JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

A CNN Model for Disease Detection in PotatoLeaves

Pappu Sah Sudi
Department of CSE
*New Horizon College of Engineering*Bangalore,
India pappusah98120@gmail.com

Nihit Gupta
Department of CSE
New Horizon College of Engineering
Bangalore, Indianihit67@gmail.com

Kamlesh Kumar Baniya
Department of CSE
New Horizon College of Engineering
Bangalore,India
jaiswalprince1224@gmail.com

Ms. Sindhuja Bangari
Department of CSE
New Horizon College of Engineering
Bangalore, India
sindhuja.nhce@gmail.com

Dr. Rachana P
Department of CSE
New Horizon College of Engineering
Bangalore, India
dr.rachanap@newhorizonindia.edu

ABSTRACT

India is a predominantly agricultural country, and its crop production rate is a key source of concern. We created a strategy to boost agricultural produce rates and prevent disease infection in plants, making farmer work easy to some extent. They can use deep neural Networks, an AI subfield, to detect damaged plants without having to do so manually, and to cure them before it's too late. This study looks at a variety of experiments and suggests that Convolution Neural Networks (CNN) are better at identifying leaf diseases than other methods. CNN also contributes the highest level of disease detection accuracy possible. Keywords: Neural Networks (NN), Convolutional Neural Network (CNN), Architecture, Deep Learning (DL).

I. INTRODUCTION

Covid -19 has an impact on all living things. Humans recognise that excellent health only protects us against diseases and viruses. Agriculture is an example of a health-oriented industry. In every country, agriculture will play a significant role. One such country is India. The majority of Indians work in agriculture. During farming, they plant a number of crops, the most adaptable of which is the potato, which accounts for 28.9% of India's total agricultural crop production [2]. The potato, as we all know, is the king of vegetables and may be paired with almost any other vegetable. White potatoes, according to a study, are the most abundant and cost-effective potassium source of any vegetable or fruit. Because we recognise the importance of potatoes, it is our responsibility to ensure that people have access to sufficient quantities of fresh potatoes. Potato plants must be properly farmed in order to produce large quantities of potatoes. Disease infection in plants should be limited for the best output, which farmers are aware of. However, because doing so manually takes time, We may be able to save time and have more confidence in our ability to detect and treat disease while it is still in its early stages. Farmers nowadays are often ignorant of a disease that has impacted their crops at an early stage. In the early stages of crop development, there is no reliable approach for detecting infections. Artificial intelligence is currently widely used. With its recent advancement, it is also beneficial to the agriculture industry. Deep Neural Networks (DNN) is a branch of artificial intelligence. Two principles are necessary to identifydisease in leaves.

1. Object Recognition
2. Image Classification

The goal of the study is to identify and diagnose the plant's sicknessEarly blight and late blight are fungal infections that mostly affect the leaves of potatoes [2].The distinctions between healthy and unhealthy leaves are depicted in the accompanying diagram.

Fig. 1: Health leaf / Early blight / Late blight

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

CNN is one of DNN's algorithms. CNN converts the input image into a vector representation more effectively. So that it can do a wide range of vision tasks. It is obvious to use the CNN as an image "encoder" by pre-training it in the image. According to recent study, CNN performs best in object recognition and image categorization. The study is organized in the following way: section I discusses relevant work, section III discusses technique, and sections IV and V explore conclusion and futurescope.

II. RELATED WORK

The study on detecting plant leaf disease is explained in the paragraph below. As can be seen, the best potential accuracy foridentifying plant leaf disease is evaluated using CNN (Convolutional Neural Network).

In [1] to detect plant ailments, researchers used images of various leaves to build a deep learning model that relies on a custom architectural convolution network.

It created a self-learning system for detecting and recognizing diseases in potato leaves, such as early blight, late blight, and healthy, in [2], with a categorization accuracy of about 98 percent across multiple test datasets. Farmers may readily diagnose diseases in their early stages using this technology, allowing them to increase crop productivity.

