IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 1, Ser. 1 (Jan. – Feb. 2026), PP 62-71
www.iosrjournals.org

Advancements In Deep Learning Techniques For Image-Based Detection Of Diseases In Leaves Of Potato: A Review

Hima Gandhi[1], Madan Lal[2], Kanwal Preet Singh Attwal[3]
[1]Department Of Computer Science And Engineering, Punjabi University, India
[2,3]Department Of Computer Science And Engineering, Punjabi University, India

Abstract:
This review article analyzes the significant role of deep learning in the automated classification and detection of potato leaf diseases, as they prominently threaten global agricultural production. Traditional approaches to identify diseases typically rely on manual inspection processes that are labor-intensive, time-consuming, and prone to human-induced error, leading to variability and producing inconsistent outcomes. In contrast, deep learning frameworks utilize convolutional neural networks (CNNs), leveraging transfer learning techniques to enhance diagnostic accuracy while reducing computational overhead. By using an extensive dataset of annotated potato leaf images, several studies report that deep learning models can effectively discriminate between healthy and diseased foliage, focusing on prevalent conditions such as Potato Late Blight and Potato Early Blight. This study highlights that deep learning based approaches demonstrate promising performance primarily under controlled and semi-controlled conditions when compared to traditional machine learning methods, while robustness across diverse real-world environments remains an active area of research. Furthermore, the article examines the practical challenges inherent in real-life agricultural environments that include variations in illumination and complex backgrounds that may hinder the process of accurate disease symptom recognition. Moreover, while existing studies evaluate pre-trained and hybrid deep learning models for plant leaf disease detection, these evaluations are often conducted on different crops, datasets, and experimental settings. This highlights a research gap in the systematic and consistent comparison of multiple pre-trained deep learning models for potato leaf disease detection under unified evaluation frameworks, particularly in the context of real-world agricultural conditions. Ultimately, the review aims to expand its scope to ensure broader accessibility to reliable diagnostic tools, thereby enabling farmers and agricultural stakeholders with effective resources to manage potato leaf diseases. By supporting the adoption of such innovative technologies, the study contributes to ongoing efforts aimed at improving agricultural resilience and productivity in the context of global challenges.

Key Word: *Potato Leaf Disease Detection, Deep Learning, Convolutional Neural Network (CNNs), Image Classification, Agricultural Automation.*

Date of Submission: 14-01-2026 Date of Acceptance: 24-01-2026

I. Introduction

Agriculture is an essential sector that provides food, agricultural products, raw materials, and other assets to the whole population [1]. Potatoes (Solanum tuberosum) are one of the most important food crops globally. They are the fourth most widely cultivated food crop across the world [2]. Potatoes are necessary food crops followed by rice, wheat and maize. Due to rich nutritional content and supply of vital carbohydrates, vitamins, and minerals like potassium, it is one of the major staple foods. During the phenological cycle of potato growth plant, it can be infected by several types of diseases which can be specific to its natural growth cycle. It can generate obstacles in the overall agricultural production leaving food scarcity for certain population. It can majorly impact the yield and reduce the quality of potato crop. The diseases which infect the potatoes can be majorly classified as Early Blight, Late Blight, Fungi, Bacteria, Virus, Pest, Nematode, Phytophthora and other microorganisms. These infections can drastically reduce yields and interrupt supply chains, negatively impacting both producers and regional food markets. The major focus of detection of potato leaf disease is to mitigate the global problem of food scarcity and sustainability in agribusiness. Early detection of disease is crucial to take appropriate preventive measures, precise diagnosis and effective control measures. For detection and classification of many types of stressed potato leaf, different approaches can be implemented. The traditional approach was based on observing directly by naked eye which is error prone and time consuming method. With advancements in technology, there is a revolution in the research and development of disease detection techniques. By leveraging machine learning, deep learning algorithms and techniques over potato leaf are suitable for early disease detection. In earlier times, there was no easy way to detect the spread of contagious plant disease using prerequisite knowledge and required prolonged hours of labor efforts. In addition, the visual perception and interpretation abilities of humans affect the precision of identification [3]. To enhance the accuracy of disease detection, advanced techniques such as machine learning, deep learning, and computer vision are increasingly utilized. These techniques are more affordable and easier to implement and update, which makes them a more practical choice for agricultural settings. By extracting meaningful features from leaf images, software-based methods classify the images as either healthy or diseased.

The adoption of deep learning technologies, especially convolutional neural networks(CNNs), has risen as a highly impactful tool for identifying plant diseases. By analyzing extensive collection of leaf images, CNNs are able to detect fine-grained visual features and distinguish between different disease categories with high efficiency, making them a strong substitute for traditional manual inspection.

The following are the summary of the contributions of this survey :

- The paper discusses a review of the existing literature focusing on traditional methods and recent advancements in automated disease detection systems.
- Additionally, it compiles and presents all publicly available annotated datasets relevant to potato leaf disease detection studies.
- The paper provides a thorough examination of the limitations in the existing models used for detecting potato leaf diseases.
- It highlights the key challenges encountered by researchers in this domain and proposes potential directions for future research.

The rest of the paper is structured as follows: Section 2 represents a review of related work. An overview of Deep Learning techniques is discussed in Section 3. The section 4 discusses metrics used. Section 5 introduces commonly used datasets in Potato leaf disease detection domain. Section 6 discusses challenges and issues in this field. Section 7 presents conclusion and suggests future research directions to enhance model performance.

