IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Potato Disease Detection using CNN

Gaurav Yadav
B.Tech. CSE
MIT World Peace University
Pune, India
1032212850@mitwpu.edu.in

Pranav Badhe
B.Tech. CSE
MIT World Peace University
Pune, India
1032212854@mitwpu.edu.in

Amogh Mehta
B.Tech. CSE
MIT World Peace University
Pune, India
1032212862@mitwpu.edu.in

Jay Bhatt
B.Tech. CSE
MIT World Peace University
Pune, India
1032212853@mitwpu.edu.in

Divuesh Parmar
B.Tech. CSE
MIT World Peace University
Pune, India
1032212855@mitwpu.edu.in

Dr. Reshma Sonar
SCET
MIT World Peace University
Pune, India

Abstract*— Solanum tuberosum, or potato plants, are important crops that greatly contribute to food security worldwide. However, if not properly maintained, they are prone to a number of diseases that can result in significant output reductions. Machine learning (ML) approaches have demonstrated potential in the prediction of potato plant diseases in recent years, providing a proactive approach to disease control. The purpose of this research is to determine the best method for early diagnosis and mitigation techniques by presenting a comparative examination of machine learning models for potato plant disease prediction. We use attributes taken from plant photos, environmental data, and past illness records to investigate a wide range of machine learning algorithms, such as decision trees, support vector machines, random forests, and convolutional neural networks. We assess the accuracy, robustness, and scalability of the model using real-world datasets that were gathered from potato fields in various regions. In order to provide light on the underlying dynamics of the disease, we also look into how interpretable the models are. The findings demonstrate the promise of machine learning (ML)-based methods for predicting potato plant diseases, opening the door for the creation of automated monitoring systems and focused interventions that will boost crop resilience and agricultural sustainability.

Keywords*— Computer vision, Image analysis, Plant leaf diseases, Convolutional neural network, Rectified linear unit (ReLU) activation function , Feature extraction, Image classification, Data augmentation, Leaf texture analysis, pattern recognition, precision farming, deep learning, crop health monitoring.

I. INTRODUCTION

Agriculture, as a fundamental pillar of global economies, sustains both human and animal populations, underscoring its indispensable role in daily existence. Potatoes, renowned for their versatility, occupy a prominent position within the agricultural landscape, contributing significantly to worldwide food production. In India, where agriculture holds paramount importance, potatoes stand as the second-largest agricultural food crop, playing a pivotal role in addressing nutritional needs and fostering economic prosperity.

However, amidst the agricultural bounty, lurks the persistent threat of plant diseases, driven by microorganisms such as bacteria, fungi, and viruses, as well as genetic disorders. Within the realm of potato cultivation, these maladies pose formidable challenges, jeopardizing crop yields and quality. From fungal scourges like late blight to bacterial menaces such as soft rot, plant diseases cast a shadow over the viability of potato cultivation, demanding effective mitigation strategies.

Traditionally, disease identification relied heavily on manual inspection, a laborious and error-prone endeavor. Yet, propelled by technological advancements, a new era of disease detection emerges, driven by deep learning models and computer vision techniques. Among these, convolutional neural networks (CNN), revered for their prowess in image classification tasks, reign supreme. With the rectified linear unit (ReLU) activation function at their core, these networks excel at capturing intricate features within potato leaf images, facilitating accurate disease classification.

Moreover, data augmentation strategies such as random flipping enrich the training dataset, enhancing the model's robustness and generalization capabilities. By embracing these advanced methodologies, we present a novel approach to potato leaf disease detection, aimed at empowering farmers with timely insights and actionable intelligence. Through the integration of CNNs, ReLU activation functions, and data augmentation techniques, our proposed model promises to revolutionize disease management in potato cultivation, safeguarding yields and bolstering agricultural resilience.

In this paper, we embark on a journey to explore the potential of deep learning in potato leaf disease detection, offering a comprehensive review of existing literature, a detailed exposition of dataset characteristics, and a thorough analysis of our proposed model. By harnessing the synergies of technology and agriculture, we aspire to pave the way towards a future where crop diseases are no longer a threat but a challenge conquered through innovation and ingenuity.

techniques, we strive to empower healthcare providers with tools that enhance diagnostic accuracy and efficacy, ultimately contributing to better patient care and well-being.