Using CNN, it aids in the identification of the illness induced in potato leaves in [3]. The illness in potato leaves was classified with 99 percent accuracy using a 700-800 picture dataset. The downside of the dataset with the above-mentioned image is that

more photographs may be used to obtain more accurate results

[4] Suggested is efficient, quick, and precise, with a high prediction rate. It has the ability to identify sickness in plants and also gives treatment for the same. To increase the plant's health and productivity, we must have a thorough understanding of the disease and its treatment. The framework is implemented using the CNN algorithm with Python and has an accuracy of 80%.

Dense Net is used to discover and classify various leaf diseases in [5] using only three plants: tomato, potato, and bell pepper. The Plant Village dataset is recognised, and all data is recorded using a mobile camera and live detection. The goal of this system or study in the future is to increase the model's accuracy rate, reduce computation for small computers such as mobile phones,and produce a mobile app that is more user-friendly and can beutilised immediately.

In paper [6] employed a variety of machine learning algorithms such as nave Bayes, KNN, and SVM to achieve accuracy of 88.67, 94.00, and 96.83 percent.

In [7] looked explored how a Convolutional Neural Network model may predict plant diseases using images from a particulardataset in the field and previous informational indexes.

A technique for identifying and ordering leaf disease is used in [8]. K-Means division is used to complete the sick part division. According to this article, deep learning algorithms could be employed in the future to detect a variety of plant ailments. The

[9] investigates different ways for developing a decent classifier for potato leaf disease using an RGB image. They recommend creating a big input dataset based on the findings of this investigation. It examines all of the features learned by the models to see if they are representative of the overall organisation of a picture rather than just the leaves and affected region.

[10] Using the MATLAB Image Processing and Bioinformatics Tool Box, creates a fast, self-contained, and error-free system for disease identification and categorization of potato leaves. The technology might be used to identify and categorise diseases onPotato Plant leaves, with the goal of eventually identifying and classifying diseases in humans.

Comparison of model performance:

Table 2 shows the numerous algorithms that have been used to identify potato disease, as well as the accuracy that each algorithm has delivered.

Table 1: Model performance comparison

S.No	Algorithm	Ref No.	Accuracy
1	Ann	[17]	85-91%
2	NN	[38]	93%
3	BPNN	[10]	92%
4	Naive Bayes	[6]	88.67%
5	KNN	[6]	94.00%
6	SVM	[6]	96.83%
7	SSD &RCNN	[1]	94.60%
8	CNN	[3]	99.09%

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

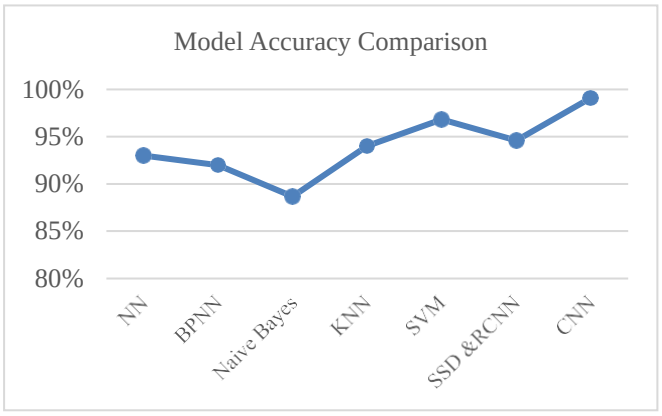

Fig. 2: Model accuracy comparison

According on the findings of the literature review, CNN outperforms alternative classification algorithms.

