International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Potato Leaf Disease Detection Using Convolution Neural Network Model

S. K. Bharadwaj[1], V. K. Jadon[2], Amit Sharma[3], Jitendra Kumar[4], D. K. Mishra[5], Ashish Shukla[6]

Submitted: 05/02/2024 **Revised**: 13/03/2024 **Accepted**: 19/03/2024

Abstract: Potatoes, being one of the most widely consumed vegetables globally, have increasingly become a focus for agricultural departments worldwide. However, despite their popularity, potato leaf diseases pose a significant threat to potato crops. A range of diseases, including early blight, late blight, and Septoria blight, can affect potato plants, manifesting symptoms in their leaves. Detecting and addressing these outbreaks early on is crucial to prevent major economic losses for farmers.

In this research paper, we propose a model that utilizes image processing techniques to identify and detect diseases in potato leaves. Our approach relies on a Convolution Neural Network (CNN), chosen for its effectiveness in image classification tasks. By leveraging CNN technology, we aim to provide accurate and efficient detection of potato leaf diseases, thereby enabling timely intervention and mitigation measures

Keywords- Leaf Disease, CNN, Potato, Accuracy, Agriculture, Deep Learning.

1. Introduction

Traditional manual detection methods are laborious, time-consuming, and prone to errors. In response, this article proposes the development of an end-to-end project utilizing Deep Learning techniques. The objective is to detect and identify potato leaf diseases, which encompass a range of illnesses that are challenging for the human eye to classify accurately. Leveraging Convolutional Neural Networks (CNNs), our project aims to provide an efficient and accurate solution for disease detection from images of potato leaves.

Fig. 1: Healthy Leaves

Fig. 2 : Late Blight Leaves

[1,4,5,6]Department of Engineering Mathematics and Computing Madhav Institute of Technology & Science, India
[2,3]Department of Applied Sciences, Anand Engineering College, Agra
[1]bharadwajsantosh@gmail.com,[2]dr.vinayjadon@gmail.com,[3]gopal_adya@yahoo.com, [4]mishradilip3826@gmail.com, [5]jkmuthele@mitsgwalior.in,[6]shuklasachin850@gmail.com
Orcid Id:[1]https;//orcid.org/0000-0003-3355-4870, [2]https;//orcid.org/0009-0000-6528-2908, [3]https;//orcid.org/0009-0001-4833-0506, [4]https;//orcid.org/0000-0003-4444-1257,[5] https://orcid.org/0000-0001-5939-5586, [6]https://orcid.org/0009-0009-6697-3140

Fig. 3: Early Blight Leaves

It seems like you're providing a summary of various research studies that focus on using different deep learning models for plant disease detection, particularly on potato, maize, corn, pepper bell, and wheat plants. Here's a rephrased version.

Ghosh et al. [1] conducted an experimental study utilizing RGB images of potato plant leaves captured under controlled laboratory conditions. Their research involved training Convolutional Neural Network (CNN) and Support Vector Machine (SVM) models. They observed variations in model performance attributed to different environmental conditions during image capture, leading them to create a consistent dataset and achieve reliable categorization accuracy's.

Tiwari et al. [2] utilized the Kaggle dataset of potato plant leaf images to classify them into healthy, early blight, and late blight categories. Employing the pre-trained VGG19 model, they achieved an impressive accuracy of 97.8%

Sumit Kumar et al. [3] introduced a deep learning approach based on CNN and region-based fully connected RCNN networks for detecting various complex plant leaf diseases. Their model demonstrated a validation accuracy of 94.6%

Shrivastav et al. [4] developed a model using Deep Convolutional Neural Networks (DCNN) with the Kaggle Plant Village dataset, simplifying plant health identification. Their model achieved an accuracy of 88%.

