Disease Detection in Potato Plants Using Deep Learning Applied on Leaf Image

Ayushi Godiya[1], Dr. Abhay Kothari[2]

[1] Research Scholar, Department of Computer Science and Engineering, Sage University Indore M.P. India
29ayushi91@gmail.com
[2] Professor, Department of Computer Science and Engineering, Sage University Indore M.P. India
abhaykothari333@gmail.com

Abstract: The paper focuses on the detection of diseases in potato plants using deep learning applied to leaf images. It discusses the challenges faced by the agriculture sector in India and emphasizes the need for innovative solutions. The paper highlights the importance of potato cultivation in India, including its adaptability to various climates and its role as a crucial food source. Furthermore, the paper explores the use of convolutional neural networks (CNNs) for disease detection in potato plants. It describes the process of data collection, including the acquisition of a dataset containing both healthy and diseased potato leaf images. The paper discusses different methods of data collection, such as obtaining readymade data from third-party vendors or using a team of annotators to collect and classify images manually. The paper also explains the image acquisition process and the application of CNNs for feature extraction and classification. It discusses the use of convolution and pooling layers, as well as the activation function ReLU, in the CNN architecture. The paper presents the training and testing accuracy results of the proposed CNN model, demonstrating its effectiveness in accurately classifying healthy and diseased potato leaves. In conclusion, the paper highlights the potential of deep learning techniques, specifically CNNs, for disease detection in potato plants. It suggests that these techniques can contribute to improving the efficiency and productivity of the agricultural industry. The findings of the paper provide valuable insights for researchers and practitioners in the field of agricultural technology.

Keywords: Deep learning, Convolutional Neural Networks (CNNs), Image acquisition, Convolution, Pooling layers, ReLU activation function, Model training, Productivity, Efficiency.

1. Introduction

Indian agriculture has a long history, dating back 11,000 years. It's been mentioned in ancient texts like the Vedas and Mahabharata. Over time, we've made great strides in food production, becoming a top producer of various agricultural products. However, we face challenges today, like climate change, a growing population, and income disparities. To tackle these issues, we're using advancements in science and technology. Agriculture is crucial for India's economy, providing employment for almost 50% of the workforce and contributing 17% to our GDP. While we've made progress, we've also experienced difficulties, including famines due to droughts. India's climate and water resources vary greatly, and we heavily rely on monsoons for water.

1.1. Agricultural research in India Agriculture is possibly the oldest sector in the country to have research as its core segment. There is evidence of research and development in agriculture in the ancient as well as medieval periods of Indian history.

Table 1. Production of agricultural commodities and cultivated area in the country in 1950-51 and 2021-22

Commodity	1950-51	2021-22	Times increase
Food grains (Mt)	51	314	6.2
Vegetables & fruits (Mt)	25	333	13.3
Milk (Mt)	17	210	12.4
Egg (billion)	1.8	122	67.8
Fish (Mt)	0.8	14.2	17.8
Net sown area (Mha)	130	140	1.1
Gross sown area (Mha)	150	198	1.3

"The government of India has taken several steps for the development of agriculture and welfare of the farmers. Emphasis has been laid on promoting agriculture sector through modern technologies to transform agricultural practices, and increase productivity and profitability. As seen in the table below, Agriculture has higher share of total budget of Government vis-a-vis other Departments [1, 2].

Table 2. Allocation of Budget for Various Ministries/Departments (2020-2023)

Ministry/Deptt. Name	2023-24	2020-21	2021-22	2022-23
Department of Agriculture, and Farmers Welfare	134399.77	123017.57	124000.00	115531.79
% Share of Deptt. w.r.t Total Central Plan	4.41	3.53	3.14	2.57
Department of School Education and Literacy	59845.00	54873.66	63449.37	68804,85
% Share of Deptt. w.r.t Total Central Plan	1.96	1.57	1.61	1.53
Department of Health and Family Welfare	65011.80	71268.77	83000.00	86175.00
%Share of Deptt. w.r.t Total Central Plan	2.13	2.04	2.10	1.91
Ministry of Housing and Urban Affairs	50039.90	54581.00	76549,46	76431.60
% Share of Deptt. w.r.t Total Central Plan	1.64	1.56	1.94	1.70
GOI'S Total Budget Outlay	3042230.09	3483235.63	3944908.67	4503097.45

2. Literature Review:

Table 3. Literature review

Author(s)	Year	Main Idea	Conclusion
Shilpa S. Selvan et al.	N/A	Highlight the challenges facing the agriculture sector in India	Recent pandemic situation has affected the agricultural sectors, which needs to be discussed and planned [3]
Rupinder Kaur and	N/A	Analyze the debt position of marginal and small farmers in rural Haryana	Majority of marginal and small farmers in rural Haryana are under debt [4]

Karamjeet Kaur			
Shahid Jibran and Azra Mufti	2019	Identify the major issues and challenges of Indian agriculture and propose possible solutions	Indian agriculture is facing several issues and challenges that require attention and solutions [5]
D Hebsiba Beula et al.	2021	Discuss the dependence of Indian agriculture on uncertain factors and the need for crop insurance	Crop insurance can be an alternative to manage production risk in Indian agriculture [6]
Anirudha Vachaspati Vempati et al.	2020	Study the challenges faced by an impoverished community in Bihar, India and propose technology-based solutions	Technological innovations can address resource poverty and contribute to poverty alleviation and sustainability [7]
Haiguang Wang et al.	2012	Explore the application of neural networks in image recognition of plant diseases	Neural networks can effectively identify and diagnose plant diseases based on image processing [8]
Sammy V. Militante et al.	2019	Develop a deep learning-based system for detecting and recognizing plant diseases	The system achieved high accuracy in detecting and recognizing plant diseases and plant varieties [9]

2.1. Challenges in Indian Agriculture

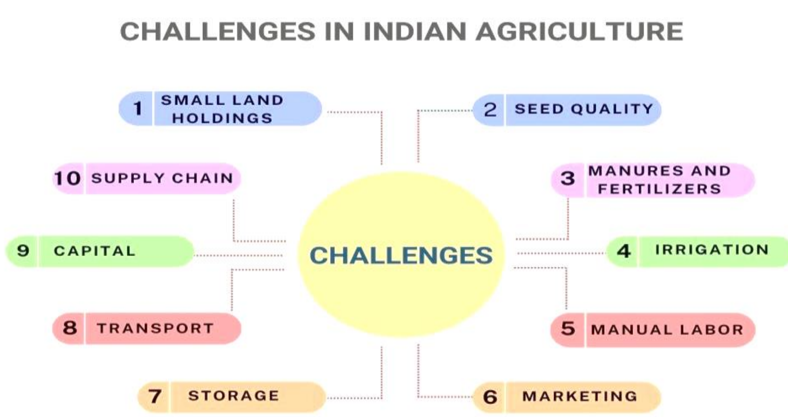


Figure 1. Solving these challenges is crucial for the future of Indian agriculture [10].