III. PROPOSED METHODOLOGY

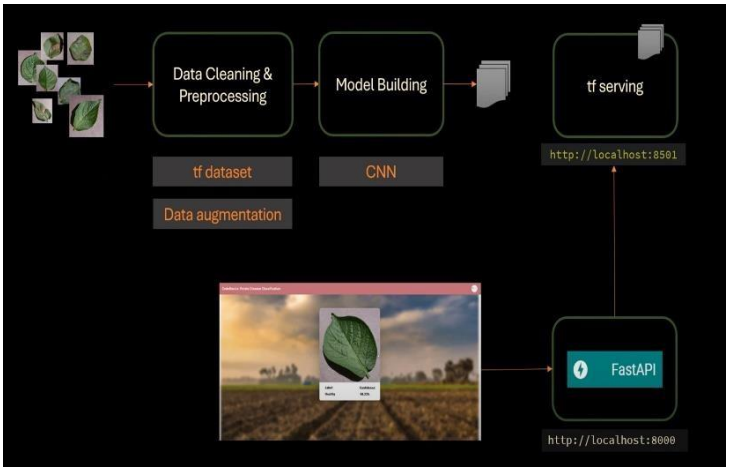

Fig. 3: - Process involved in disease classification usingCNN

Data Collection: Any algorithm's evaluation and implementation begin with this stage. For a high accuracy outcome, the more data you have, the higher the accuracy prediction. Data gathering is the first step in the proposed method. There are over 2150 images in this dataset, with a variety of leaf patterns (Early blight disease, late blight disease, and healthy leaf).

Table 2 describes the dataset for detecting potato leaf disease. It is divided into three categories. There are three types of blight: early, late, and healthy. This dataset also divides diseases into categories and counts the number of pictures it contains.

The accuracy comparison between several algorithms is shown in the figure below cheval 2. CNN has the most accuracy, as shown in the graph below.

Table 2: Number of images with diseases

Disease	Type of Disease	No. of Image
Early Blight	Fungal	1000
Late Blight	Fungal	1000
Healthy	No Disease	150

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

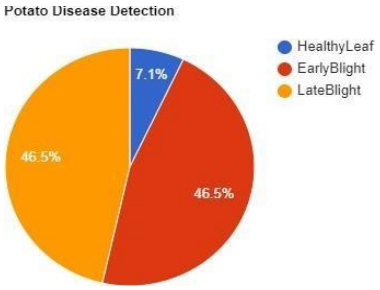

Fig. 4: Each label with number of images

Pre-processing: For accurate results from any dataset, preprocessing is essential. Because multiple photos with varied patterns can affect the classification result, only the leaf of the potato plant is focused in this proposed approach's images, which are all cropped to the same size.

Noise cancellation: Only the primary features that are required for the classification process are maintained in the image before it is sent on to the next step. All other features are eliminated. Filtering is the process of making an image smoother andlowering the contrast in order to capture just the most crucial and relevant characteristics.

Fig. 5: - Dataset

Figure 5 provides an exact image of the leaf; no other things are visible in the photograph except the leaf, which is the samesize as everything else.

Feature Extraction: - After removing noise from pictures, we must do feature extraction. Feature extraction was used to reduce the size of an image without sacrificing crucial details. It also gets rid of the useless features.

Classification: - Any classification algorithm or deep neural network, such as ANN, CNN, or SVM, can be used to classify a picture to a particular target class label. When compared to other algorithms, CNN is a neural network that is used for picture categorization since it has the highest accuracy.

Model Building: Convolutional neural networks (CNNs) are a type of deep learning neural network that belongs to the convolutional neural network family. CNNs are a huge step forward in image recognition technology. They are commonly used in the classification of images and are routinely used to analyze visual imagery.
Convolutional Neural Networks: These are two-dimensional neural networks that receive input in two dimensions. Because images can be represented in comparable ways, using neural networks for image pre-processing makes sense. Convolution is unusual in that it produces visible characteristics in photographs. Before constructing a feature map that summarizes the existence of observed features, it filters the input. These networks learn the filters in the context of a given prediction problem during training. By multiplying the filter with the input array once, a single value is produced. When the filter is applied, it produces a two-dimensional vector called features. Once those are built, they are sent to the feature map via non-linearity, such as the ReLU, for the outputs of the fully linked layer. The architecture of CNN is seen in Figure 1. A CNN's architecture is as follows:

1. Convolutional (conv) layer
2. ReLU (Rectified linear unit) layer
3. Pooling layer,

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

4. Pair of fully-connected (FC) layers.

The majority of the calculations are done in the convolution (Conv) layer. These layers operate as feature extractors when a feature is identified in the input image, generating the feature map. This layer uses a kernel based on image characteristics to execute convolution operations, as well as attempting to extract picture features using the kernel. Pooling layers and Relu.