Shrestha et al. [5] designed a CNN-based model aimed at early-stage plant disease detection, working across various plant species and implementing the model in Python programming. Their model achieved an accuracy of 88.8%

Rangaraju et al. [6] focused on maize plant leaf disease detection using a computer vision system. They captured images with a 13 MP camera and achieved an overall testing accuracy of 86.70%. S K Hassan [7] implemented DCNNs to detect plant leaf diseases early on using the Kaggle Plant Village dataset, covering 14 different crops and 38 classifiers, including the healthy class. They achieved high accuracies ranging from 97.02% to 99.56% using various models Pranay et al. [8] worked with an open-source dataset of corn leaf images, training a CNN model with 4000 images and achieving 98% accuracy on 2000 test images Mangal et al. [9] utilized pepper bell plant images from the Plant Village dataset to distinguish between healthy and diseased leaves. Employing the canny edge technique for segmentation, GLCM technique for feature extraction, and CNN classifier, they attained an accuracy of 97.82%. Gaikwad et al. [10] investigated wheat plant leaf images and internet-sourced images, employing the KNN cluster method for segmentation and utilizing texture, shape, and color features for classification. They achieved accuracies of 80.21% and 89.23% using Feed Forward Neural Network and SVM, respectively.

2. Methodology

Various methodology are used for potato leaf disease classification for example image acquisition, image pre-processing, image augmentation, feature extraction and classification. Feature extraction and classification are preformed using CNN. The steps of potato leaf disease classification are presentation in figure 1.

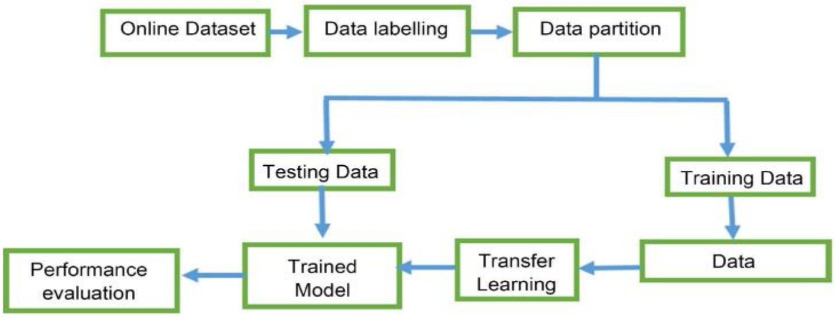

Fig. 4: Overview architecture of the proposed study.

2.1. **Image Acquisition**: The text highlights the significance of acquiring a suitable dataset for potato leaf disease recognition. It suggests three methods for dataset collection: utilizing ready-made data from third-party vendors or platforms like Kaggle, manually capturing images from various fields using a camera (acknowledging the time-consuming nature of this approach), and gathering images from websites featuring potato-related content

In this project, a ready-made dataset from Kaggle was employed due to its extensive collection of diverse image datasets available online. Kaggle serves as a valuable resource for researchers seeking high-quality datasets for various purposes

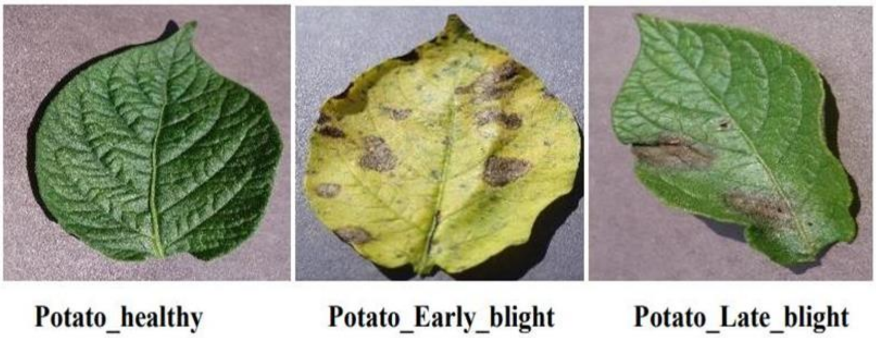

Fig. 5.1: Potato leaf image samples e.g Healthy, Early Blight, Late blight.

It is a popular source for many types of image collections. The dataset is split into a training dataset, a validation dataset, and a test dataset.

There are three types of potato leaf datasets in our dataset e.g. healthy, early blight, and late blight. Alternaria solani is the root cause of early blight. Both hills and lowlands contain it. Concentric rings define the angular, oval shape of this brown to black necrotic patch. On the leaf, several dots combine and disperse. Phytophthora Infestans is the cause of late blight in potatoes. It affects tubers, leaves, and stems. Leaf spots become larger, more numerous, and eventually turn purple-brown before becoming completely black. Under the leaves' surface, white growth appears. The sample image from our potato leaf disease dataset is presented in Figure 5.2.