II. Literature Review

This section presents the contribution of the researchers in the field of potato leaf disease detection. Recent progress in artificial intelligence has highlighted its substantial potential in the domain of plant disease image classification. Researchers have investigated a broad spectrum of methodologies, encompassing conventional machine learning algorithms, state-of-the-art deep learning architectures, and hybrid approaches integrating convolutional neural networks (CNNs) with vision transformers. These techniques have consistently demonstrated effectiveness in accurately identifying diverse plant diseases using image-based datasets. Early research in this domain predominantly relied on traditional machine learning approaches. Hylmi, Wiharto and Suryani [4] employed a multi-class support vector machine (SVM) for potato leaf disease detection using texture, color, and shape-based features. Their method, which involved leaf spot extraction along with RGB histograms, GLCM, and contour analysis, reported an accuracy of 97.56%. Similarly, another study by Nishad, Mitu and Jahan [5] applied K-Means clustering to improve potato production and several augmentation techniques like VGG16, VGG19, and ResNet50 network model, however VGG16 achieved 97% accuracy. Rusli et al. [6] developed a classification framework combining image processing with artificial neural networks to enable early detection of potato leaf diseases and minimize agricultural losses. K-means clustering algorithm was used to segment features of disease and then extracted using Gray Level Co-occurrence Matrix(GLCM) and then fed to ANN for classification, thus obtaining an accuracy of 94%. A study by Iqbal and Talukder [7] introduced an image processing and machine learning-based system for classifying Early Blight (EB) and Late Blight (LB). Using 450 images from the PlantVillage dataset and testing seven different classifiers, the Random Forest algorithm emerged as the most effective, attaining a 97% accuracy rate in differentiating diseased from healthy leaves. For instance, in study by Panigrahi et al. [8], multiple traditional machine learning classifiers, including Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest (RF) were applied for maize disease classification tasks. Among these models, Random Forest achieved the highest classification accuracy of 79.2%, surpassing the performance of the other evaluated techniques.

The progression of research has seen deep learning, particularly CNN architectures, rise to prominence. For example, a study by Shrestha et al. [9] utilized a CNN-based approach for plant disease detection. 15 cases were fed to the model, out of which 12 were diseased and 3 healthy. With an extensive dataset of 3000 high resolution RGB images, the model classified achieved an accuracy of 88.8% .

Further advancements in the field have involved the application of pre-trained models. The study by Sauda, Kaur and Lal [10] highlight the use of the PlantVillage dataset containing healthy and diseased tomato leaf images across three categories. Inception-ResNet-v2, Inception-v3, and ResNet-50 models were evaluated, where Inception-v3 achieved the highest accuracy of 99.5%. Inception-ResNet-v2 further demonstrated increasing accuracy as the number of epochs progressed, indicating strong learning capability. The study by Hong, Lin and Huang [11] highlights the implementation of multiple CNN architectures, including ResNet50, Xception, MobileNet, ShuffleNet, and DenseNet121_Xception, for classifying 9 types of tomato leaf disease classification. By applying transfer learning, the researchers reduced both training time and computational cost. Among the tested models, DenseNet121_Xception achieved the highest recognition accuracy of 97.10%. More recently, hybrid models that combine CNNs with vision transformers (ViTs) have been introduced, leveraging the spatial feature extraction strength of CNNs alongside the global attention mechanism of transformers to improve accuracy and generalization. Building on these advances, pure ViT-based approaches have also been explored. For instance, Borhani, Khoramdel and Najafi [12] proposed a lightweight ViT framework for real-time plant disease detection, highlighting its potential as an efficient alternative to conventional CNN architectures. Arshad et al. [13] introduced a hybrid deep learning framework, termed PLDPNet, which leveraged deep features extracted from VGG19 and Inception-V3, while employing vision transformers for the final classification stage. Their approach achieved an accuracy of 98.66% and an F1-score of 96.33%, demonstrating the effectiveness of combining CNN and transformer-based architectures for potato disease detection. However, the study also emphasized a key limitation that is the restricted availability of well-annotated potato disease datasets which constrained the generalizability and scalability of their model. Furthermore, Shaheed et al. [14] proposed an enhanced architecture named EfficientRMT-Net, which combined the strengths of Vision Transformers with ResNet50. The framework reported an accuracy of 97.65% on a generic image dataset and achieved 99.12% accuracy when applied to potato leaf images from the Plant Village dataset. Another study by Sinamenye, Chatterjee and Shrestha [15] introduced a hybrid deep learning framework, EfficientNetV2B3 and ViT models for potato disease detection, integrating the feature extraction capability of EfficientNetV2B3 with the contextual learning power of ViTs. Trained on the Potato Leaf Disease Dataset, which captures realistic agricultural conditions, the model attained an accuracy of 85.06%, marking an 11.43% improvement over prior approaches. This demonstrates the promise of hybrid CNN–ViT architectures for enhancing disease identification in complex field environments.

III. Deep Learning Techniques In Image Classification

Deep Learning is a subset of machine learning with an algorithm inspired by the structure and function of the brain, which is called an Artificial Neural Network (ANN). The computation of multilayer neural networks becomes feasible as it uses many hidden layers in the network, thus having the expression "Deep". Unlike machine learning, it eliminates the requirement of manual preparation of data due to its capacity of feature extraction. This complex pattern recognition ability for disease identification and classification with context to agriculture has emerged as a powerful approach. This field leverages deep convolutional neural networks(CNNs) to perform automated learning from raw image inputs, enabling precise recognition and classification of fine differences in plant heath conditions.

