Potato Leaf Disease Detection and Classification using CNN

Dr. Suman Kumar Swarnkar
Assistant Professor
Department of Computer Science & Engineering,
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India
Sumanswarnkar17@gmail.com

Mr. Abhisek Chakraborty
Department of Computer Science & Engineering,
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India
abhishek.chakraborty@ssipmt.com

Mr. Arijeet Dutta
Department of Computer Science & Engineering,
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India
dutta.arijeet@ssipmt.com

Ms. Nachiketa Mohanty
Department of Computer Science & Engineering,
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India
mohanti.nachiketa@ssipmt.com

***Abstract*— All of us are familiar with potatoes as a vegetable. In India, growing potatoes has become quite popular in recent years. However, diseases like early blight and late blight are impeding potato production and driving up production costs. To boost potato production and digitize the system, the goal is to create an automated and quick disease detection method. Our primary objective is to use leaf images to diagnose potato disease using the CNN algorithm. This paper provides a visual representation of how automated systems based on machine learning will identify and categorize potato leaf diseases. The most effective method for identifying and analyzing these disorders is image processing. In this analysis, picture division is carried out; over 2000 images of healthy and unhealthy potato leaf are taken from Kaggle, and a few pre-made models are used to identify and classify healthy and diseased leaves. With 30% test data and 70% train data, the algorithm successfully predicts one of them with an accuracy of 91.41% in testing. Our results demonstrate that CNN outperforms all currently available tasks for detecting potato disease. Soft computing technology is desired by plant pathologists for the precise and dependable diagnosis of plant diseases.**

Keywords—Potato disease, late blight, early blight, CNN, image processing, V3, VGG16, VGG19, Feature extraction.

DOI Number: 10.48047/NQ.2022.20.12.NQ77741 **NeuroQuantology2022;20(12): 4147-4151**

I. INTRODUCTION

All throughout the world, people are familiar with potatoes, which are also a staple diet in many nations. Another name for potatoes is the root of all veggies. Because India is a country with a large agricultural sector and produces a variety of crops, potatoes play a significant role in our economy. India is the world's second-largest producer of potatoes. About 43,000,000 tons of potatoes were produced in India in 2018. The key factor is expanding potato production because the demand is steadily rising worldwide, and it is necessary to sell as much as our region can.

However, the reality is that due to a dangerous sickness that has been affecting the produce and export in recent years The disease can occasionally be seen on the infected potato leaf. The plant's leaves will occasionally develop spots as well. The early and late blights are the most typical potato diseases. The majority of the early blight signs are tiny, black lesions, while the late blight symptoms can appear blistering boiling water, gradually rotting, and drying out. Using CNN technique, these pathologies will be distinguished from potato leaves.

Particularly advantageous for farmers. There are three different categories of processed photographs available. They are healthy, early, and late blight. The total amount of two areas of pictures are isolated, one for training and the other for testing. Around 75% of the images are in the training section.

Disease identification and segmentation are crucial, however because soybean diseases are active in the real world, conventional segmentation algorithms like ask-means and color-based segmentation cannot produce segmentation results with speed and accuracy.

However, deep convolutional neural networks (CNNs), a recent development in machine learning, has shown that learned representations are more efficient and effective. The key benefit of as a result of representation learning, computer programs can automatically recognize objects in large collections of photos traits that can accurately identify photographs.

Few researchers have suggested using CNN to classify plant diseases and identify leaves. Using straightforward leaf photos of healthy and ill plants, the reported convolutional neural network models execute plant and disease, detection, and classification tasks with an accuracy of 99.53%. Convolutional neural networks are used by all

of these suggested methods as both feature extractors and classifiers. We found that combining the CNN network with a shallow classifier could improve network performance.