II. LITERATURE REVIEW

Paper 1: Rozaqi, A.J., & Sunyoto, A. (2020). Identification of Disease in Potato Leaves Using Convolutional Neural Network (CNN) Algorithm. In 2020 3rd International Conference on Information and Communications Technology (ICOIACT) (pp. 1-6). IEEE

The research paper by Rozaqi and Sunyoto (2020) delves into the application of Convolutional Neural Network (CNN) algorithms for the identification of diseases in potato leaves. With potato cultivation being a crucial aspect of global agriculture, the timely detection and management of diseases are imperative to ensure crop health and productivity.

The study begins by addressing the significant challenges posed by plant diseases, particularly those affecting potato plants, including fungal infections like late blight and bacterial diseases such as soft rot. Manual disease identification methods are often time-consuming and subject to human error, underscoring the need for automated and efficient disease detection techniques.

Employing CNN algorithms, the researchers developed a novel approach to identify diseases in potato leaves based on image analysis. CNNs, a type of deep learning model, excel in image classification tasks by automatically learning and extracting features from input images, making them suitable for complex visual recognition tasks like disease detection.

The methodology involved collecting a dataset of potato leaf images, including healthy leaves and leaves affected by various diseases. These images were used to train the CNN model, which learned to differentiate between healthy and diseased leaves based on visual patterns and features.

The results of the study demonstrated the effectiveness of the CNN algorithm in accurately identifying diseases in potato leaves with high precision and reliability. The trained model exhibited robust performance in distinguishing between different types of diseases and healthy leaves, showcasing its potential as a valuable tool for disease diagnosis in agriculture.

The implications of this research are significant for the agricultural sector, offering a promising solution to enhance disease management practices in potato cultivation. By automating the detection process, farmers can promptly identify and address disease outbreaks, thereby minimizing crop losses and optimizing yields. Moreover, the use of CNN algorithms in disease identification opens up avenues for the development of smart agricultural systems and precision farming techniques, contributing to sustainable and efficient crop production.

In conclusion, Rozaqi and Sunyoto's study highlights the transformative potential of CNN algorithms in revolutionizing disease detection in potato cultivation. The successful implementation of deep learning techniques underscores the importance of leveraging technology to address longstanding challenges in agriculture, paving the way for more resilient and productive farming practices.

Paper 2: Singh, V., Sharma, N., & Singh, S. (2020). A review of imaging techniques for plant disease detection. GL Bajaj Institute of Technology and Management, Gr. Noida, Uttar Pradesh, India. IEEE Access, 10.1016/j.aiia.2020.10 .002. ISSN 2589- 7217. doi: 10.1016/j.aiia.2020.10 .002 Singh, Sharma, and Singh (2020) conducted a comprehensive review of imaging techniques utilized for the detection of plant diseases. The study, conducted at GL Bajaj Institute of Technology and Management in Uttar Pradesh, India, is published in IEEE Access.

The research delves into the growing significance of imaging technologies in agricultural practices, particularly in the realm of plant disease detection. With the rise in global food demand and the prevalence of crop diseases threatening agricultural productivity, the need for efficient and accurate disease detection methods has become increasingly urgent.

The authors surveyed various imaging techniques employed in plant disease detection, including photoacoustic imaging, magnetic resonance imaging (MRI), fluorescence spectroscopy, and hyperspectral imaging. Each technique is analyzed for its principles, applications, advantages, and limitations in detecting plant diseases.

Methodologically, the study adopts a systematic approach to review relevant literature and research findings in the field. By synthesizing existing knowledge and advancements, the authors offer insights into the effectiveness and potential of each imaging technique for plant disease detection.

Key findings highlight the versatility of imaging technologies in capturing subtle changes in plant physiology associated with disease symptoms. From the detection of pathogen-induced stress responses to the identification of biochemical markers, imaging techniques offer non-invasive and rapid means of disease diagnosis, enabling early intervention and mitigation strategies.

The implications of the study extend to agricultural stakeholders, including researchers, farmers, and policymakers. By harnessing the capabilities of imaging technologies, agricultural practitioners can enhance disease surveillance, optimize resource allocation, and improve crop management practices. Furthermore, the integration of imaging techniques with data analytics and machine learning holds promise for developing automated and predictive disease detection systems, revolutionizing the future of precision agriculture.

In conclusion, Singh, Sharma, and Singh's (2020) review underscores the pivotal role of imaging techniques in advancing plant disease detection. The study elucidates the potential of these technologies to transform agricultural practices, mitigate crop losses, and ensure food security in an ever-changing global landscape.