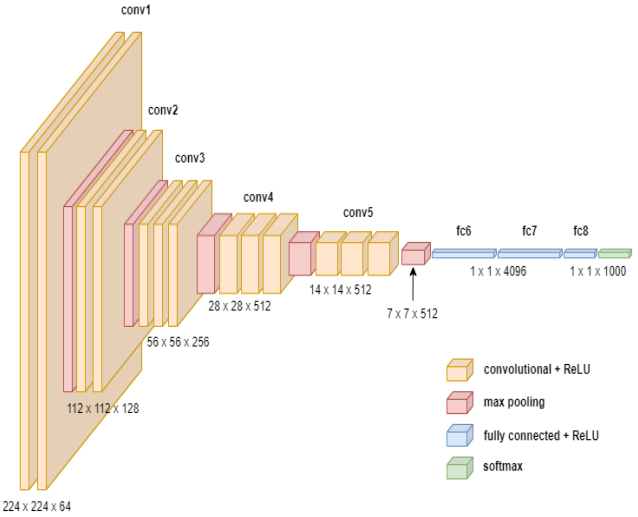

Fig. 6: Architecture of Convolutional Neural Network

Relu layer: Relu is a layer of activation that works in aformulaic manner.

Relu= maxf_0(0,x)

The non-linearity is caused by the Relu operation, which will becomputed after the convolution process is completed.

All negative values are converted to 0 by the RELU layer, which speeds up the training process. By sharing the weight of neurons in feature maps, the pooling layer serves to lower the spatial resolution of feature maps. The pooling layer is used to computeaverage pooling, maximum pooling, multi-resolution order less pooling, and stochastic pooling.

Flatten Layer: A two-dimensional dataset is reduced to a single feature vector using this layer.

Fully Connected Layer: The neurons in this layer are perfectly linked with those in the preceding layers. The softmax technique is used by the fully connected layer to find an input picture utilising a feature vector as input. Back propagation will be employed to lower the error value in this layer.

The CNN working method in Plant leaf detection categorization is shown in fig. 4. It will take a 2D vector representation of an image as input and conduct the operations listed above before categorising the image

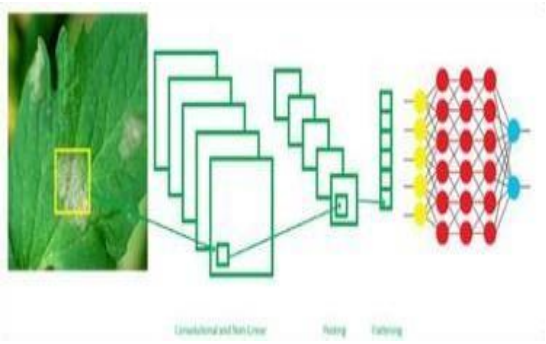

Fig. 7: CNN illustration in plant leaf detection

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

The following diagram illustrates the numerous phasesinvolved in image classification.

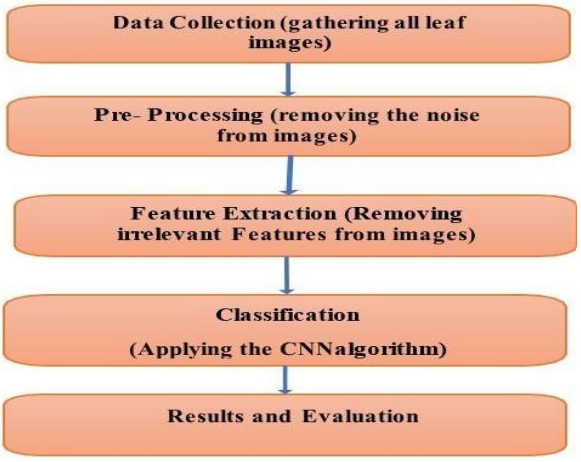

Fig 8: Some basic algorithm for Image Classification

IV. RESULTS AND COMPARISIONS:

The leaf classification model was implemented and we were able to classify an image which was clear that leaf was either infected or healthy. By displaying the confusion matrix and watching the diagonal values for measuring the amount of accurate classifications, an individual might determine the model's accuracy.