Convolutional Neural Networks

Convolutional Neural Networks(CNNs) are the predominant architectures in deep learning, learning directly from data and eliminates manual feature extraction. CNN is based on the concept of artificial neural networks, consisting of one or more hidden layers along with input and output layers. These layers facilitate the learning process to identify patterns and relationships in data [16]. The incoming data is traversed though input layer in form of pixel values followed by hidden layers. These hidden layers are responsible for applying convolution function for extraction of local features including edges, textures and shapes. The convolutional layer is responsible for the matrix multiplication or dot product. Furthermore, Activation function such as Rectified Linear Unit (ReLU) introduces non-linearity and converts negative values to zero, thereby network learns complex patterns. For input data X with k filters with wj and bj presenting weight and bias with f as activation function. yj denotes the final computed output of neuron j after applying the weighted sum and activation [17]. * represents the convolutional process in equation(1) as follows:

$$y_j = \sum_i f\left(x_i * w_j + b_j\right),\ j=1,2,\ldots k \qquad 1$$

Pooling layers are added further to minimize the spatial dimensions of feature maps while pertaining significant data. To avoid overfitting and reduce computation, max pooling and average pooling functions are applied. These layers together form a hierarchy of attributes allowing CNN to detect complex image patterns, significant for image classification tasks. Taking p × p window size, activation value as xij at location (i,j) with N as total items in S. z is the average value of all activations in the window as given in equation(2) :

$$z = \frac{1}{N} + \sum_{(i,j)\epsilon S}^{\infty} x_{i,j},\ i, j=1,2,\ldots,p \qquad 2$$

Towards the end of architecture, CNN includes batch normalization layer for activation normalization and fully connected layers to flatten and for combination of features extracted. Fully connected layer transforms each feature map to one-dimension feature vector. The output or final layer produce prediction using activation function. SoftMax function is used for multi-class classification and binary classification can be achieved using Sigmoid function. In given equation(3), ϒ and z defines input and output descriptor, w is weight and b represent bias for fully connected layer.

$$\Upsilon = \sum_i f(wx+b) \qquad 3$$

Thereby, model trained on CNN significantly reduces the rate of error between predicted and actual values. Through parameter optimization, error function is minimized and performance of CNN improves. Consequently, CNN is a robust artificial neural network model extendable to higher accuracy for image processing and detection.

The basic architecture of CNN consists of convolutional layers using filters to detect features in input images, an activation function introducing non linearity to the convolution outputs, pooling layer reduces results and pertaining features followed by fully connected layers for detection and classification tasks. [18]

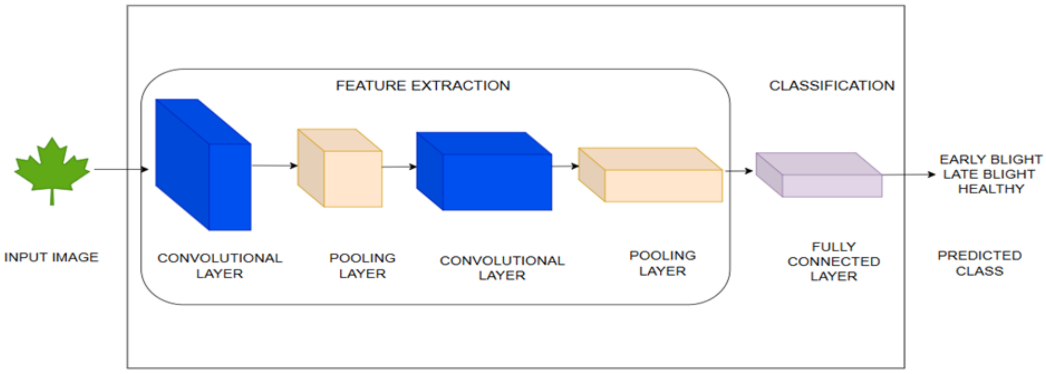

Figure 1: A simple CNN architecture, comprised of just five layers

Transfer Learning

Transfer Learning is a Convolutional Neural Network(CNN) and Deep Learning method. The knowledge is reusable and transferred from a pre-trained model using weights. The models obtained work faster and is more robust with even less amount of training data. Transfer Learning consists of two steps: pre-training and fine-tuning. The model is pre-trained usually on ImageNet dataset with consists of over 1.2 million images with more than 1,000 classes for basic understanding of data and features. Transfer Learning is performed on the baseline pre-trained model to selectively freeze its layers. Freezing make the layers untrainable and they are not further considered for backward propagation but only for forward passes during inference. Frozen layers consider original pre-trained weights which reduces overfitting and makes the model lightweight. For fine tuning, few layers are unfreezed for adaptation to new data and improvement of performance. Some methods are to freeze CNN layers and train only classifier, to freeze particular layers then unfreeze and fine tune the rest or only performing fine-tuning on the model. For instance, firstly freezing all layers, then unfreeze and fine-tune. For leaf disease detection, pre-trained models such as VGG16, Inception, ResNet etc. have been applied on different datasets. These datasets can be in controlled or uncontrolled environment.