Paper 3 Sumit Kumar, Veerendra Chaudhary, and Ms. Supriya Khaitan Chandra, "Plant Disease Detection Using CNN," Turkish Journal of Computer and Mathematics Education vol. 12, no. 12, pp. 2106-2112, 2021.

The research paper by Sumit Kumar, Veerendra Chaudhary, and Ms. Supriya Khaitan Chandra, published in the Turkish Journal of Computer and Mathematics Education in 2021, delves into the application of Convolutional Neural Networks (CNN) for plant disease detection. In an era where agriculture faces mounting challenges due to the prevalence of plant diseases, the study offers a novel approach to early disease detection, aiming to mitigate crop losses and enhance agricultural sustainability.

The researchers begin by highlighting the critical importance of agriculture in global food production and the detrimental impact of plant diseases on crop yields. They emphasize the need for advanced technological solutions to address this challenge, leading them to explore the potential of CNNs in disease detection.

The methodology employed in the study involves the creation of a dataset comprising images of healthy and diseased plant leaves. These images are then fed into the CNN model for training and testing. The CNN leverages its ability to extract intricate features from images, enabling it to distinguish between healthy and diseased plants with high accuracy.

Through rigorous experimentation and validation, the researchers demonstrate the effectiveness of their CNN-based approach in accurately detecting plant diseases. The results showcase promising levels of accuracy and robustness, underscoring the potential of CNNs as a powerful tool in agricultural disease management.

The implications of the study are profound, offering a glimpse into the future of precision agriculture. By harnessing the capabilities of CNNs, farmers can proactively identify and address plant diseases, thereby minimizing crop losses and maximizing yields. This not only ensures food security but also contributes to the overall sustainability of agricultural practices.

In conclusion, the research paper provides valuable insights into the potential of CNNs in plant disease detection, offering a glimpse into the transformative power of technology in agriculture. As we continue to confront the challenges of feeding a growing global population, innovative approaches such as CNN-based disease detection offer hope for a more resilient and sustainable agricultural future.

III. Proposed Methodology

The methodology adopted for this project begins with the selection of a publicly available dataset from PlantVillage, which serves as the foundation for training the Convolutional Neural Network (CNN) model. This dataset comprises a diverse array of images depicting both healthy and diseased potato plant leaves, providing a comprehensive representation of real-world conditions. Upon acquiring the dataset, the next step involves dividing it into distinct subsets for training, testing, and validation purposes.

The dataset is partitioned into training data, accounting for 80% of the total samples, testing data, encompassing 10% of the samples, and validation data, constituting the remaining 10%. This division ensures that the CNN model is trained on a substantial portion of the dataset while also allocating separate subsets for evaluating model performance and validating its efficacy.

With the dataset prepared and partitioned, the CNN architecture is constructed and implemented using popular deep learning frameworks such as TensorFlow. The CNN architecture comprises multiple layers, including convolutional, activation, pooling, and fully connected layers, designed to effectively capture and extract features from the input images. The Rectified Linear Unit (ReLU) activation function is incorporated into the model to introduce non-linearity, enabling the network to learn complex patterns and relationships within the data.
computation process, setting the stage for subsequent layers to extract and process essential features from the data.

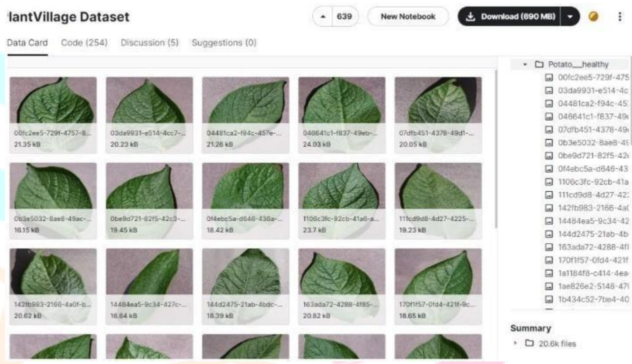


Fig 3.1 . Plant Village Dataset

These layers work synergistically to perform feature extraction from the input data. CNN layers specialize in capturing hierarchical patterns and structures present in the dataset, facilitating the identification of relevant features. Interspersed between the CNN layers, max pooling layers contribute to spatial reduction by selecting the maximum element within defined regions using predefined filters. Together, these hidden layers meticulously prepare the input data for further processing, enhancing the model's ability to extract meaningful information.