Table 4. Types of potato found in India

Kufri Sindhuri	Bihar, Gujarat, Maharashtra, Punjab, Uttar Pradesh, Karnataka, and Jammu Kashmir are the leading producers of this potato cultivar.	Early blight resistance is moderate, and it can withstand the Potato Leaf Roll Virus. This potato has an average yield of 40 tonnes per acre [11].
Kufri Chandramukhi	Bihar, Gujarat, Haryana, Punjab, Madhya Pradesh, Himachal Pradesh, Uttar Pradesh,	This potato has an average yield of 25 tonnes per acre [12].

	Orissa, and West Bengal are the main producers.	
Kufri Jyoti	Bihar, Maharashtra, Gujarat, Haryana, Punjab, Uttar Pradesh, Karnataka, and West Bengal are the main producers.	This potato has an average yield of 20 tonnes per acre. Early and late blight resistance is moderate in this potato.
Kufri Lauvkar	Mainly grown in Maharashtra, Madhya Pradesh and Karnataka.	Yields faster in warm climates. Average yield is about 30 tonnes per hectare.
KufriBadshah	Mostly grown in Jammu Kashmir, Haryana, Punjab, Gujrat, Uttar Pradesh and Madhya Pradesh.	Average yield is about 50 tonnes per hectare.
KufriBahar	Mostly grown in Haryana, Uttar Pradesh, Himachal Pradesh and Jammu Kashmir.	Average yield of this potato is about 45 tonnes per hectare.
KufriLalima	This potato crop is mostly grown in Uttar Pradesh and Bihar.	average yield is about 40 tonnes per hectare. This potato is moderately resistant to early blight.
Kufri Jawahar	Mainly grown in Haryana, Punjab, Gujrat, Madhya Pradesh and Karnataka.	average yield is about 40 tonnes per hectare. This potato is moderately resistant to late blight disease.
Kufri Sutlej	This crop is mostly grown in Bihar, Uttar Pradesh, Haryana, Punjab and Madhya Pradesh.	yield is about 40 tonnes per hectare. This potato is moderately resistant to late blight disease.
Kufri Ashoka	Mostly grown in Bihar, Punjab, Haryana, Uttar Pradesh and West Bengal.	average yield is about 40 tonnes per hectare.
Kufri Pukhraj	Mostly grown in Bihar, Punjab, Haryana, Uttar Pradesh, Maharashtra, Himachal Pradesh, Madhya Pradesh, Orissa and West Bengal.	average yield is about 40 tonnes per hectare. This potato is resistant to early bight disease and moderately resistant to late blight disease.
Kufri Chipsona-1	Mostly grown in Bihar and Uttar Pradesh.	average yield is about 40 tonnes per hectare. This potato is resistant to late blight disease and tolerant to frost.
Kufri Chipsona-2	Mostly grown in Bihar and Uttar Pradesh.	average yield is about 35 tonnes per hectare. This potato is resistant to late blight disease and tolerant to frost.
Kufri Anand	Mostly grown in plains Bihar and Uttar Pradesh.	yield is about 35-40 tons per hectare.

Potatoes are unique because they can grow from a piece of another potato, producing genetic clones. They can also grow from seeds to create genetically different tubers. Potatoes can thrive in various climates, from sea level to high mountains. They are incredibly efficient, producing more food per unit of water compared to other major crops. Potatoes are grown in over 100 countries and have become a crucial food source, especially in South America, Africa, and Asia. In recent decades, potato production in developing countries has outpaced other food crops, making it a vital source of food security for millions of people. [13]

2.1.1. Nutrition, Medicinal Properties and It's Uses

A raw potato is mostly water, with carbohydrates (mainly starch) making up most of its content. It's low in fat and has some protein. It's a good source of vitamin B6 and vitamin C. Cooking reduces vitamin levels but doesn't change other

---

nutrients much. Potatoes are often considered high on the glycemic index, so they're avoided by some on low-GI diets. However, the GI varies based on factors like the type of potato, how it's prepared, and what you eat with it. Reheating and cooling potatoes can create resistant starch, lowering their GI. While potatoes have lost popularity due to low-carb diets, they are still nutritious, providing fiber, vitamins, minerals, and other health benefits.

Table 5. Nutritional Importance of Potato Plant

Health Benefits	Importance
Blood pressure	● Potassium, calcium, and magnesium are all present in the potato. These have been found to decrease blood pressure naturally.
Heart health	● Potatoes contain significant amounts of fiber. Fiber helps lower the total amount of cholesterol in the blood, thereby decreasing the risk of heart disease. ● The potato's fiber, potassium, vitamin C, and vitamin B6 content, coupled with its lack of cholesterol, all support heart health.
Inflammation	● maintaining the structure of cellular membranes ● transmitting nerve impulses ● the absorption of fat ● early brain development
Metabolism	● Potatoes are a great source of vitamin B6. This plays a vital role in energy metabolism, by breaking down carbohydrates and proteins into glucose and amino acids.
Immunity	● Research has found that vitamin C may help reduceTrusted Source the severity and duration of a cold. Potatoes are a good source of vitamin C.

2.1.2. Potato Plant Dieses

Table 6. There are three distinctive characteristics of plant disease images taken in real-world scenarios [14].

Disease Name	Identification Characteristics	Detection Techniques	Possible Harmful Effects
Late Blight	Dark, water-soaked lesions on leaves and stems, white fungal growth on undersides of leaves	Visual inspection, laboratory testing	Reduced yield, rotting tubers
Early Blight	Brown lesions with concentric rings on leaves, stems, and tubers	Visual inspection, laboratory testing	Reduced yield, premature defoliation
Blackleg	Black, soft rotting stems, wilting plants	Visual inspection, laboratory testing	Stunted growth, plant death
Potato Virus Y	Mottled or yellowing leaves, stunted growth	Serological tests, PCR	Reduced yield, tuber deformities
Potato Cyst Nematode	Stunted growth, yellowing leaves, cysts on roots	Soil sampling, laboratory testing	Reduced yield, root damage