Fig. 5.2: Potato leaf image samples e.g Healthy, Early Blight, Late blight.

2.2. **Image Processing: Various** image preprocessing techniques are employed before feature extraction of an image to improve performance, such as resizing the image, filtering the images, removing the noise, changing the colour, data augmentation, normalization, and image segmentation. After being captured, plant leaf photos are typically noisy. These noisy images are very tough to recognize. So we should remove the noise from the initially collected noisy image dataset for proper recognition of those images. Then we have to resize the image to get the same sized image. An error is likely to occur if the image is not the same size as the model is built-in code using a library like TensorFlow. To make it simpler for algorithms to learn from the data, scale pixel values to a common range, often between 0 and 1 is performed by using normalization. Image preprocessing can speed up model inference while reducing the need for model training.

2.3. **Image Augmentation:** Augmentation involves expanding the dataset by employing various techniques. In disease classification, the number of images or data is increased by using rotation, flipping, shifting, randomly changing the brightness, zooming, etc. However, a significant differentiation exists between image augmentation and image preprocessing. The text discusses the application of image preprocessing techniques on both training and test datasets, while highlighting that image augmentation is specifically applied to the training data. It emphasizes that capturing every possible real-world event in an image is impractical for a model. Therefore, by augmenting the training data, the sample size can be expanded, incorporating new situations that may be challenging to encounter in the real world. This approach enables the model to learn from a broader range of events and enhances its ability to generalize to different scenarios.Image augmentation is an important issue for addressing the over-fitting problem in deep learning. It is used to avoid over-fitting problems and also improves classification accuracy. Several data augmentation samples are shown in Figure 1-3

2.4. **Feature Extraction:** It is used for finding patterns from images that are useful for disease identification from the image. The convolution layers and pooling layers are combined to create the feature extraction process, which is then followed by fully linked layers and softmax classification layers. Based on the provided inputs, the softmax classifier identifies the outputs. It reduces the dimension and also eliminates redundant data. There is no loss of relevant and significant features of image data when it reduces the dimension and the number of resources needed for processing. A feature vector is formed based on a similar feature that is used for the recognition and categorization of an object. The features are directly extracted by CNN from the source image. The images are classified using another neural network, such as SVM, DT, RF, etc., but they do not extract the features from the images directly like CNN. The input image is used for extracting the feature

2.5. **Classification**: CNN has been the most common technique for classification because of its strong feature extraction capabilities used in the CNN process for potato leaf disease classification. The common classification process is shown in the second portion of the CNN working process represented in Figure6