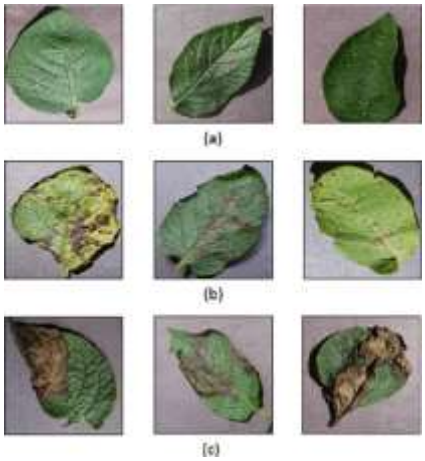


Fig. 1. Infected Potato leaves

II. LITERATURE REVIEW

I. Krishna Mitra: Using Machine Learning to Identify Diseases in Plants is a research paper. Here, they used CNN model methodology's Tensor Flow Framework to carry out their project. This model's benefit was that it only calculated leaf area and fungi induced illnesses in Sugarcane is easily identifiable. The drawback of was that it required High Computational Complexity to implement. In a research paper titled "Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment". Researchers used Fuzzy c-mean clustering, Neural Network in implementing the model. Does not need special training to farmers as dataset has images in different angles was the main advantage of the model. The only disadvantage was that the images captured by untrained farmers were not oriented and contain cluster of leaves with background visible in several segments.

In a research paper titled "Potato Disease Detection Using Machine Learning". Image processing was the used technology over here. Use of CNN.

Model gained 90% validation accuracy was the major pro in this project. Large Training Model Is Required was the major drawback of this model.

The deep learning method mask R- CNN was suggested in a study titled "Detecting the Infectious Area and Disease Using Deep Learning in Tomato Plant Leaves." The advantage of this was that disease identification in plant leaves may be improved and speeded up by utilizing the R-CNN mask. Its drawback was that it required extra.

II. Santanu Phadikar and Jaya Sil (2008): combined image growth and image segmentation techniques to describe a software prototype system for disease diagnosis.

III. Ajay A. Gujar (2010): Regularization and extraction technology was researched by et al. They described the Eign features of this technology and found that it provided greater accuracy than other detection feature technologies.

IV. Geng Ying (2008): et al. investigated image processing techniques. Cucumber powdery, speckle, and downy mildews were employed as study samples for this purpose in order to compare the effects of simple and medium filters.

V. H. Al-Hairy & et al. (2010): explains the three ways to diagnose leaf disease: 1) To identify the affected using K-means Clustering, a leafy portion. 2) Using color cooccurrence to repair the damaged portion of the leaf way for analyzing texture. 3) To use neural networks (NNs) to find and categories the different types of diseases. in depth.

Obtaining RGB photos of leaves and applying a color transformation structure come first. Following that image is segregated using the K-means clustering methodology, with the green pixels' values masking and removed before using Otsu's method to get the object's threshold value. The zero value is set for the RGB images when converting the color co-occurrence approach. Following that, infected clusters were converted to Hue Saturation Value (HSV), and each image's texture analysis was performed using the SGDM matrix. Finally, neural networks were used to implement the solution after realizing the problem.

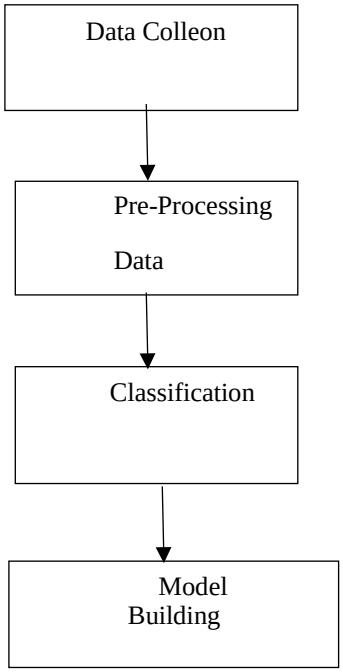
VI. J. Garcia Arnal Barbedo et al. (2018): utilized photographs to identify and recognize illnesses using the Annotated Plant Pathology Databases.

VII. Adnan Mushtaq et al. (2019): proposed a model for identifying plant diseases and suggests remedies through coevolutionary neural networks. et al, Raghavendran S.

VIII. Liu, LOuyang, W and Wang, X et al. (2020): more than 300 research publications were analyses, covering a variety of viewpoints on the detection of generic objects, such as frameworks for identification, representations of object attributes, creation of object proposals, context modelling,training approaches, and metrics for evaluation.

III. METHODOLOGY

In this project's study has gone through several stages. in the shape of an analytical framework. There are four steps in the suggested research framework, which are as follows:



A. *Data Collection*

The images of potato leaves sorted into three categories—healthy leaves, early blight, and late blight—make up the dataset used in this project.

The "Plant Village Dataset" dataset was downloaded from the Kaggle website. Data used is as follows:

Samples	Number
Healthy leaf	150
Early blight	500
Late blight	500
Total	1150

TABLE I. TABLE DATASET DETAILS

CNN is used to compute potential maps by using the activation functions. The function is:

$$y_j^l = f(z_j^l)$$

here, y_j^l is the forthcoming graph, & $f(z_j^l)$ is the activation function.