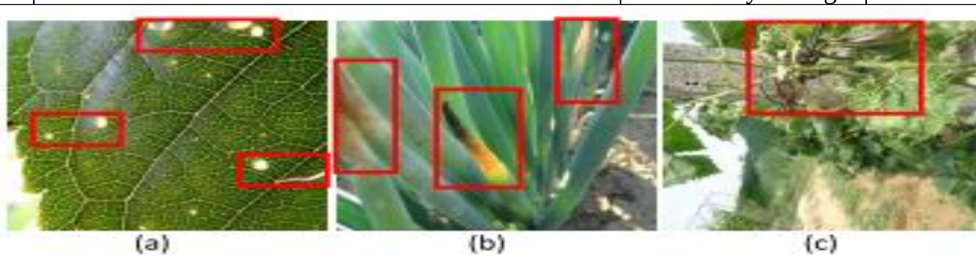


Figure 2. Sample of Plant Disease Characteristics

3. Different Plant Disease Detection Technique

Manual disease detection through human experts to identify and recognize plant diseases is a usual practice in agriculture. With the improvements in technology, automatic detection of plant diseases from raw images is possible through computer vision and artificial intelligence researches. Currently, the introduction of deep learning methods turns out to be popular. Deep learning is the advanced methods of machine learning that uses neural networks that works like the human brain. Traditional methods involve the use of semantic features as the classification method. The methodology in the study involves three key stages: acquisition of data, pre-processing of data and image classification. Image pre-processing involves re-sized images and enhancement before supplying it for the classification model [15].

4. Proposed work

4.1. Data Collection: For our project we required the dataset of potato plant leaf images. Which include healthy as well as the diseased images.

There are basically tree types of data collection

1. Readymade data: We can but the dataset from the third-party vendors or to download it from online platform like kaggle.
2. From the team of annotators: The job of the annotators to collect the data images from framers and annotates those images either as healthy potato images or having disease. This team will go to the field and ask the farmers to take the pictures or they can take the pictures themselves and then they can classify with the help of farmer or by some domain knowledge. Which means this a method to collect the data manually? This option is expensive and required budge.
3. Web scrapping scripts to collect images from internet: For this we need to go to the different website which have potato images and collect those images and use the tool like docano. There are so many tools are available which can help you to annotate the data.

For our research we are going to use the kaggle dataset which is freely available on their website with the classification of healthy and disease plant images.

4.2. **Image Acquisition:** Image dataset used for training the model was acquired in the Plant Village repository [15]. A python script was used to download images of the plant diseases from the repository. The acquired dataset consists of approximately 35,000 images with 32 different classes plant varieties and diseases.

4.3. **Image Pre-processed:** images are reduced image size and image crop to a given input. It processes and enhances the image to its needed color scale. The study uses colored and resized images to 96x96 resolution for processing.

4.4. **Classification:** Classification uses fully connected layers and for feature extraction it uses convolutional and pooling layers. The classification process classifies the plant leaf if it is infected with the disease or not, identifies the type of plant disease and recognize the plant variety.

5. Methodology:

The CNN architecture comprises several key elements, including convolutional layers, pooling layers, and fully connected layers, among others. In the architectural layout, one or more fully connected layers are typically followed by a series of convolution layers and a pooling layer. The process by which input data undergo transformations into output at various levels is known as forward propagation. It's worth noting that the convolution and pooling techniques

---

described in this section are not limited to just two-dimensional (2D) CNNs; they are also applicable to three-dimensional (3D) CNNs [16].

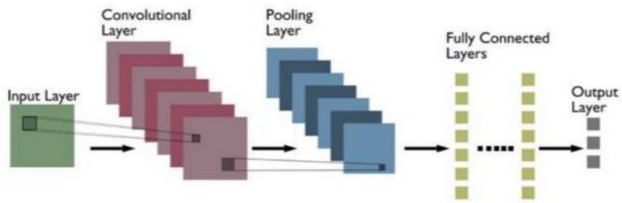


Figure 3. CNN Architecture Model

5.1. Convolution layer

A specific linear operation known as convolution plays a pivotal role in feature extraction. In convolution, a tensor, essentially an array of numbers, is convolved across a small array of numbers termed a kernel. Within the CNN architecture, the convolution layer is a fundamental component responsible for conducting this feature extraction process. By performing element-wise multiplication between each element of the kernel and the input tensor, a feature map is generated. The output value at the corresponding position in the output tensor is computed by summing up these products [17].

This process marks the initial stage of isolating essential components from an image. In the convolution layer, a set of filters is utilized to carry out convolution. Each image can be visualized as a matrix of pixel values. For instance, consider a 5x5 image where each pixel is represented by either 0 or 1. Additionally, a 3x3 filter matrix is introduced. To obtain the convolved feature matrix, the dot product is calculated by sliding the filter matrix over the image.

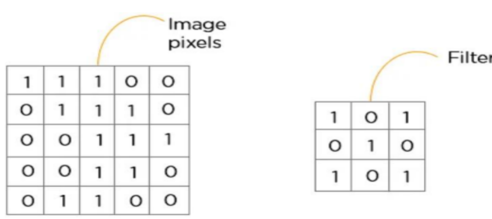


Figure 4. Working of Convolution Layer

- **ReLU Activation Function**: The acronym "ReLU" stands for "Rectified Linear Unit." After feature maps have been generated, they are then passed on to a ReLU layer. ReLU functions on a per-component basis, converting each negative pixel value to 0. This introduces non-linearity into the network, resulting in a rectified feature map as the final output. Below is a graphical representation of a ReLU function. The process involves applying various convolutions and ReLU layers to the original image to uncover distinctive features. Graphically, ReLU can be expressed as R(z) = max(0, z), where "z" represents the input value [18].

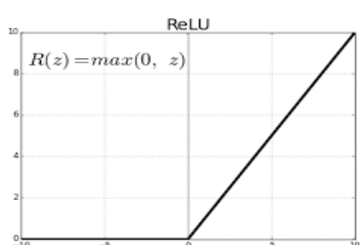
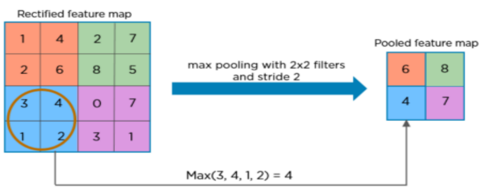


Figure 5. Activation Function ReLU

3.3.1.2 Pooling layer

A pooling layer employs a traditional down sampling technique to reduce the in-plane dimensionality of feature maps. This process serves multiple purposes, including providing translation invariance against small shifts and distortions and reducing the number of learnable parameters. It's important to note that pooling layers do not possess any learnable parameters themselves, although they utilize hyperparameters like filter size, stride, and padding in a manner similar to convolutional layers [19].

During the downsampling phase of pooling, the dimensionality of the feature map is reduced. The resulting pooled feature map is generated by combining the updated feature map with a pooling layer. To identify distinctive features within the image, such as corners and edges, the pooling layer employs various filters.