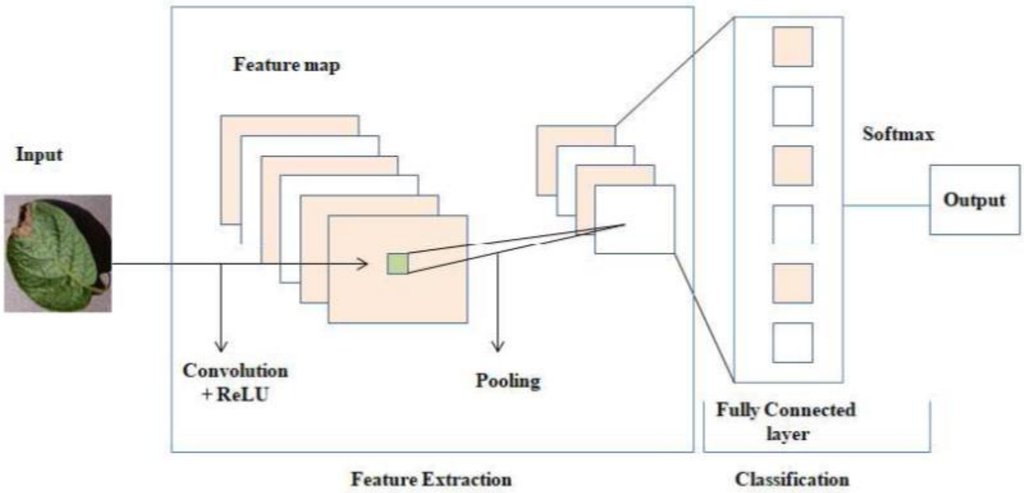

Fig. 6: CNN working process

2.6. **Evaluation and Recognition:** We used some performance metrics for measuring the performance of potato leaf disease classification problems or any classification problems in machine learning and deep learning. Confusion matrix is used for this purpose. It is a popular benchmark for accuracy or error calculations in classification problems. The performance metrics are accuracy, precision, and recall are calculated by using a confusion matrix [18,19]. The confusion matrix displays a summary of actual versus predicted values, encompassing True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) within a matrix structure.[20]. It is highly crucial for evaluating a model's performance as it reveals the model's accuracy and its rate of errors. Different evaluation metrics e.g. accuracy, precision, and recall are available for testing a model's performance [21].

3. **Data Pre-processing**: A diverse dataset comprising high-resolution images of potato leaves affected by various diseases, as well as images of healthy leaves, was collected from multiple sources.

4. Careful curation and validation ensured the inclusion of broad spectrum of potato leaf diseases, capturing different stages and severity levels.

5. The process begins with Image Preprocessing, wherein images are resized uniformly to a predetermined resolution (e.g., 256x256 pixels) to ensure consistency across the dataset. Following this, Data Augmentation is performed to increase the quantity of data. This involves generating multiple realistic variants of each training sample, artificially expanding the size of the training dataset. Data augmentation helps prevent overfitting by introducing diversity into the training data

6. Augmentation includes Image augmentation techniques, such as rotation, flipping, zooming, and shearing, were employed to expand the dataset.

7. **Model Building:** We generate CNN neural network using Max Pooling and Conv layers.

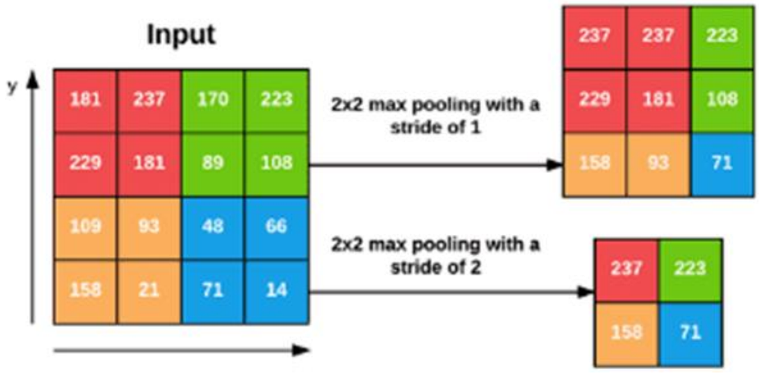

Fig. 7: CNN Pooling process

8. **Input Layer:** The initially reaccepts preprocess dim ages of potato leaves as input.

9. **Convolution Layers: Sequential** convolutional layers are stacked to extract hierarchical features from the input images. Each convolution layer is followed by ReLU activation to introduce non-linearity.

10. **Pooling Layers:** After convolution, max-pooling layers down sample feature maps, preserving essential information while reducing spatial dimensions.

11. **Output Layer:** The final layer employs a softmax activation function to produce probability distributions across disease classes, facilitating accurate classification.

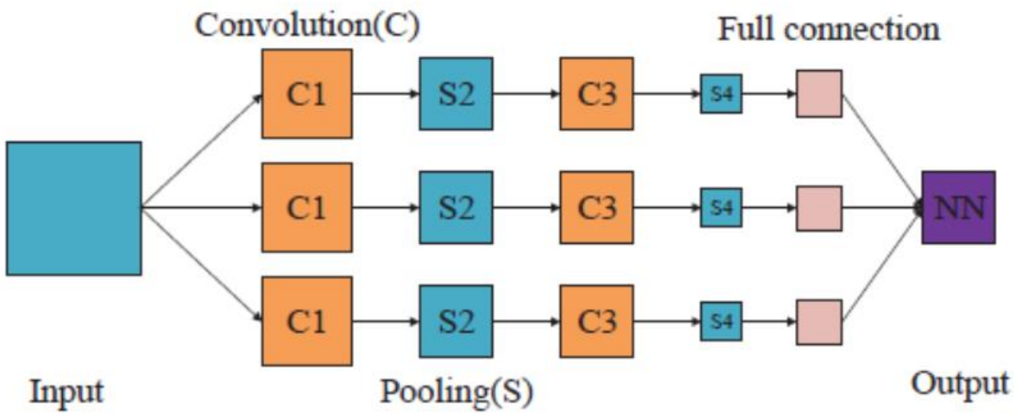

Fig8: The CNN model consisting of convolution layers, pooling and full connection layer [17].