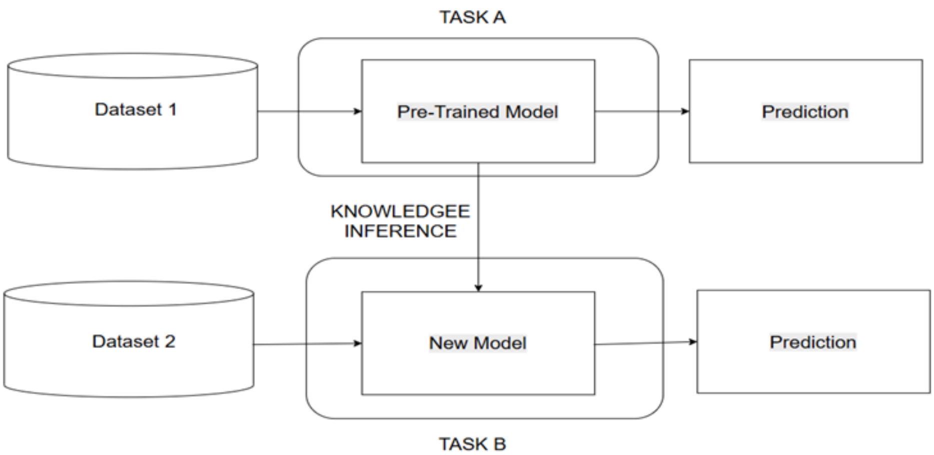

Figure 2: Transfer Learning

Pre-trained deep learning architectures

VGG: VGG is a Convolution Neural Network(CNN), developed by University of Oxford. VGG defines Visual Geometry Group which was introduced in 2014. It is trained on ImageNet dataset to classify various objects like pencil, vehicles, fruits, various animals etc. By leveraging transfer learning techniques, the pre-trained model can be developed quickly with high accuracy. The architecture consists of multiple convolutional layers and fully connected layers. It comprises of 16 or 19 weight layers, thereby VGG16 or VGG19 with small 3×3 convolutional filters. The structure is simple with advanced focus on feature extraction. Despite the advantages, some of the limitations are excessive memory and time consuming training process.

Inception-V3: Inception-V3 is a 48-layer deep CNN. It takes an input of size 299×299. The networks decrease representation of bottleneck. The network partitions 5×5 convolution into 3×3 for increasing computation efficiency. The number of modes in the output is equal to number of classes in the dataset. Thereby, using Inception-V3 rich feature representation can be learned for extensive range of images. [10]

ResNet: Residual Networks(ResNet) consists of residual connections in CNN layers, which is capable to overcome vanishing gradient problem. Model dimensions are reduced using bottleneck layers. The bottleneck layer comprises convolutions as 1×1 to reduce feature dimensions, followed by 3×3 to extract spatial features and again 1×1 to restore dimensions. ResNet takes less time and produce accurate results. These are used for the purpose of object detection, image classification and leaf disease detection. Limitations are significant memory consumption and high computational cost. Some of the ResNet variants are ResNet50, ResNet50v2,ResNet101, ResNet101v2, ResNet152 and ResNet152v2.

AlexNet: AlexNet is a pre-trained model based on LeNet-5 neural network. It embeds a deeper layer with more convolutional layers with filter in every layer. The model has five convolutional layers and three fully connected layers. To prevent overfitting, dropout function is used. ReLU is used as activation function to decrease computation expense. [19]

MobileNet: MobileNet is a lightweight Convolutional Neural Network(CNN) for mobiles and embedded vision applications. To lower computational cost and increase accuracy, MobileNet employ depth wise separable convolutions. They integrate depthwise and pointwise convolutions into a single efficient operation. These are designed to classify images in real-time for resource deficient devices. They are incorporated in facial popularity, medical diagnostics and autonomous systems for efficient inference. MobileNet has several variants including MobileNetV1, MobileNetV2, MobileNetV3, MobileNetV3-Small and MobileNetV3-Large.

DenseNet: In DenseNet, every layer is linked to every other layer. Thus, it is a densely connected CNN. The DenseNet network is based on residual connections. All the input layers get stacked as input goes on getting deeper with more channels. DenseNet, through concatenated (.) attribute use feature maps of each layer are used as inputs to all following layers. But with many features in last layer it can lead to feature map explosion. To mitigate this situation dense block is created with prespecified number of layers. DenseNet with using only few parameters can achieve high accuracy. DenseNet can be employed for different purposes like clinical imaging, disease identification, pattern recognition using different architectures like DenseNet161,DenseNet169 and DenseNet201.

Vision Transformers (ViTs) in Leaf Disease Detection:

Transformers are Deep Neural Networks (DNN), originally proposed for domain of natural language processing(NLP). They use self-attention mechanism to capture dependency between elements is sequence like words or sentences [20]. These mechanisms enable model to learn relationship between pixels of an image by identifying crucial information not captured by CNNs. Transformers leverage entire sequences using encoding and decoding blocks. Additionally, they work in a parallel manner capturing actual meaning of content unlike traditional LSTM and RNNs. Despite the advantages, one of the major challenges is high computational demand related to large datasets of images. But ViTs are capable to mitigate this gap.