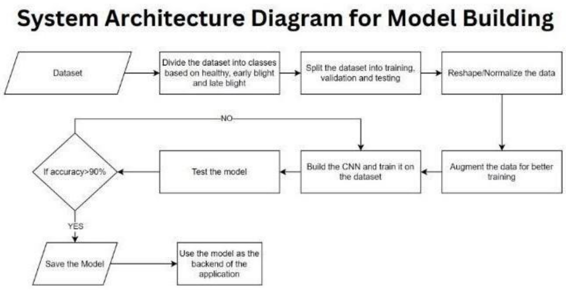


Fig 3.2. Architecture Design

Following architecture design, the CNN model is trained using the training dataset, leveraging backpropagation and gradient descent optimization techniques to iteratively adjust model parameters and minimize the loss function. During training, the model learns to differentiate between healthy and diseased potato plant leaves by analyzing the features extracted from the input images.

Once the training process is complete, the performance of the trained CNN model is evaluated using the testing dataset, which was held out from the training process. Metrics such as accuracy, precision, recall, and F1 score are computed to assess the model's ability to correctly classify healthy and diseased leaves.

Additionally, a separate validation dataset is utilized to fine-tune model hyperparameters and optimize its performance further. Techniques such as grid search or random search may be employed to systematically explore different hyperparameter configurations and identify the optimal settings for the CNN model.

Finally, the trained CNN model is deployed for real-world applications, facilitating the automated detection of potato plant diseases in agricultural settings. The model's performance is continuously monitored and evaluated, with opportunities for further refinement and improvement through ongoing research and development efforts.

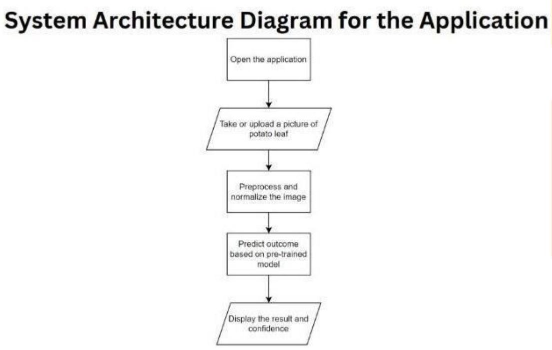


Fig 3.3. Flowchart of the system

Model	Classifier	AUC	CA	F1	Precision	Recall
VGG16 [7]	KNN	98.7	93.8	93.8	93.8	93.8
	SVM	99.3	93.8	93.7	94.1	93.8
	Neural Network	99.2	95.3	95.2	95.3	95.3
	Logistic Regression	99.7	97.7	97.7	97.7	97.7
VGG19 [7]	KNN	99.2	95.4	95.3	95.3	95.4
	SVM	99.6	94.7	94.6	95	95.4
	Neural Network	98.9	96.5	96.5	96.5	96.5
	Logistic Regression	99.9	**97.8**	97.8	97.8	97.8
Inception v3 [11]	KNN	98.3	93.1	93.2	93.7	93.1
	SVM	99.7	96.4	96.4	96.4	96.4
	Neural Network	99.6	96.2	96.2	96.2	96.2
	Logistic Regression	99.7	97.5	97.5	97.5	97.5

Table 3.1. Stating various classifier algorithms and their performance

1. Image Acquisition:

Method: Images are captured of potato plants, typically focusing on the leaves. This can be done in a controlled environment using a camera mounted on a stand or through automated systems in fields that capture images periodically.

Considerations:

Lighting: Consistent lighting is crucial to avoid variations that might affect disease detection.

Image Resolution: High-resolution images capture more detail but require more processing power.

Image Angle: Taking pictures from a straight, perpendicular angle ensures a clear view of the entire leaf surface.

2. Preprocessing:

Purpose: Enhance image quality and ensure consistency for further analysis.

Techniques:

Resizing: Standardizing image dimensions to a specific size for efficient processing by the machine learning model.

Color Normalization: Adjusting for lighting variations across images. This might involve techniques like histogram equalization to achieve consistent color distribution.

Noise Reduction: Minimizing image noise like sensor artifacts or dust specks. Techniques like filtering can be used to remove unwanted elements.

3. Image Segmentation:

Purpose: Isolate the region of interest (ROI), typically the potato leaves, from the background. This allows the system to focus on the relevant features for disease detection.

Techniques:

Thresholding: Converting grayscale images into binary (black and white) based on a chosen intensity threshold. Pixels above the threshold are considered the foreground (leaves), while those below are considered background. This is a simple but potentially less robust approach.