12. **Algorithm**–CONVOLUTIONAL NEURAL NETWORK:

We utilize the Convolution Neural Network (CNN) algorithm, a sophisticated deep learning technique designed for processing images and videos. Unlike Artificial Neural Networks (ANNs), CNNs excel at learning intricate features efficiently to classify objects within input images or video.

Compared to ANNs, CNNs possess several advantageous features. They leverage sparse connections, parameter sharing, and equivariance representation, leading to a significant reduction in training parameters and improved generalization capabilities. Inspired by the human nervous system and the visual cortex organization, CNN architecture is tailored for image processing task

CNNs are trained on extensive datasets comprising thousands of images, enabling them to extract features effectively. This characteristic contributes to the exceptional accuracy of CNN models in computer vision tasks. Deep CNNs consist of tens or hundreds of hidden layers, enhancing their ability to extract intricate features from images due to their increased complexity

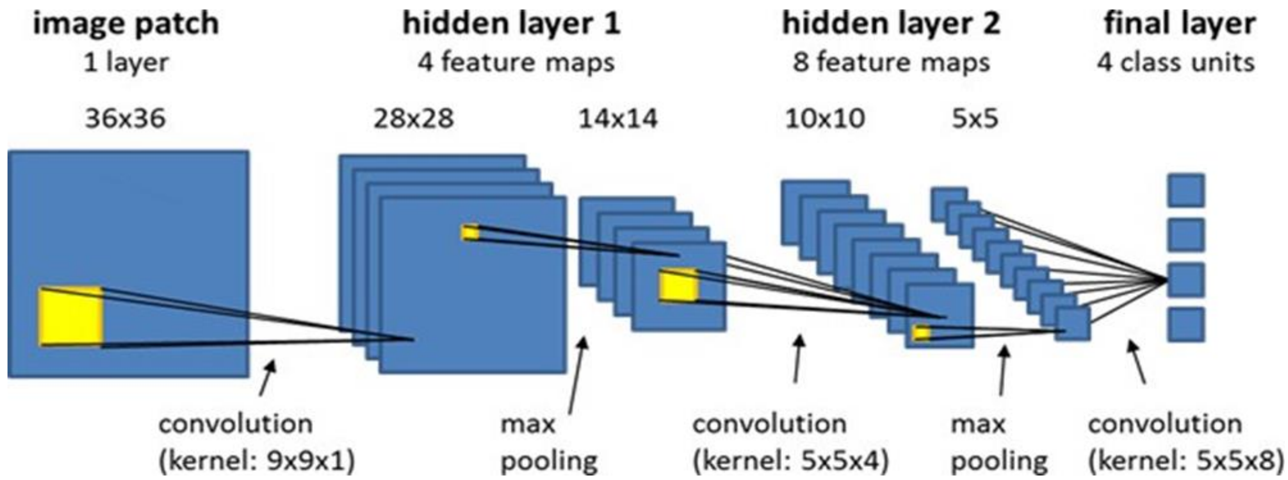

Fig9 :The CNN model consisting of different types of layers

13. Numerical Result & Graph-

When applying Convolution Neural Networks (CNNs) for potato leaf disease detection using deep learning, the process typically involves steps like data collection, preprocessing, model development, training, evaluation, and inference. We fit our model over training data and find the results and Based on it compute the accuracy scores. The train test validation is shown below.

This will generate two graphs showing the training and validation loss, as well as accuracy over epochs