Documents are treated in CNN in 2-D convolution operations.

$$O = \frac{(W-F+2P)}{(S+1)}$$

Here, height, input height, filter size, padding & stride are given by O, W, K, P, S.

B. *Pre-Processing Data*

Currently, 1150 potato leaf photos from three classes—healthy, early

Dataset	80:20		70:30	
	Train	Val	Train	Val
Late Blight	400	100	350	150
Early Blight	400	100	350	150
Leaf Healthy	122	30	105	45
Total	922	230	805	345

Table 2. test-train dataset details

C. *Classification*

After the above processes, further we need to classify images using Convolutional Neural Network (CNN) architecture. CNN is a supervised learning method where identification of image by training existing dataset and targeting image variables. The convolutional layer in CNN helps neural network to recognize potato leaves based on the attributes that they have. Neural network uses pixels in the picture to recognize images of potato leaves. In this project we will be using image with size 150X150X3; where it will have three channels, that are red, green, and blue. The leaf image will be convoluted with a filter firstly. Then pooling will be applied to reduce the resolution of the image with keeping its quality intact. MaxPooling will be used on output image. In the next step, this layer us being flattened. This will change the feature map resulting from pooling into vector form. This project has the proposed model for CNN architecture in identifying diseases in potato leaves using 4 convolutional layers and 4 MaxPooling.

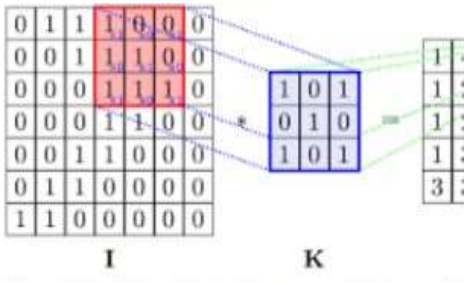


Fig.2. Represents the convolution us in K over the image

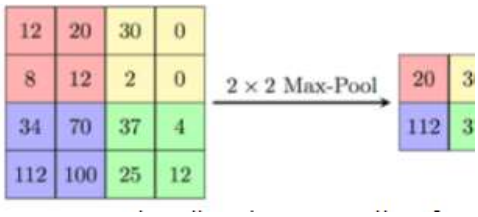


Fig.3. describes the max-pooling from window and reducing the dimension form

D. Model Building

The sequential model is created using CNN. We have rectified linear units for this. Additionally,
SoftMax is utilized as a predicting activation.

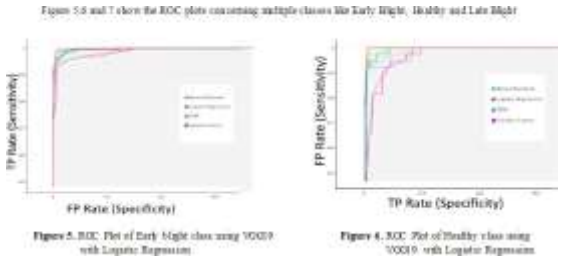
which depends on the maximum likelihoods. The equation for the SoftMax function is given below: Here, X*T* W signifies X and W's internal product.

RESULT

A plant village dataset that includes 1000 photographs of early blight, 1000 images of late blight, and 152 images of healthy potato leaves was used to create the 2152 images of leaves used in this proposed model. The training portion of the dataset has 1700 images (70%) while the test portion contains 452 images (3%). Several previously trained models,
including inceptionV3, VGG16, and Among the VGGs utilized for feature extraction, VGG19 produced the best results, as shown in table 3 below. For classification, a variety of classifiers including KNN, SVM, Neural Network and logistic regression are used. of which
Logistic Regress ion provides the most advanced solution.
has a 97.7% classification accuracy. Performance measures like AUC (Area Under the Curve), CA (Classification Accuracy), Precision, Recall, and F1-Score are calculated to assess the effectiveness of the model.