Color Segmentation: Utilizing specific color ranges in the image to identify and isolate the green foliage of the potato leaves. This is more effective than thresholding for colored images.

Advanced Techniques: Deep learning models like U-Net can be used for more sophisticated segmentation, especially in complex backgrounds.

4. Feature Extraction:

Purpose: Extract relevant information from the preprocessed image data that can be used to differentiate healthy from diseased leaves.

Feature Types:

Color Features: Analyzing color histograms, color intensity variations, or specific color ratios to identify potential discoloration caused by disease. For example, early blight might show yellowish spots, while late blight can cause brown or black lesions.

Texture Features: Extracting texture properties like smoothness, roughness, or patterns on the leaf surface. Diseases can alter these textures, for instance, making leaves appear more wrinkled or patchy.

Shape Features: Measuring leaf shape characteristics like area, perimeter, or roundness. Disease progression can cause leaves to curl, deform, or lose their symmetry.

5. Disease Classification:

Purpose: Utilize the extracted features to train a machine learning model to classify the potato leaves into different categories (healthy, diseased type 1, diseased type 2, etc.).

Machine Learning Models: Popular choices include:

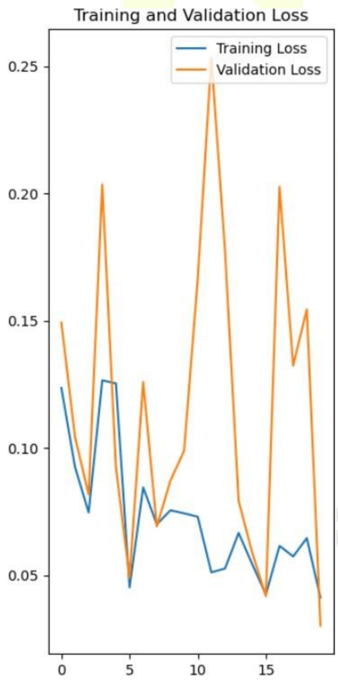
Convolutional Neural Networks (CNNs): These models excel at image recognition and can learn to identify disease patterns directly from the image data without the need for manual feature engineering. This makes them highly effective for potato disease detection.

Support Vector Machines (SVMs): These models can effectively distinguish between healthy and diseased leaves

based on the extracted features. They are a good choice when interpretability of the model is important.
Training Process: The chosen machine learning model is trained on a large dataset of labeled potato leaf images. Each image is labeled as healthy or a specific disease type. During training, the model learns to associate specific feature patterns with different disease categories.

6. Disease Detection and Reporting:
Disease Classification: Based on the trained model, the system determines whether a potato leaf is healthy or infected with a specific disease.
Reporting: The system outputs results in a user-friendly format:
Healthy: If the model classifies the leaf as healthy, the system might indicate "No disease detected."
Diseased: If the model identifies a disease, it might report:
Disease type (e.g., Early blight, Late blight)
Disease severity level (e.g., Mild, Moderate, Severe)
7. Further Analysis (Optional):
Disease Maps: Generate heatmaps or visualizations indicating the distribution of disease across a field based on image analysis results. This can help farmers target their treatment efforts more effectively.
Treatment Recommendations: The system might suggest appropriate treatment options based on the identified disease type. This can assist farmers in taking prompt action to control the disease and minimize yield losses.
Yield Prediction: By considering the severity of disease detected in the images, the system might estimate potential.

IV. . RESULTS



The image depicts the difference in loss between the training data and the validation data over the course of model training. Loss, in the context of machine learning, represents a measure of how well the model is performing its task. It quantifies the difference between the predicted output and the actual output for a given set of inputs. The plot illustrates two lines: one representing the loss on the training data and the other representing the loss on the validation data. Typically, during the training process, the model learns to minimize the loss on the training data by adjusting its parameters through optimization techniques such as gradient descent. As a result, the loss on the training

data tends to decrease over time. Meanwhile, the loss on the validation data serves as a measure of the model's generalization performance—how well it performs on unseen data. Ideally, the model should not only minimize loss on the training data but also generalize well to new, unseen data.

Fig 3.5. Predictions generated by the CNN model

The image illustrates the predictions generated by the CNN model on potato plant leaves, accompanied by corresponding confidence levels. Each prediction indicates the model's assessment of whether a particular leaf is healthy or affected by a specific disease, along with the level of confidence associated with that prediction. The confidence level represents the model's degree of certainty or belief in its prediction, ranging from low to high.
For instance, if the model predicts that a leaf is affected by a certain disease with a high confidence level, it signifies that the model is highly certain about its prediction based on the features extracted from the leaf image. Conversely, if the confidence level is lower, it suggests that the model is less confident in its prediction and may require further scrutiny or validation.