Vision Transformers have recently elicited interest in computer vision with significant potential. ViTs analyze visual input by non-conventional techniques. Self-attention mechanism is major constituent which models interactions across distinct areas of input image emphasizing only relevant aspects. This adoption increases accuracy and reduces the dependence on extensive pre-processing, thereby streamlining the disease detection pipeline. Input image is parted into relevant size patches interacting with each other and further processed by dense feed- forward layers in transformer blocks. As a result, the output layer produces data for classification or other tasks. ViTs applications are new with more effective approach that enable model to emphasize important parts without multiple convolutional layers. ViTs recognize minute color change, texture alterations or irregularity in shape of leaves and stem. This approach is mostly beneficial where symptoms of disease appear irregularly on different parts of stem, leaves or other parts and require model that recognize patterns across different regions.

The primary objective of Image Transformers is analysis of images of various sizes effectively. The ViTs framework classifies input image into patches of fixed size and then transformed to continuous vectors by linear embedding method. Patch representation integrates positional embeddings for retaining the information of position. The transformer encodes receives a stream of image patches and consists of alternating layers of Multi-Head Self-Attention(MSA) and Multi-layer perceptron(MLP). Layer Normalization is applied prior to MSA and MLP layer to ensure optimization stability and to accelerate training. The Residual connections are thus incorporated after every layer for overall high performance. Despite high efficiency of CNNs, they come with drawbacks especially while working with colossal datasets and need for capturing long-range image dependency like texture of leaf or lesion structure of plant surface. Another challenge is CNNs rely on fixed receptive fields which makes it difficult for capturing relationships between distant regions of an image without incurring substantial computational cost. This limitation highlights the potential of Transformer-based architectures for complex and high-dimensional agricultural data.

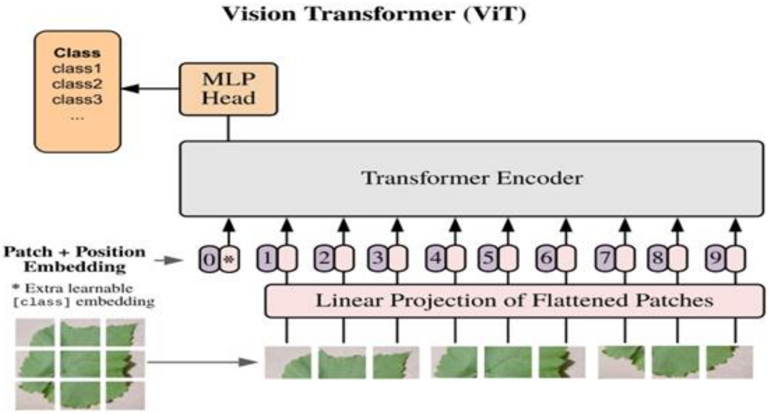

Figure 3: Vision Transformers Architecture[17]

IV. Performance Metrics For Disease Classification:

Evaluation metrics are essential for determining the performance and robustness of disease classification models. The metrics listed below are commonly applied in potato leaf disease detection tasks. The following are key metrics commonly used in disease classification models, particularly in the context of potato leaf disease detection [21].

Accuracy: Accuracy can be calculated as the ratio of correct predictions to the total predictions. As this metric is used commonly to evaluate model performance, it does not provide a complete understanding of a model's effectiveness, especially for imbalanced datasets. For example, in a dataset where healthy samples are significantly more, a model that primarily predicts "healthy" may achieve high accuracy but ineffective to identify diseased plants. In the following equations, TP and TN denotes True Positive and True Negative values respectively. FP and FN denote False Positive and False Negative values respectively.

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \qquad 4$$

Precision: Precision is described as the ratio of true positive prediction to the sum of true positives and false positives. This metric is important in disease detection as it express the ability of model to identify only those instance that exhibit disease symptoms. A high precision value signifies fewer false positives, thereby supporting more reliable decision-making and reduces unwarranted interventions.

$$\text{Precision}=\frac{TP}{TP+FP} \qquad 5$$

Recall: Recall represents the proportion of true positives predictions among all actual positive cases. It is also referred to as sensitivity. For disease detection applications, a high recall value is crucial to ensure that infected plants are correctly identified that enables timely intervention and treatment. Conversely, low recall indicates that many diseased cases may remain undetected, increasing the risk of disease spread.

$$\text{Recall}=\frac{TP}{TP+FN} \qquad 6$$

F1 Score: F1 score reflects the equal measure of trade-off and is defined as the hormonic mean of precision and recall. It provides a thorough evaluation of the performance of the model, by combining false positives and false negatives into a single evaluation measure. It is attested for the applications which require equality between precision and recall, remarkably for class imbalance.

$$\text{F1 Score}=2*\frac{\text{Precision*Recall}}{\text{Precision+Recall}} \quad 7$$

Confusion Matrix: It evaluates the performance of model through distinct classes evaluating any potential bias and classification errors.

These metrics altogether, help for evaluating and refinement of disease detection models. Furthermore, establishing necessary requirements for practical agricultural applications by complying by precision and reliable results. Also, for the purpose of visualization, the training and validation curves are plotted to identify training and validation loss and accuracy over the curves of epochs. This whole process of training is monitored for detecting overfitting or underfitting.

V. Publicly Available Datasets For Potato Leaf Disease Detection

The publicly available datasets play a crucial role in training and evaluating deep learning models for disease detection. Well-organized and properly annotated datasets allow researchers to build, compare, and refine models with greater accuracy. In potato disease detection studies, several open access datasets, most prominent the PlantVillage Dataset and additional relevant datasets.