Table 2 comparison report with other models

Model Proposed	**Classification Accuracy**
Image segmentation + backpropagation neural network[2]	92%
Image segmentation + Support Vector Machine[3]	95%
Proposed Approach	**97.8%**



$$P(x) = \frac{e^{x^T w^l}}{\sum_{k=1}^{k} e^{x^T w^l}}$$

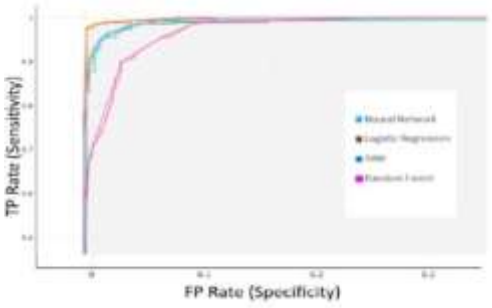


Fig.7. ROC plot of Late blight class using VGGI 9 with Logistic Regression

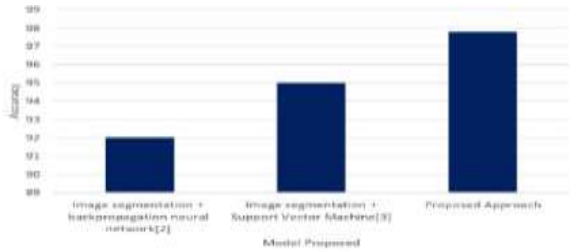


Fig.8. Classification accuracy

V. CONCLUSION

This study proposes a classification-based approach to identify late blight, early blight, and healthy leaf photos of potato plants using deep learning techniques and convolution neural networks. We discovered that CNN works well for this kind of item classification. This
Model improves validation accuracy by 91.41%. This kind of undertaking, in our opinion, will be extremely important for the agriculture industry. the majority of farmers in the
Indian villages lack literacy and have incomplete knowledge of the disease. We believe that this work has the potential to alter the Indian potato farmer. In order to accomplish classification, studies have been done on photos of both healthy and sick leaves. The proposed method successfully distinguishes between three different types of potato leaf diseases, it is concluded.

VI. FUTURE SCOPE

The potato grower in India. Studies on images of both healthy and ill leaves have been conducted in order to

achieve classification. It is determined that the suggested method successfully distinguishes between three different types of potato leaf diseases.

The database will subsequently be connected to the interface over the internet. We'll also try to apply these kinds of suggested methods to a variety of tasks, like identifying different kinds of plants with leaves.

REFERENCES

[1] MAJJI V APPLALANAIDU, G. KUMARAVELAN "A Review of Machine Learning Approaches in Plant Leaf Disease Detection and Classification" In 2021 IEEE Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021), pp 716-724 IEEE, 2021

[2] Md. Khalid Rayhan Asif, Md. Asfaqur Rahman, Most. Hasna Hena "CNN based Disease Detection Approach on Potato Leaves" in 2020

Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020], pp 428-432 IEEE,2020

[3] Parul Sharma, Yash Paul Singh Berwal, Wiqas Ghai "KrishiMitr (Farmer's Friend): Using Machine Learning to Identify Diseases in Plants" In 2018 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), pp 29-34. IEEE 2018.

[4] Mustafa Merchant, Vishwajeet Paradkar, Meghna Khanna, Soham Gokhale. "Mango Leaf Deficiency Detection Using Digital Image Processing and Machine Learning" 2018 3rd International Conference for Convergence in Technology (I2CT), pp 1-3. IEEE 2018.

[5] Melike Sardogan, Adem Tuncer, Yunus Ozen. "Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm" 2018 3rd International Conference in Computer Science and Technology. pp 382-385. IEEE 2018.

[6] H. Al-Hiary, S. Bani-Ah Mad, M. Reyalat, M. Braik And Z. A Lrahamneh, Fast and Accurate Detection and Classification Of Plant Diseases, IJCA, 2011, 17(1), 31-38, IEEE-2010

[7] Geng Ying, Li Miao, Yuan Yuan &Hu Zelin [2008] A Study on the Method of Image PreProcessing for Recognition of Crop Diseases, International Conference on Advanced Computer Control, 2008 IEEE.

[8] Ajay A. Gurjar, Viraj A. Gulhane," Disease Detection on Cotton Leaves by Eigenfeature Regularization and Extraction Technique", International Journal of Electronics, Communication & Soft Computing Science and Engineering (IJECSCSE) Volume 1,Issue 1

[9] Santanu Phadikar & Jaya Sil [2008] Rice Disease Identification Using Pattern Recognition Techniques, Proceedings Of 11th International Conference on Computer and Information Technology, 25-27

[10] Garima Tripathi and Jagruti Save : "AN IMAGE PROCESSING AND NEURAL NETWORK BASED ROACH FOR DETECTION AND CLASSIFICATION OF PLANT LEAF DISEASES", Volume 6, Issue 4, April (2015), pp. 14-20.