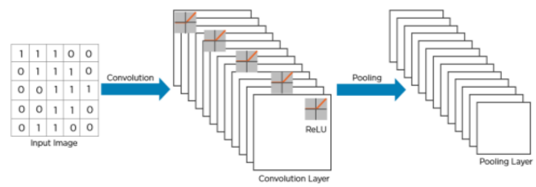


Figure 7. Working of Pooling Layer

Figure 8. Structure of CNN

The subsequent step in the process is known as "flattening." During this phase, the 2-Dimensional arrays obtained from the pooled feature maps are transformed into a single, elongated, continuous linear vector. This flattened matrix serves as the input for the fully connected layer, which is responsible for the task of classifying the image [20].

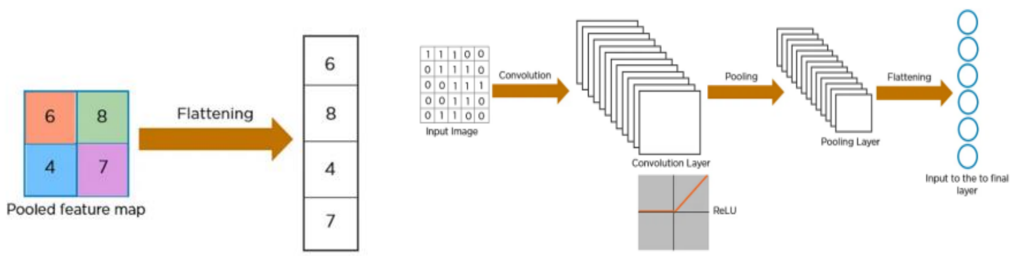
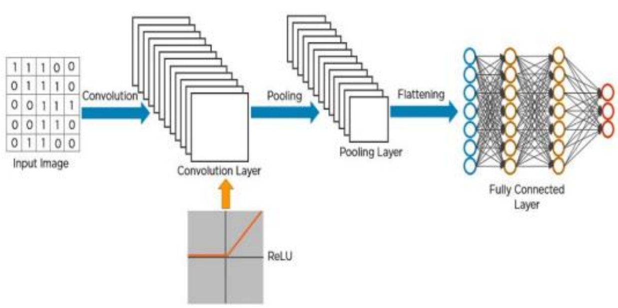


Figure 9. Process of Flattering



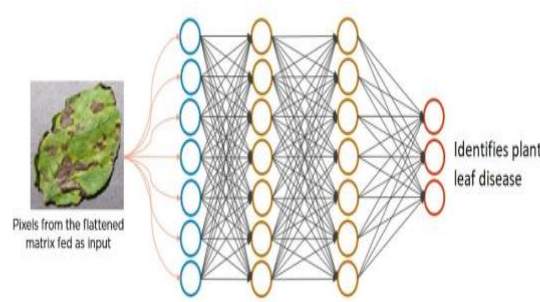


Figure 10. Structure of CNN with Flattering

Figure 11. CNN Process

Step 1: In the first step, the image's pixels are inputted into the convolutional layer, which performs the convolution operation.

Step 2: This operation yields a convolved map as the result.

Step 3: The convolved map is then processed using a ReLU function, generating a refined feature map.

Step 4: Multiple convolutions and ReLU layers are applied to the image to identify various features.

Step 5: In the fifth step, multiple pooling layers equipped with different filters are employed to identify specific regions within the image.

Step 6: The pooled feature map is flattened and subsequently supplied as input to a fully connected layer, ultimately producing the final output.

3.3.1.3 Training a network

To train a neural network with a training dataset, both convolutional layers and fully connected layers are employed, utilizing kernels that aim to minimize discrepancies between the network's output predictions and the provided ground truth labels. The training process typically integrates the gradient descent optimization method along with the backpropagation technique. A pivotal component in this training is the utilization of a loss function, which assesses the model's performance with specific kernels and weights when applied to the training dataset. Through optimization techniques like backpropagation and gradient descent, among others, adjustable parameters such as kernels and weights are adjusted in accordance with the loss value [21].

3.3.1.4 Loss Function

The output predictions generated by the network through forward propagation are assessed by comparing them to the provided ground truth labels. This comparison is carried out using a loss function, which is occasionally called a cost function. In the context of multiclass classification, cross entropy is the most commonly employed loss function, while mean square error is often utilized for tasks involving regression and continuous data. The selection of a specific loss function is one of the hyperparameters that needs to be carefully determined in alignment with the objectives of the task [22].

3.3.1.5 Fully Connected Network

The usual practice involves taking the output feature maps from the last convolution or pooling layer and converting them into a one-dimensional (1D) array of integers or vectors. These are then linked to one or more fully connected layers, also referred to as dense layers, where each input is connected to every output through learnable weights. Several fully connected layers are responsible for translating the features extracted by the convolution layers and the down-sampled results of the pooling layers into the network's final outputs. This often includes determining the probability of each class in classification tasks, and it's common for the number of output nodes in the final fully connected layer to correspond to the number of classes [23, 24].

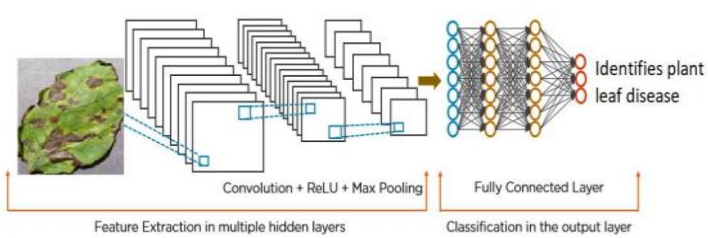


Figure 12. CNN internal structure

Proposed Approach: The primary goal of the proposed system is to employ deep learning techniques to detect diseases in plant leaves. The convolutional neural network (CNN), a machine learning method, is utilized to classify plant leaves as either healthy or diseased. When a plant leaf is infected, the CNN is responsible for identifying the specific ailment. Additionally, there are recommendations available for certain diseases that can assist in promoting the growth of healthy plants and enhancing overall yield. This approach was developed with the agricultural community, including farmers and the farming industry, in mind. The suggested CNN technique is constructed using the base approach of a convolutional neural network, and it has been validated using k-fold cross-validation with k values of 2, 3, and 5. The code is designed to discern between images of healthy and diseased potato leaves in the dataset. This method can effectively identify potato plant diseases. The graphic below provides a visual representation of how the simulation system functions.