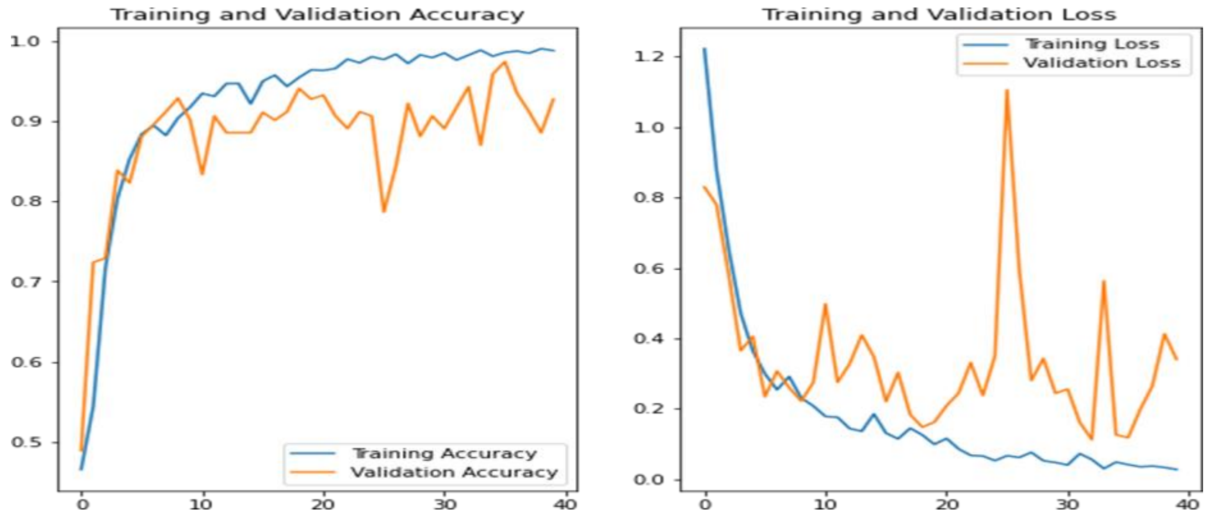

Fig 10: CNN Train test accuracy validation

Training and Validation Accuracy Graph: A line plot illustrating the model's accuracy over training epochs. This graph demonstrates the convergence of the model during training and potential indications of over fitting or under fitting.

Fig 11: Disease Potato leaf

Table 01: compression of different methods

Author	**year**	**Vegetable/ plant leaf**	**Methodology Used / Classifer**	**Accuracy (%)**
Mohanty et al.[11]	2016	Plant Village dataset for 26 different diseases with 38 classes	AlexNet	85.53%
Hu. Yh. Et al[16]	2016	Potato	Hyperspectral Imaging	95%
Islam M. etl[15]	2017	Potato leaf	Segment and Multi SVM	95%
Manya et al [13]	2019	RTK-GNSS Potato leaf	ResNet 18	93.78%
Ghosh etl.[14]	2021	Potato leaf	Pre-trained SNet	93.78%
Proposed Paper		Potato Leaf	CNN	98%

14. Conclusion:

In summary, this paper has delved into the utilization of Convolutional Neural Networks (CNN) and its significant impact on precision agriculture, particularly in the identification of diseases affecting potato leaves. The successful integration of CNN-based disease identification systems has far-reaching implications, offering valuable insights for enhancing agricultural practices.

By focusing on early detection and accurate classification of diseases such as early blight, late blight, and healthy conditions in potato leaves, our developed model has exhibited an impressive classification accuracy of **approximately 98%.** This high level of precision is indicative of the potential of CNN models to serve as a transformative tool in the realm of potato leaf disease identification.

The implications of our research extend beyond the real ms of academia, pointing towards tangible applications in real-world scenarios. The ability to identify and categorize diseases with such accuracy opens the door to timely interventions, which, in turn, can play a crucial role in minimizing yield losses and optimizing crop management strategies.

As we move forward, the findings of this study underscore the pivotal role that CNN models can play in revolutionizing agricultural practices. The integration of advanced technologies in disease identification not only aids in the improvement of crop yield but also contributes to the broader initiative of advancing precision agriculture. In conclusion, our research positions CNN models as a corner stone technology, emphasizing their potential to drive positive changes in the field and shape the future of agriculture.

References:

[1] Soma Ghosh, Renu Rameshan and and Dileep A. D. “An Empirical Study on Machine Learning Models for Potato Leaf Disease Classification using RGB Images”, ICPRAM 2021 - 10th International Conference on Pattern Recognition Applications and Methods.