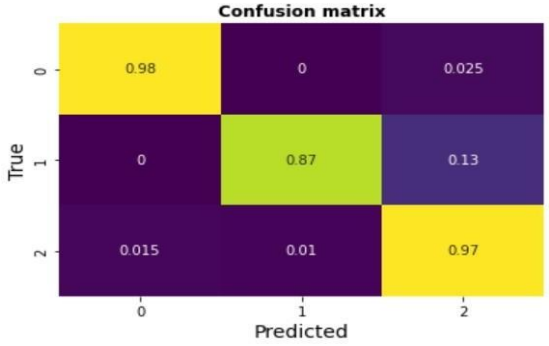

Fig. 9: Confusion Matrix

Our model outperforms the competition in all classes. Late blight affects 13% of otherwise healthy leaves. We can address this by adding more photos for healthy class. But, for now, 87 percent accuracy isn't terrible! Thus concluded result was morethan the manual prediction (farmer prediction).

Fig. 10: End result of model

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1929 - 1935
https://publishoa.com
ISSN: 1309-3452

In the end result we were able to build neural network model and successfully deployed it and got the result. In above figure it is clearly observed that leaf image is imported to the interface and result is displayed with the message of which disease plantleaf is infected of.

IV. CONCLUSION

Our end to end system neural network is capable of classifying an image of leaf whether it is healthy leaf or infected one. A convolutional neural network encoder decoder was included in the model, which aid in leaf image classification. This publication summarized the findings of numerous papers that used deep neural networks to detect illness in potato leaves. The two kinds of potato leaf disease are Early blight disease and Late blight disease. When compared to other deep neural networks, CNN was discovered to have the highest accuracy in recognizing and classifying illnesses in potato leaves such as early blight disease, late blight disease, and non-infected. According to this study, Artificial Neural Networks (ANN) had an accuracy of 85 percent, Support Vector Machines (SVM) had an accuracy of 88.89 percent, and CNN [3] had an accuracy of
99.07 percent. As a result, it has been discovered that Convolutional Neural Networks provide the highest level of accuracy, and in the realm of Deep Learning, it is a relatively recent technology that has been used to recognize images and distinguish unusual features from them.

REFERENCES

[1] sumit Kumar, Veerendra Chaudhary, Ms. Supriya Khaitan Chandra, Sumit Kumar,"Veerendra Chaudhary, Ms. Supriya Khaitan
Chandra",[2021]

[2] Divyansh Tiwari, Mritunjay Ashish, Nitish Gangwar, Abhishek Sharma, Suhanshu Patel,'Dr. Suyash Bhardwaj,"Potato Leaf Diseases Detection Using Deep Learning",[2020]

[3] Sanjay K Hoysala, Santosh K S, Srihari D R, Suhith Shekar Chandra, Ajay Prakash B V," Recognition and Classification of Leaf Disease in Potato plants",[2019]

[4] Pitchai R1, Sharath Kumar G2, Ashutosh Varma D3, Madhu Babu
CH4," Potato Plant Disease Detection Using Convolution Neural Network",[2020]

[5] A T Madhavi,MMYugesh,Srinidhi k, C S SVijayasundar,"Detection of leaf disease using convolutional neural network" ,[2020]
S Jeyalakshmi, R Radha, "An effective approach to feature extraction for classification of plant diseases using machine learning",[2020]

[6] Gautam Trivedi , Ankur Yaduvanshi , M.Hema," IDENTIFICATION OF DISEASES IN PLANTS USING DEEP LEARNING",[2019]

[7] R. Mounika, Dr. p. Shayamala Bharathi," DETECTION OF PLANT LEAF DISEASES USING IMAGE PROCESSING",[2020]

[8] Soma Ghosh, Renu Rameshan and Dileep A. D.," An Empirical Study on Machine Learning Models for Potato Leaf Disease Classification using RGB Images",[2021]

[9] Mr. Girish Athanikar , Ms. Priti Badar, " Potato Leaf Diseases Detection and Classification System ",[2016]

[10] V. Veerasamy et al., "LSTM Recurrent Neural Network Classifier for High Impedance Fault Detection in Solar PV Integrated Power System," in IEEE Access, vol. 9, pp. 32672-32687, 2021, doi: 10.1109/ACCESS.2021.3060800.

[11] N. Kumar, Y. S. Sneha, J. Mungara and S. G. Raghavendra Prasad, "A Survey on Data Mining Methods Available for Recommendation System," 2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), 2017, pp. 1- 6, doi: 10.1109/CSITSS.2017.8447672.