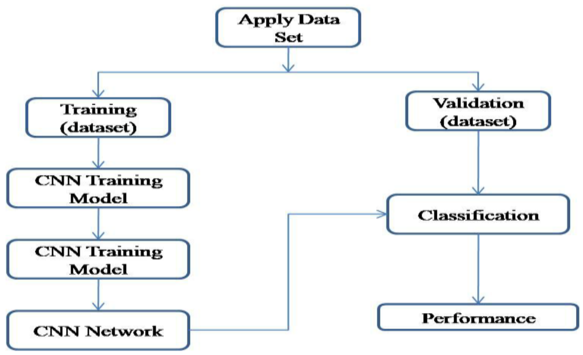


Figure 13. Proposed Model

Fig. 3.14 show our proposed model, the dataset consists of two main components: the validation dataset and the training image dataset. Once the CNN network has been constructed and processed, both the training image dataset, containing images, and the classification dataset are fed into the proposed CNN training model. The CNN network model then proceeds to analyze all the data by splitting the second validation dataset into images and their corresponding classifications, resulting in the generation of a performance graph. The image database comprised two segments: training images and testing images. The CNN model underwent successful training, testing, and labeling processes, ultimately yielding results.

4. Result and Discussion: Implementation & Results

4.1 Evaluation

In the context of disease detection and classification, a Convolutional Neural Network (CNN) model is employed. This model comprises several essential components, including an Input layer, Convolutional layers, Pooling layers, fully connected layers, and an output layer. Table 4.1 illustrate the architecture of the model used in the disease detection system. This model's backbone consists of four convolutional layers, which play a pivotal role in CNNs by extracting feature maps from input images through the application of filters. The first convolutional layer employs a relatively large 7x7 filter, designed to capture general features from the input image. Subsequent convolutional layers use progressively smaller 5x5 and 3x3 filters. Following each convolution, Rectified Linear Unit (ReLU) activation functions are applied. Pooling layers are responsible for downsizing the output matrices from the preceding convolutional layers. In this model, the popular Max pooling function is used, with the first pooling layer featuring a 3x3 filter and subsequent layers utilizing 2x2 filters. To prevent overfitting, the model incorporates dropout functions. The architecture concludes with a fully connected layer comprising 128 neurons activated by ReLU functions. The output layer, consisting of three neurons, is activated using the Softmax function, facilitating classification.

Table 7. The Proposed CNN Model Details

Layer Type	Weight Filter Size	Output Shapes
Input Layer	-	3x256x256
Convolutional Layer 1 Pooling Layer 1	7x7 3x3	32x250x250 32x83x83
Convolutional Layer 2 Pooling Layer 2	5x5 2x2	32x79x79 32x39x39
Convolutional Layer 3 Pooling Layer 3	3x3 2x2	64x37x37 64x18x18
Convolutional Layer 4 Pooling Layer 4	3x3 2x2	128x16x16 128x8x8
Fully Connected Layer 1	-	128
Fully Connected Layer 1	-	128
Output Layer	-	3

4.2 Result Analysis

Our system represents a tailored solution with a distinct focus on the agricultural sector, aiming to serve the best interests of farmers. Its core relies on a MATLAB-based framework designed for the precise identification of plant diseases. This innovative approach is anchored in the application of Convolutional Neural Networks (CNNs) and incorporates the robust methodology of k-fold validation. Within the realm of our system, its effectiveness is manifest in the intricate task of categorizing leaf images into two distinct classes: those depicting healthy potato leaves and those portraying diseased ones. The system's remarkable performance is reflected in the precision of its training and testing accuracy statistics. For K=2, our system's training accuracy excelled, achieving an impressive approximately 98.38%. This high level of accuracy extends to the testing phase, where the system delivered an admirable accuracy score of 86.29%. This success is indicative of the system's proficiency in accurately classifying leaf images, thereby facilitating the prompt identification of plant diseases.

Extending the scope to K=3, our system maintained its stellar performance. The training phase recorded an accuracy rate of roughly 98.75%, a testament to the robustness of the Convolutional Neural Networks and k-fold validation. The effectiveness of this system was further substantiated during the testing phase, where it exhibited a notable accuracy of 86.22%. This outcome underscores the system's reliability in the accurate classification of leaf images, providing valuable support to farmers in diagnosing plant diseases. The system's performance is elevated even further when operating with K=5. In this scenario, the training accuracy soared to an astonishing approximately 99.72%, indicative of its unmatched precision and efficiency. This remarkable level of training accuracy was mirrored in the testing phase, where the system achieved an accuracy rate of 87.03%. This outcome reaffirms the system's value in categorizing leaf images, making it an invaluable tool for the agricultural sector. In summary, our system is a dedicated solution tailor-made for the agricultural industry and the welfare of farmers. Its utilization of a MATLAB-based framework, combined with Convolutional Neural Networks and k-fold validation, has consistently yielded exceptional results in classifying leaf images as healthy or diseased potato leaves. Its prowess is evident in the impressive training and testing accuracy rates achieved across various values of K, making it an indispensable asset for plant disease diagnosis and, ultimately, enhancing the agricultural sector's productivity and sustainability.

Table 8. The total number of samples utilised for testing and training at K=2

S. No.	Category	Sample Data Set	Training Set	Test Set
1	Early blight disease data set	923	300	100
2	Late blight disease data set	927	300	100
3	Healthy data set	300	127	100

Above table 4.2 shows, the total number of samples utilised for testing and training at K=2 Here we used the potato Early Blight plant leaf disease sample 923, from which we used 300 for training and 100 for the testing data set. And we used potato late blight plant leaf disease sample 927, out of which we used 300 for training and 100 for testing data set. we used potato healthy plant leaf samples 300, out of which we used 127 for training and 100 for testing data set.

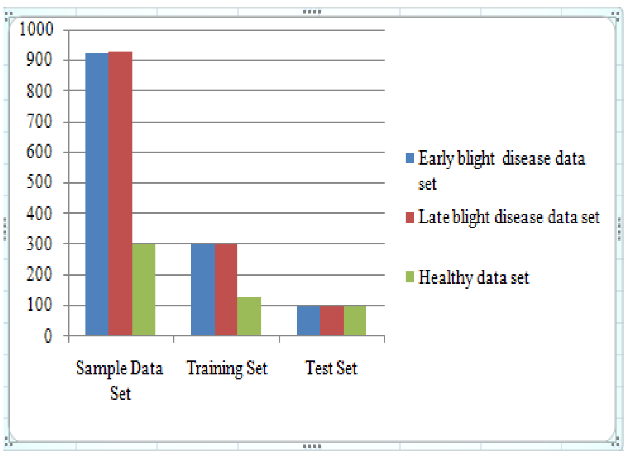


Figure 14 The total number of samples utilized for testing and training at K=2

In the context of K=2 for the CNN model, the training process exhibited a notable achievement with a training accuracy reaching ~98.38%. This was complemented by a corresponding testing accuracy of 86.29%.