[2] Divyansh Tiwari, Mritunjay Ashish etl.” Potato Leaf Diseases Detection Using Deep Learning”, Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2

[3] Sumit Kumar et., Plant Disease Detection Using CNN, Turkish Journal of Computer and Mathematics Education Vol.12 No.12 (2021), 2106-2112

[4] PrakanshuShrivastav , Vaibhav Avsthietl, “Plant Disease Detection using Convolutional Neural Network”, Internation Journal of Advanced Research(IJAR) January 2021

[5] Garima Shrestha, Deepsikha etl. “Plant Disease Detection using CNN”, Proceedings of 2020 IEEE Applied Signal Processing Conference (ASPCON), ISBN: 978-1-7281-6882-1

[6] Gokaraju Rangaraju, VNR Vignana Jyothi, Plant Disease Detection Using Convolutional Neural Networks , International Journal of Advanced Trends in Computer Science and Engineering, ISSN 2278-3091, June 2020

[7] S k Mahmudul Hassan , Arnab Kumar Maji etl. “Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach”, Electronics 2021, 10, 1388. https://doi.org/10.3390/electronics10121388, June 2021

[8] Garlapati Pranay, Aditya Kurubaetl. Detection of Disease and Damage Control for Crops using Convolution Neural Networks, Journal of Xi'an University of Architecture & Technology, ISSN No : 1006-7930, Volume XII, Issue XI, 2020

[9] Sakshi Mangal and Pratiksha Meshram, PLANT DISEASE IDENTIFICATION USING DEEP LEARNING CLASSIFICATION MODEL: CNN, Journal of University of Shanghai for Science and Technology, ISSN: 1007-6735, Volume 23, Issue 1, January – 2021

[10] Varsha P. Gaikwad and Dr. Vijaya Musande,Wheat Disease Detection using Image Processing, 2017 1st International Conference on Intelligent Systems and Information Management(ICSIM),978-1-5090-4264-7/17/$31.00©2017IEEE, 10.1109/ICISIM.2017.8122158

[11] Shrad P. Mohanty, David P. Hughes and Marcel Salathe, Sept 2016 ”Using Deep Learning for Image-Based Plant Disease Detection” Front. Plant Sci 7:1419 doi: 10.3389/fpls.2016.01419, Volume 7.

[12] Varsha P. Gaikwad, Vijaya Musande, 2020, “Plant leaf damage detection based in colour and texture using deep learning – ALEXNET”, Solid State Technology Volume: 63 Issue: 2 Publication

[13] Manya Afonso, Pieter etl. Blackleg Detection in Potato Plants using Convolutional Neural Networks, https://doi.org/10.1016/j.ifacol.2019.12.481

[14] Soma Ghosh, Renu Rameshan and Dileep A. D ,An Empirical Study on Machine Learning Models for Potato LeafDisease Classification using RGB Images, , ICPRAM 2021 - 10th International Conference on Pattern Recognition Applications and Methods.

[15] Islam, M.; Dinh, A.; Wahid, K.; Bhowmik, P. Detection of potato diseases using image segmentation and multiclass support vector machine. In Proceedings of the 2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE), Windsor, ON, Canada, 30 April–3 May 2017; pp. 1–4.

[16] Hu YH, Ping XW, Xu MZ, Shan WX, He Y, "Detection of Late Blight Disease on Potato Leaves Using Hyperspectral Imaging Technique," PubMed, 36(2), 2016, pp. 515-519.

[17] Xiaofei, ,Xutao li, Y.K. Lau, Xiaofeng Zhang, "Hyperspectral Image Classification With Deep Learning Models" IEEE(2018), 1-16.

[18] Asif, M.K.R.; Rahman, M.A.; Hena, M.H. CNN based disease detection approach on potato leaves. 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2020, pp. 428–432

[19] Rashid, J.; Khan, I.; Ali, G.; Almotiri, S.H.; AlGhamdi, M.A.; Masood, K. Multi-level deep learning model for potato leaf disease recognition. Electronics 2021, 10, 2064.

[20] Singh, S.; Agrawal, V. Corrigendum: Development of ecosystem for effective supply chains in 3D printing industry–an ISM approach (2021 IOP Conf. Ser.: Mater Sci Eng. 1136 012049). IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2021, Vol. 1136, p. 012078.

[21] Sholihati, R.A.; Sulistijono, I.A.; Risnumawan, A.; Kusumawati, E. Potato leaf disease classification using deep learning approach. 2020 international electronics symposium (IES). IEEE, 2020, pp. 392–397.