PlantVillage Dataset: One of the most widely utilized datasets in plant disease research is the PlantVillage dataset, which comprise a diverse collection of images spanning multiple crop species, including potato. The potato category contains images of common leaf diseases such as Early Blight consisting of 1000 samples, Late Blight consisting of 1000 samples and Healthy with 152 samples making 2152 in total, each characterized by distinct visual patterns. This comprehensive image repository is highly valuable for training robust deep learning models with strong generalization capability across varying disease manifestations. [22]

Domain-Specific Datasets: Several domain-specific datasets are readily available for research focusing on specific geographic regions or include a particular disease strain relevant to that area. For instance, some datasets focus on potato crops from specific climates or locations where particular disease is pervasive. The dataset size can vary, but describes prominent insights into region-specific disease and influence on the environment, supports models adapting to regional differences in disease expression.

1.PotatoCare Dataset (Version 2): The dataset is introduced in 2025 with images over ten thousand with class labels 10 presenting diverse diseases caused by potato pests along with healthy samples. It is compiled from various sources and captures a wide range of symptom variations under both field and laboratory conditions, offering diverse visual representations of each disease. [23]

2.Potato Disease Leaf Dataset (PLD): The dataset is collected from Central Punjab region of Pakistan. A multi-level deep learning framework was applied to this dataset. Initially, YOLOv5-based image segmentation was employed to automatically extract potato leaves from complex field images that reduced background interference. Subsequently, a Convolutional neural network(CNN)- based classifier was used to identify diseases. [24]

Field-Based Datasets: These datasets are acquired from real field environment directly rather than controlled laboratory settings, thereby it captures substantial variability arising from background complexity with diverse directions and distance of the images, illumination changes and image perspective. These datasets facilitate the development of resilient models by exposing them to the variability inherent in field-level challenges. It encourages researchers to develop robust computer vision models to detect leaf pests in the presence of background complexity and occlusions. [25]

Table no 1: Illustrates the overview of publicly accessible datasets for potato disease detection research affecting the visible symptoms of disease and download URLs.

	Dataset Description	Primary Datasets	Disease Classes (Potato)	Attributes	Applications	Challenges & Considerations	Link
[22]	General-purpose Dataset	PlantVillage Dataset	3 classes: Early Blight, Late Blight, Healthy	2152 high-resolution images, multi-class, large scale in	Generalization focused model development for training and validation	Possible class imbalance among disease categories	https://www.kaggle.com/datasets/emmarex/plantdisease/data

				controlled settings			
[23]	Domain-Specific Potato Dataset	PotatoCare	10 disease classes: Black Scurf, Blackleg, Blackspot Bruising, Brown Rot, Common Scab, Dry Rot, Healthy Potatoes, Pink Rot, Miscellaneous and Soft Rot .	10,117 images compiled from diverse sources.	Focused on generalizability to create diverse representative collection of images.	Imbalanced classes as Blackspot bruising have more samples than Blackleg and Pink Rot.	https://data.mendeley.com/datasets/7vm7xskfg4/2
[24]	Domain-Specific Potato Dataset	Potato Disease Leaf Dataset(PLD)	3 classes: Early Blight, Late Blight, Healthy	4,062 field-acquired images from Central Punjab, Pakistan.	Focused on potato-specific disease recognition in real-world condition	Limited disease categories and regional bias may affect generalization.	https://www.kaggle.com/datasets/rizwan123456789/potato-disease-leaf-datasetpld
[25]	Field-Based Datasets	Field-collected potato disease images	7 classes: Virus, Phytophthora, Pest, Fungi, Nematode, Bacteria, Healthy	3,076 high-resolution images, diverse environmental conditions; natural lighting, background variability, multiple viewing angles	Diverse conditions of potato pests and diseases in real-world settings	Varied background and occlusion, Varied angle of view, Diverse object placement and orientation, Varied lighting condition	https://www.kaggle.com/datasets/nirmalsankalana/potato-leaf-disease-dataset/code, https://data.mendeley.com/datasets/ptz377bwb8/1

VI. Challenges In Deep Learning Based Potato Disease Detection

Inadequate Data and Quality

The critical issue in development of high performance deep learning models is data scarcity and quality. Deep Learning algorithms require a massive amount of data with labeled images for proficient learning to distinguish features. Thus, to assemble good quality dataset can be challenging especially for regions having minimum resources lacking technical experts and data infrastructure. Another issue is data imbalance, as some diseases are less common which leads to lower samples of a particular category. Thus, results lead to biased model performing well on majority class but poor performance for less common diseases. Data which is noisy containing poor resolution images, varying lighting conditions like shadows and backgrounds, inconsistently labelled moreover complicate training and decrease model accuracy. The proposed solution to this can be implementation of data augmentation to elaborate datasets available using transfer learning to leverage pre-trained models and by applying techniques like class balance. Regardless of these approaches, to obtain a high quality, balanced and representative data remains a crucial obstacle. [26]

General Issues

The prominent challenge for deployment of models is generalization for different regions, seasons and environmental settings. The difference in model performance for particular regions with certain conditions leads to discrepancy in variability of model performance exposed to unfamiliar conditions. For instance, specific environmental conditions having different lighting, humidity, temperature can affect appearance of leaf. Potato varieties grown for different areas exhibits varied symptoms of disease. Generalization challenges are addressed by diverse training data across environmental conditions and potato variety, along domain adaptation and data augmentation techniques to increase robustness of model.