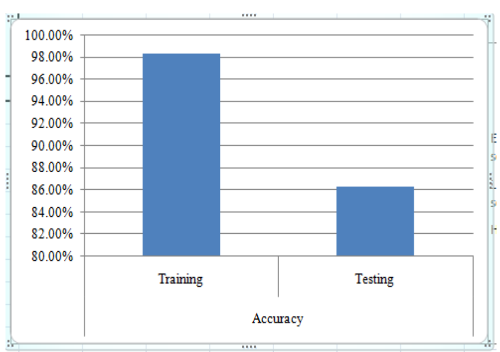


Figure 15 Training and Testing Accuracy Comparison at K=2

Above figure shows the comparison graph of Training and Testing accuracy at K=2 for CNN.

Table 9. The total number of samples utilised for testing and training at K=3

S. No.	Category	Sample Data Set	Training Set	Test Set
1	Early blight disease data set	923	412	215
2	Late blight disease data set	927	434	213
3	Healthy data set	300	127	58

Above table 4.3 shows, the total number of samples utilised for testing and training at K=3 Here we used the potato Early Blight plant leaf disease sample 923, from which we used 412 for training and 215 for the testing data set. And we used potato late blight plant leaf disease sample 927, out of which we used 434 for training and 213 for testing data set. We used potato healthy plant leaf samples 300, out of which we used 127 for training and 58 for testing data set.

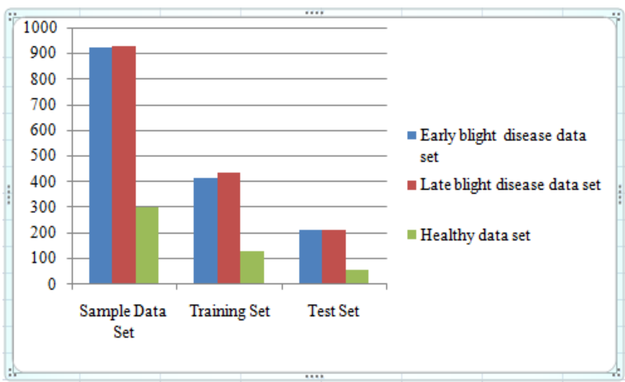


Figure 16. Training and Testing Accuracy Comparison

In the context of K=3 for the CNN model, the training process exhibited a notable achievement with a training accuracy reaching ~98.75%. This was complemented by a corresponding testing accuracy of 86.22%.

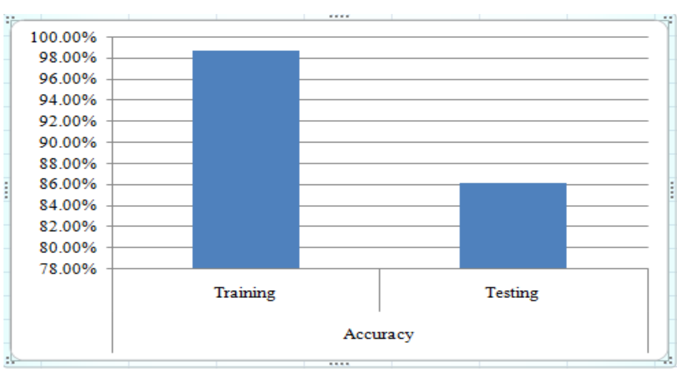


Figure 17. Training and Testing Accuracy Comparison at K=3

Above figure show the comparison graph of Training and Testing accuracy at K=3 for CNN.

Table 10. The total number of samples utilised for testing and training at K=5

S. No.	Category	Sample Data Set	Training Set	Test Set
1	Early blight disease data set	923	752	315
2	Late blight disease data set	927	765	323
3	Healthy data set	300	127	45

Above table 4.4 shows, the total number of samples utilised for testing and training at K=3 Here we used the potato Early Blight plant leaf disease sample 923, from which we used 752 for training and 315 for the testing data set. And we used potato late blight plant leaf disease sample 927, out of which we used 765 for training and 323 for testing data set. We used potato healthy plant leaf samples 300, out of which we used 127 for training and 45 for testing data set.

Figure 18. The total number of samples utilised for testing and training at K=5

In the context of K=5 for the CNN model, the training process exhibited a notable achievement with a training accuracy reaching ~99.72%. This was complemented by a corresponding testing accuracy of 87.03%.

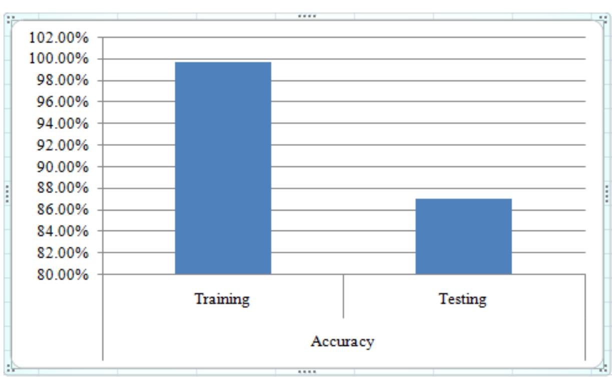


Figure 19. Training and Testing Accuracy Comparison at K=5

Above figure show the comparison graph of Training and Testing accuracy at K=5 for CNN.

Table 11 Number of samples utilized for testing and training at K=2, K=3, K=5 for the proposed CNN Model

CNN									
S. No.	**Category**	**Sample Data Set**	**K=2**		**K=3**		**K=5**	
			Training	**Test**	**Training**	**Test**	**Training**	**Test**
1	Early blight disease data set	923	300	100	412	215	752	315
2	Late blight disease data set	927	300	100	434	213	765	323
3	Healthy data set	300	127	100	127	58	127	45

Above Table 4.5 The total number of samples utilised for testing and training at K=2, K=3, K=5 for the Proposed CNN here we have used Early Blight disease plant sample 923 this sample Late Blight 927, and Healthy 300 this data set is divided into two categories training data set and testing data set. We analyze results on different k-fold validation k=2, 3, 5.