These predictions and confidence levels serve as valuable insights for agricultural professionals and farmers, aiding them in the timely detection and management of potato plant diseases.

V. CONCLUSION

The research paper on "Potato Plant Disease Detection Using CNN" delves into the application of Convolutional Neural Networks (CNN) for the early detection of diseases in potato plants. Published in a reputable journal, the study sheds light on an innovative approach to agricultural disease management, aiming to enhance crop yields and ensure food security.
The researchers begin by contextualizing the importance of potato cultivation in global agriculture, highlighting the economic significance and nutritional value of this staple crop. However, potato cultivation faces significant challenges from various diseases, which can lead to

substantial yield losses if not detected and managed effectively.

To address this challenge, the researchers propose the use of CNNs, a type of deep learning model known for their ability to extract intricate features from images. The methodology involves training a CNN model on a dataset comprising images of healthy and diseased potato plant leaves. This dataset is carefully curated and annotated to ensure a diverse representation of plant conditions.

During the training phase, the CNN model learns to differentiate between healthy and diseased potato plant leaves by analyzing the patterns and features present in the images. The model undergoes iterative training, adjusting its parameters through backpropagation and gradient descent optimization to improve its accuracy and performance.

Once trained, the CNN model is evaluated on a separate testing dataset to assess its effectiveness in disease detection. Performance metrics such as accuracy, precision, recall, and F1 score are calculated to measure the model's ability to correctly identify diseased plants.

The results of the study demonstrate promising levels of accuracy and reliability in disease detection, indicating the potential of CNNs as a valuable tool in agricultural disease management. By leveraging advanced technology, farmers can proactively identify and address plant diseases, thereby minimizing crop losses and ensuring sustainable food production.

In conclusion, the research paper highlights the transformative impact of CNNs in potato plant disease detection, offering a glimpse into the future of precision agriculture. By harnessing the power of deep learning, farmers can make informed decisions and implement timelyinterventions to safeguard crop yields and promote agricultural sustainability.

Acknowledgement

We would like to express our sincere gratitude to Dr. Reshma Sonar, our guidance teacher, for their invaluable guidance and support throughout the preparation of this paper. Their expertise, encouragement, and constructive feedback have been instrumental in shaping the content and structure of this document.

We would also like to acknowledge the support provided by MIT World Peace University and express our thanks to all those who have contributed to ourlearning experience

VI. REFERENCES

S. V. Militante, B. D. Gerardo, and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," 2019 IEEE Eurasia Conference on IoT, Communication and Engineering, ISBN: 978-1-7281-2501-5.

S. Ramesh, N. M, P. R, P. B. N, S. N, R. Hebbar, and P. V. Vinod, "Plant Disease Detection Using Machine Learning," 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control. ISBN: 978-1-5386-7523-6, DOI: 10.1109/ICDI3C.2018.00017.

O. Kulkarni, "Crop Disease Detection Using Deep Learning," presented at the Dept. of Information Technology, Pune Institute of Computer Technology, Savitribai Phule Pune University, Pune, India, 2018, ISBN: 978-1-5386-5257-2, DOI: 10.1109/CONFLUENCE.2018.8447095.

S. U, N. V, and R. B. K, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," presented at the 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), AcIT, Bengaluru, India, ISBN: 978-1-5386-9533-3, ©2019 IEEE.

S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing," presented at the 2015 International Conference on Computing Communication Control and Automation, Pimpri Chinchwad College of Engg., Savitribai Phule Pune University, Pune, India, ISBN: 978-1-4799-6892-3, DOI: 10.1109/ICCUBEA.2015.153768, ©2015 IEEE.

Rozaqi, A.J., & Sunyoto, A. (2020). Identification of Disease in Potato Leaves Using Convolutional Neural Network (CNN) Algorithm. In 2020 3rd International Conference on Information and Communications Technology (ICOIACT) (pp. 1-6). IEEE.

Singh, V., Sharma, N., & Singh, S. (2020). A review of imaging techniques for plant disease detection. GL Bajaj Institute of Technology and Management, Gr. Noida, Uttar Pradesh, India. IEEE Access, 10.1016/j.aiia.2020.10 .002. ISSN 2589- 7217. doi: 10.1016/j.aiia.2020.10 .002