Computational efficiency

Another major challenge for Convolutional Neural Networks (CNNs) is computational requirement. For the purpose to train and implement, model require necessary computational power that include high performing GPUs, substantial memory and extended processing times. Deep learning models are computationally intensive, hinder real-time applications. In context of agriculture, stakeholders like small-scale farmers and cooperatives, lack access to cutting-edge computing resources. In addition, for the mobile applications of deep learning models, real-time processing is complicated and demand lightweight models with low latency, difficult to achieve without performance degradation. Optimized techniques consisting if model pruning, quantization and knowledge distillation mitigate these challenges but achieving an optimal trade-off between efficiency and accuracy is however complex. Ensuring efficient model performance on affordable hardware is essential for the large-scale adoption of deep learning in agriculture. [27]

VII. Conclusion And Future Directions

In conclusion, the adoption of deep learning methodologies for the detection and classification of potato leaf diseases represents a transformative progression in precision agriculture, as this approach utilizes technology to optimize yield and resources for specific need of individual plants. This review indicates that deep learning models, particularly Convolutional Neural Networks (CNNs), constitute a scalable and automated alternative to conventional manual inspection paradigms, which are labor-intensive and susceptible to observer bias. By effective learning hierarchical and discriminative feature representations from leaf imagery, these models enable high-throughput and accurate diagnosis of potato diseases. Such capabilities contribute to early-stage disease identification, facilitating timely agronomic interventions to reduce crop loss and maximize yield outcomes.

Nevertheless, despite notable performance gains, several technical and operational challenges constrain the widespread deployment of deep learning systems in real-world agricultural contexts. These challenges include limited model generalization under heterogeneous environmental conditions, insufficient availability of high-quality and class-balanced annotated datasets, and substantial computational demands associated with training and inference. Ensuring reliable model performance across heterogeneous agricultural environments is a key determinant of real-world applicability. Furthermore, to achieve high levels of accuracy in disease detection, deep learning models require access to large-scale, high quality, and class-balanced annotated datasets for model training and validation. The insights presented in this review highlight the need for developing deep learning frameworks that are flexible and resource-efficient. By addressing these limitations, deep learning has capability to advance potato disease management, augmenting agricultural production while reinforcing global food security.

Looking ahead, the application of deep learning for potato leaf detection is poised for significant advancements. A central objective will be to strengthen model generalization by leveraging varied and representative datasets. This objective can be achieved by adopting semi-supervised and self-supervised learning strategies, which enable effective utilization of unlabeled or synthetically generated data, thereby reducing reliance on extensive annotated datasets. Such methodology has the potential to alleviate current challenges associated with data scarcity in this domain.

The findings of this review further emphasize the growing effectiveness of deep learning techniques in achieving accurate and early detection of potato leaf diseases. The comparative analysis of existing models and methodologies provides valuable guidance for researchers in selecting appropriate architectures, dataset, and evaluation strategies for future work. These insights support agricultural technologists and practitioners in developing reliable, automated disease detection systems that facilitate timely intervention and improved crop management. By identifying current trends, strengths, and limitations of deep learning approaches, this review also helps delineate key directions for future advancement.

Moreover, the development of lightweight models is critical to enable real-time disease detection in resource deprived environments. Techniques like pruning, quantization, and knowledge distillation can enhance the adaptability of computationally efficient models while maintaining high predictive accuracy, which is particularly important for smallholder farmers with limited computational resources.

Future research should also prioritize the integration of domain-specific agronomic knowledge into deep learning frameworks. By incorporating expert insights, these systems can be aligned to the unique requirements of distinct agricultural communities, thereby improving their practical utility and effectiveness. Synergistic collaboration between agronomy professionals and machine learning experts is essential for designing solutions that balance both technical robustness and practical application. Additionally, leveraging multi-modal data sources such as combining image data with environmental factors like temperature, soil moisture, humidity, pH level can significantly increase accuracy and reliability of potato disease detection models. This multi-faceted approach enhances the understanding of the major determinants of potato health, thereby facilitating more informed and effective decision making for farmers. Overall, the study contributes both theoretically, by consolidating existing knowledge, and practically, by facilitating the development of intelligent decision-support tools for sustainable agriculture.

References

[1]. I. Attri, L. K. Awasthi And . T. . P. Sharma, “Machine Learning In Agriculture: A Review Of Crop Management Applications,” Multimedia Tools And Applications, Vol. 83, Pp. 12875–12915, 2023.
[2]. M. E. Çalişkan, A. Bakhsh And K. Jabran, Potato Production Worldwide, Academic Press, 2023.