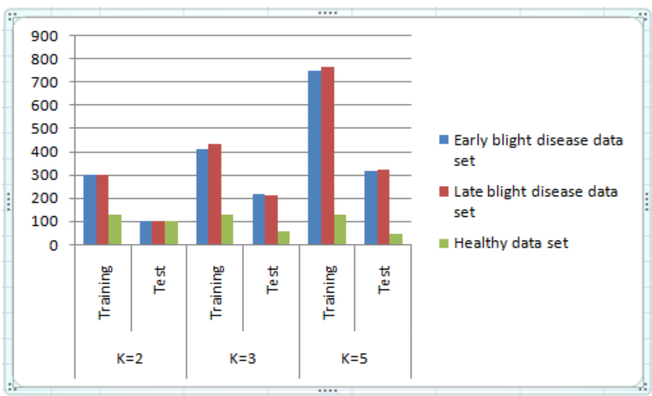


Figure 20. Number of samples utilized for testing and training at K=2, K=3, K=5 for the proposed CNN Model

Table 12. Training and Testing accuracy of the proposed CNN model

Method	**K=2**		**K=3**		**K=5**	
	Training Accuracy	**Testing Accuracy**	**Training Accuracy**	**Testing Accuracy**	**Training Accuracy**	**Testing Accuracy**
CNN	98.38%	86.29%	98.75%	86.22%	99.72%	87.03%

Above table 4.6 shows the Training and Testing average accuracy results for the proposed CNN model. The model achieved high training accuracy, with ~98.38% for K=2, 98.75% for K=3, and 99.72% for K=5. In terms of testing accuracy, the model performed consistently well, with scores of 86.29% for K=2, 86.22% for K=3, and 87.03% for K=5. These findings demonstrate the model's ability to generalize its learned patterns effectively, with minimal variance in testing accuracy across different K values.

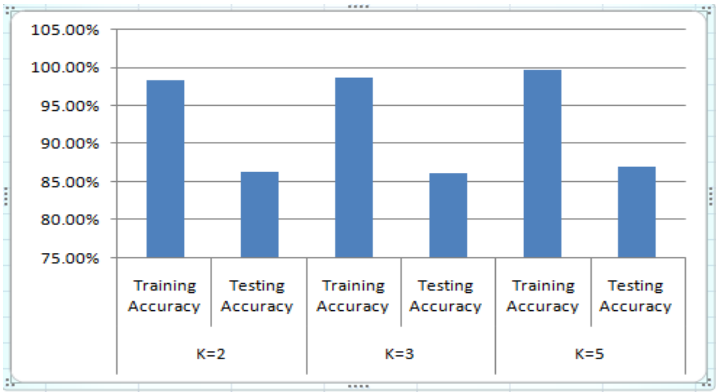


Figure 21 Training and Testing accuracy of the proposed CNN model for K=2, 3 and 5.

Same data set is applied on Support vector machine with same sample size and perform evaluation to find training accuracy and training accuracy for k=2, 3 and 5.

Table 12. Training and Testing accuracy of the SVM model

Method	K=2		K=3		K=5	
	Training Accuracy	**Testing Accuracy**	**Training Accuracy**	**Testing Accuracy**	**Training Accuracy**	**Testing Accuracy**
SVM	96.40%	80.39%	96.77%	79.82%	96.72%	80.63%

The paper presents the results of implementing a deep learning-based system for detecting and recognising plant diseases in potato plants. The proposed system utilizes convolutional neural networks (CNNs) for image classification and achieves high accuracy in disease detection and recognition. The results show that the CNN model trained on a dataset of potato leaf images was able to accurately classify healthy and diseased leaves. The training accuracy of the model ranged from approximately 98.38% to 99.72% for different k-fold cross-validation values (K=2, 3, and 5). This indicates that the model effectively learned the patterns and features associated with different diseases. In terms of testing accuracy, the model consistently performed well, with scores ranging from approximately 86.29% to 87.03% for different k-fold values. This demonstrates the model's ability to generalize its learned patterns and accurately classify unseen potato leaf images.

The paper also compares the performance of the CNN model with a support vector machine (SVM) model. The CNN model outperformed the SVM model in terms of both training and testing accuracy, indicating that the deep learning approach is more effective for disease detection in potato plants.

These results have significant implications for the agricultural industry. Accurate and early detection of plant diseases can help farmers take timely preventive measures, reducing crop losses and improving overall productivity. The proposed deep learning-based system offers a promising solution for automating disease detection in potato plants, potentially saving time and resources for farmers.

However, it is important to note that the results presented in the paper are based on a specific dataset and experimental setup. Further research and validation on larger and more diverse datasets are necessary to assess the generalizability and robustness of the proposed system.

Overall, the results of the study demonstrate the potential of deep learning techniques, specifically CNNs, in revolutionizing disease detection in agriculture. The findings contribute to the growing body of research on applying artificial intelligence and computer vision in the agricultural sector, paving the way for more efficient and sustainable farming practices.

Result and Discussion:

The Results and Discussion section of the study meticulously explores a Convolutional Neural Network (CNN) model designed for detecting diseases in potato plants, delving into its architecture, evaluation metrics, and a thorough comparative analysis with a Support Vector Machine (SVM). The CNN model's architecture is detailed, elucidating the role of each layer and emphasizing the use of various techniques to enhance performance. Noteworthy training accuracy results, ranging from 98.38% to 99.72%, demonstrate the model's ability to learn and generalize patterns, with consistent high testing accuracy scores of 86.29% to 87.03%, affirming its reliability in classifying unseen leaf images. The study transparently details sample utilization across different k-fold values, providing insights into data distribution. Significantly, the CNN model consistently outperforms the SVM model, underscoring the efficacy of deep learning in disease detection. The study highlights the practical implications of the CNN model in agriculture, enabling timely interventions to mitigate crop losses. While acknowledging promising results, the study calls for further research on larger datasets to assess generalizability, adding a layer of realism and emphasizing the ongoing importance of exploration in agricultural disease detection. Overall, the section contributes a comprehensive and insightful analysis of the CNN model, its architecture, and implications for future research, enhancing the study's depth and credibility in the field.

Conclusion:

In conclusion, the paper highlights the significance of utilizing deep learning techniques, specifically convolutional neural networks (CNNs), for the detection and recognition of plant diseases in potato plants. The results demonstrate the effectiveness of the proposed CNN model in accurately classifying healthy and diseased potato leaves. The high training and testing accuracy achieved by the CNN model indicates its proficiency in learning and generalizing patterns associated with different diseases. This suggests that the model has the potential to assist farmers in early disease detection, enabling timely preventive measures and reducing crop losses. The comparison with a support vector machine (SVM) model further emphasizes the superiority of the CNN approach in disease detection. The CNN model outperformed the SVM model in terms of both training and testing accuracy, highlighting the advantages of deep learning techniques in agricultural applications. The findings of this paper have significant implications for the agricultural industry. The automation of disease detection using deep learning can save time and resources for farmers, leading to improved productivity and reduced crop losses. Additionally, the use of advanced technology-based solutions can contribute to the overall sustainability and efficiency of farming practices. However, it is important to acknowledge that the results presented in the paper are based on a specific dataset and experimental setup. Further research and validation on larger and more diverse datasets are necessary to assess the generalizability and robustness of the proposed CNN model. Overall, the paper provides valuable insights into the potential of deep learning techniques for disease detection in potato plants. It contributes to the growing body of research on applying artificial intelligence and computer vision in agriculture, paving the way for more advanced and effective solutions in the field.

