and Classification," *Frontiers in Artificial Intelligence*, 2026.
- [29] D. P. Hughes and M. Salathé, "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics Through Machine Learning and Crowdsourcing," *arXiv preprint arXiv:1511.08060*, 2015.
- [30] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A Dataset for Visual Plant Disease Detection," in *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, Hyderabad, India, 2020, pp. 249–253.
- [31] M. A. Hossain, M. R. Islam, S. A. Hossain, and M. A. Rahman, "Potato Disease Detection Using Deep Learning Techniques on Publicly Available Datasets," *Computers and Electronics in Agriculture*, vol. 216, 2024.

- [32] M. M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, "Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine," *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1–4, 2017.