[3]. P. Andal And M. Thangaraj , "Comprehensive Review Of Methods For Leaf Disease Identification," Discover Artificial Intelligence, Vol. 5, 2025.
[4]. M. S. Hylmi, W. And E. Suryani, "Detection Of Potato Leaf Disease Using Multi-Class Support Vector Machine Based On Texture, Color, And Shape Features," International Conference On Electrical And Information Technology (Ieit), Pp. 20-24, 2022.
[5]. M. A. R. Nishad, M. A. Mitu And N. Jahan, "Predicting And Classifying Potato Leaf Disease Using K-Means Segmentation Techniques And Deep Learning Networks," Procedia Computer Science, Vol. 212, Pp. 220-229, 2022.
[6]. A. H. T. Rusli, B. C. C. Meng, N. S. Damanhuri, N. A. Othman, M. H. Othman And W. F. A. W. Zaidi, "Potato Leaf Disease Classification Using Image Processing And Artificial Neural Network," IEEE 12th International Conference On Control System, Computing And Engineering (Iccsce), Pp. 107-112, 2022.
[7]. M. A. Iqbal And K. H. Talukder, "Detection Of Potato Disease Using Image Segmentation And Machine Learning," International Conference On Wireless Communications Signal Processing And Networking (WiSPNET), Pp. 43-47, 2020.
[8]. K. P. Panigrahi, H. Das, A. K. Sahoo And S. Moharana, Maize Leaf Disease Detection And Classification Using Machine Learning Algorithms, 2020, Pp. 659-669.
[9]. G. Shrestha, D. M. Das And N. Dey, "Plant Disease Detection Using CNN," IEEE Applied Signal Processing Conference (Aspcon), Pp. 109-113, 2020.
[10]. S. Sauda, L. Kaur And M. Lal, "Comparison Of Pre-Trained Deep Model Using Tomato Leaf Disease Classification System," In Machine Intelligence And Smart Systems, 2021, Pp. 553-566.
[11]. H. Hong, J. Lin And F. Huang, "Tomato Disease Detection And Classification By Deep Learning," International Conference On Big Data, Artificial Intelligence And Internet Of Things Engineering (ICBAIE), Pp. 25-29, 2020.
[12]. Y. Borhani, J. Khoramdel And E. Najafi, "A Deep Learning Based Approach For Automated Plant Disease Classification Using Vision Transformer," Scientific Reports, Vol. 12, No. 1, 2022.
[13]. F. Arshad, M. Mateen, S. Hayat, M. Wardah, Z. Al-Huda, Y. H. Gu And M. A. Al-Antari, "PLDPNet: End-To-End Hybrid Deep Learning Framework For Potato Leaf Disease Prediction," Alexandria Engineering Journal, Vol. 78, Pp. 406-418, 2023.
[14]. K. Shaheed, I. Qureshi , F. Abbas, S. J. Sohail Jabbar , Q. Abbas, H. Ahmad And M. Z. Sajid, "EfficientRMT-Net—An Efficient Resnet-50 And Vision Transformers Approach For Classifying Potato Plant Leaf Diseases," Sensors, Vol. 23, No. 23, 2023.
[15]. J. H. Sinamenye, A. Chatterjee And R. Shrestha , "Potato Plant Disease Detection: Leveraging Hybrid Deep Learning Models," Bmc Plant Biology, Vol. 25, No. 1, 2025.
[16]. T. Ozcan and E. Polat, "BorB: A Novel Image Segmentation Technique for Tomato Plant Leaf Detection in Images Captured in Complex Outdoor Environments," Vol. 13, Pp. 71822-71839, 2025.
[17]. I. Pacal, I. Kunduracioglu, M. H. Alma, M. Deveci, S. Kadry, J. Nedoma, V. Slany And R. Martinek , "A Systematic Review Of Deep Learning Techniques For Plant Diseases," Artificial Intelligence Review, Vol. 57, 2024.
[18]. R. Nash And K. T. O'shea, "An Introduction To Convolutional Neural Networks.," Arxiv Preprint, 2015.
[19]. A. Bajpai, M. Tyagi, M. Khare And A. Singh, "A Robust And Accurate Potato Leaf Disease Detection System Using Modified Alexnet Model," International Conference On Computer And Communication Engineering (Iccce), Pp. 264-269, 2023.
[20]. S. Islam, H. Elmekki, A. Elsebai, J. Bentahar, N. Drawel, G. Rjoub And W. Pedrycz, "A Comprehensive Survey On Applications Of Transformers For Deep Learning Tasks," Expert Systems With Applications, Vol. 241, 2024.
[21]. A. Kaur, G. S. Randhawa, F. Abbas, M. Ali, T. J. Esau And A. A. Farooque, "Artificial Intelligence Driven Smart Farming For Accurate Detection Of Potato Diseases: A Systematic Review," IEEE Access, Vol. 12, Pp. 193902-193922, 2024.
[22]. D. P. Hughes And M. Salathé, "An Open Access Repository Of Images On Plant Health To Enable The Development Of Mobile Disease Diagnostics," 2016.
[23]. S. Islam And T. Afrin, "Potatocare: Deep Learning Based Potato Disease Dataset," 2025.
[24]. [Online]. Available: Https://Www.Kaggle.Com/Datasets/Rizwan123456789/Potato-Disease-Leaf-Datasetpld.
[25]. N. H. Shabrina, S. Indarti, R. Maharani, D. A. Kristiyanti, I. N. Prastomo And T. A. M, "A Novel Dataset Of Potato Leaf Disease In Uncontrolled Environment," Data In Brief, Vol. 52, 2024.
[26]. V. Balafas, E. Karantoumanis, M. Louta And N. Ploskas, "Machine Learning And Deep Learning For Plant Disease Classification And Detection," IEEE Access, Vol. 11, Pp. 114352-114377, 2023.
[27]. A. Shrotriya, A. K. Sharma, S. Prabhu And A. K. Bairwa, "An Approach Toward Classifying Plant-Leaf Diseases And Comparisons With The Conventional Classification," IEEE Access, Vol. 12, Pp. 117379-117398, 2024.