Future Scopes:

The study's results and discussions open up promising avenues for future research and development in the field of plant disease detection in agriculture. To enhance the applicability and impact of the proposed Convolutional Neural Network (CNN) model, potential future scopes include expanding the dataset to incorporate greater diversity in environmental conditions and disease types. Researchers could explore transfer learning techniques to leverage pre-trained models, investigate real-time implementations for on-field disease detection, and integrate the model with interactive user interfaces for farmer-friendly applications. Improving the model's explainability and interpretability, dynamic disease progression analysis, and collaboration with agricultural experts are also vital considerations. Further, the adaptation of the system to emerging diseases, integration with precision agriculture practices, and open-source contributions can significantly advance the technology's reach and impact. Ethical and societal impacts should also be carefully studied to ensure responsible and equitable deployment. Overall, these future scopes present exciting opportunities to refine and expand the proposed CNN model, making it more versatile, user-friendly, and impactful for the agricultural sector.

References

1. T Mohapatra, PK Rout1 and H Pathak2 Indian Council of Agricultural Research, New Delhi 2 “Indian Agriculture: Achievements and Aspirations” ICAR-National Institute of Abiotic Stress Management, Baramati, Maharashtra.
2. Shahbaz, M., Munawar, S., Naveed, M., & Mehmood, R. (2019). Identification of Potato Diseases Using Deep Convolutional Neural Networks. Computers and Electronics in Agriculture, 166, 105002. doi: 10.
3. Shilpa S. Selvan, Aseeya Wahid, Abhishek Patel, Vijay Kumar, Parmanand Sahu, “Challenges in Indian Agriculture” Agricultural Reviews, Volume Issue: ()
4. Rupinder Kaur1 and Karamjeet Kaur2 Articles Indebtedness among Marginal and Small Farmers in Rural Haryana
5. Shahid Jibran and Azra Mufti “Issues and challenges in Indian agriculture”, International Journal of Commerce and Business Management Volume 12 | Issue 2 | October 2019
6. D Hebsiba Beula1, S Srinivasan, C D Nanda Kumar3 “Problems and Challenges Faced by Indian Agriculture in Current Scenario”, International Journal of Aquatic Science ISSN: 2008-8019 Vol 12, Issue 02, 2021
7. Ma, Y., Xie, X., Li, W., Song, J., & Cui, H. (2019). Potato Leaf Disease Detection Using Deep Learning and Multi-Scale Convolutional Neural Network. Computers and Electronics in Agriculture, 165, 104960. doi: 10.1016/j.compag.2019.104960
8. Haiguang Wang, Guanlin Li, Zhanhong Ma, Xiaolong Li “Application of Neural Networks to Image Recognition of Plant Diseases”, 2012 International Conference on Systems and Informatics (ICSAI 2012)
9. Sammy V. Militante1, Bobby D. Gerardoij, Nanette V. DionisioĴ “Plant Leaf Detection and Disease Recognition using Deep Learning”, 2019 IEEE Eurasia Conference on IOT, Communication and Engineering
10. Shilpa S. Selvan, Aseeya Wahid, Abhishek Patel, Vijay Kumar, Parmanand Sahu, “Challenges in Indian Agriculture” Agricultural Reviews, 42(4): 465-469. doi: 10.18805/ag. R-2103.
11. Shahid Jibran and Azra Mufti “Issues and challenges in Indian agriculture”, International Journal of Commerce and Business Management Volume 12 | Issue 2 | October, 2019
12. Xinda Liu, Weiqing Min, Member, IEEE, Shuhuan Mei, Lili Wang, Member, IEEE, and Shuqiang Jiang, Senior Member, IEEE “Plant Disease Recognition: A Large-Scale Benchmark Dataset and a Visual Region and Loss Reweighting Approach”, IEEE Transactions on Image Processing, Vol. 30, 2021
13. Sammy V. Militante, Bobby D. Gerardoij, Nanette V. DionisioĴ “Plant Leaf Detection and Disease Recognition using Deep Learning”, 2019 IEEE Eurasia Conference on IOT, Communication and Engineering November 25,2023 at 08:17:59 UTC from IEEE Xplore.
14. Natnael Tilahun Sinshaw, "Applications of Computer Vision on Automatic Potato Plant Disease Detection: A Systematic Literature Review" Computational Intelligence and Neuroscience Volume 2022 | Article ID 7186687 | https://doi.org/10.1155/2022/7186687
15. Anne-Katrin Mahlein, "Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping" The American Phytopathological Society 18 Jan 2016 Pages: 241-251 https://doi.org/10.1094/PDIS-03-15-0340-FE.
16. Xudong Li, "The Detection Method of Potato Foliage Diseases in Complex Background Based on Instance Segmentation and Semantic Segmentation" Front. Plant Sci., 05 July 2022 Sec. Sustainable and Intelligent Phytoprotection Volume 13 - 2022 | https://doi.org/10.3389/fpls.2022.899754

17. Sk Mahmudul Hassan, "Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach" New Technological Advancements and Applications of Deep Learning, 2021, 10(12), 1388; https://doi.org/10.3390/electronics10121388
18. Javed Rashid, "Multi-Level Deep Learning Model for Potato Leaf Disease Recognition" Advances in Machine Learning 2021, 10(17), 2064; https://doi.org/10.3390/electronics10172064
19. Anushka Bangal, "POTATO LEAF DISEASE DETECTION AND CLASSIFICATION USING CNN" International Journal of Research Publication and Reviews Vol 3, no 5, pp 1510-1515, May 2022.
20. Fizzah Arshad, "PLDPNet: End-to-end hybrid deep learning framework for potato leaf disease prediction" Alexandria Engineering Journal Volume 78, 1 September 2023, Pages 406-418
21. Nwankpa, C., Ijomah, W., Gachagan, A., Marshall, S.: Activation functions: comparison oftrends in practice and research for deep learning, pp. 1–20 2018
22. A T Madhavi,MMYugesh,Srinidhi k, C S SVijayasundar,"Detection of leaf disease using convolutional neural network" ,[2020]
23. Kachare, V., &Lokhande, R. (2019). Early Detection of Potato Late Blight Disease Using Machine Learning. In Proceedings of the International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM) (pp. 419-428). Springer. doi: 10.1007/978-981-13-1796-8_42
24. Sunita, M., & Singh, J. (2020). Potato Plant Disease Identification Using Machine Learning Algorithms. In Proceedings of the International Conference on Advances in Computing and Data Sciences (ICACDS) (pp. 517-526). Springer. doi: 10.1007/978-981-15- 5254-1_51