[11] P. Revathi, M. Hema Latha, Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques, ISBN, 2012, 169-173, IEEE.

[12] Ungsumalee Santagati, Aekapop Bunpeng. "Potato Leaf Disease Classification Based on Distinct Color and Texture Feature Extraction" 2019 19th International Symposium on Communications and Information Technologies (ISCIT). pp 82-85. IEEE 2019.

[13] Islam, Monzurul, Anh Dinh, Khan Wahid, and Pankaj Bhowmik. "Detection of potato diseases using image segmentation and multiclass support vector machines." In 2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE), pp. 1-4. IEEE, 2017.

[14] D. P. A. A. Mr. Suman Kumar Swarnkar, 'Improved Convolutional Neural Network based Sign Language Recognition', International Journal of Advanced Science and Technology, vol. 27, no. 1, pp. 302–317, 2019.

[15] V. Verma, Suman Kumar Swarnkar, and V. Swarnkar, 'Iris Recognition System Using Adaptive Neuro Fuzzy Inference System Classifier', 2017.

[16] B. Chandel and Suman Kumar Swarnkar, 'STOCK PREDICTION USING FINANCIAL DATA AND NEWS SENTIMENTAL ANALYSIS', 2021.

[17] Suman Kumar Swarnkar, A. Ambhaikar, A. Guru, and A. Sheikh, 'IOT based device for social distancing to stop spreading coronavirus'. AU Patent 2,021,102,602, 01-Jul-2021.

[18] V. K. Swarnkar, A. Ambhaikar, and Suman Kumar Swarnkar, 'BIG DATA SECURITY ENHANCEMENT BASED INTRUSION DETECTION SYSTEM USING K-MEAN CLUSTERING OF DECOMPOSITED FEATURES'. IN Patent 202,121,004,801, 02-Dec-2021.

[19] Suman Kumar Swarnkar, A. Ambhaikar, V. K. Swarnkar, and U. Sinha, 'Optimized Convolution Neural Network (OCNN) for Voice-Based Sign Language Recognition: Optimization and Regularization', in Information and Communication Technology for Competitive Strategies (ICTCS 2020), 2020, p. 633.

[20] S. Agarwal, J. P. Patra, and Suman Kumar Swarnkar, 'Convolutional Neural Network Architecture Based Automatic Face Mask Detection', International Journal of Health Sciences, no. SPECIAL ISSUE III, p. 6232629, 2022.

[21] Suman Kumar Swarnkar, Vivek Chandrakar, 'A Naive Bayes Approach for Word Prediction using NLP', International Journal for Scientific Research and Development, vol. 6, no. 7, pp. 401–407, 2018.

[22] Suman Kumar Swarnkar, Ruby Chandrakar, 'Telemedicine Cryptography using DNA Sequence', International Journal for Research in Applied Science and Engineering Technology, vol. 5, no. 12, pp. 2122–2129, 2017.

[23] Suman Kumar Swarnkar, Bindeshwari Chandel, 'STOCK PREDICTION USING FINANCIAL DATA AND NEWS SENTIMENTAL ANALYSIS', Journal of Xi' an Shiyou University, Natural Science Edition, vol. 16, no. 9, pp. 279–285, 2020.

[24] Suman Kumar Swarnkar, G. S. Chhabra, A. Guru, B. Janghel, P. K. Tamrakar, and U. Sinha, 'Underwater Image Enhancement Using D-Cnn', NeuroQuantology, vol. 20, no. 11, pp. 2157–2163, 2022.

[25] Suman Kumar Swarnkar, A. Guru, G. S. Chhabra, P. K. Tamrakar, B. Janghel, and U. Sinha, 'Deep learning techniques for medical image segmentation & classification', Universidad Tecnica de Manabi, vol. 6, no. S10, pp. 408–421, 2022.

[26] Suman Kumar Swarnkar, Bhumika Chandrakar, 'A Novel Appraoch for Brain Tumor Detection using SOM and Fuzzy Classifier', International Journal of Emerging Technologies and Innovative Research, vol. 5, no. 6, pp. 413–420, 2018.

Suman Kumar Swarnkar, Ms. Preeti Bhole and Mr. Virendra Swarnkar, 'HYBRID BASED RECOMMENDATION SYSTEM USING SWITCHING TECHNIQUE', International Journal for Science and Advance Research In Technology, vol. 3, no. 11, pp. 103–109, 2017.

eISSN1303-5150