Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: https://doi.org/10.22214/ijraset.2022.45458

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

Potato Diseases Detection Using Machine Learning Techniques

Harisha J[1], Mrs. Renuka Malge[2]
[1]PG Scholar, VTU, CPGSB, Muddenahalli, Chikkaballapur-562101
[2]Asst. Professor, Dept. Of CSE (MCA), VTU, CPGSB, Muddenahalli, Chikkaballapur-562101

Abstract: It is becoming more crucial than ever to do thorough study on actual rural occurrences, especially with the development of agricultural equipment and the application of artificial intelligence to the detection of plant diseases. The quantity and quality of potatoes are significantly impacted by several illnesses including early blight and late curse, and translating these leaf infections manually takes time and is uncomfortable. The output of potatoes can be increased by using skilled and automated identification of these problems during in the sprouting stage until it requires a high degree of competence. For the purpose of identifying a small number of plant diseases, several models have already been developed. In this work, a model that uses ready-made models like VGG19 to fine-tune (transfer learning how) the extraction of pertinent components from the dataset is presented. Then, with the use of several classifiers, it was shown that, with a score of 97.8% just on test dataset, strategic relapse beat those in term of order precision.
Keywords: Catching highlights, calculating relapse, starting with V3, VGG16, VGG19, and adjusting.

I. INTRODUCTION

There are several types of vocations on the globe, but farming is unquestionably the most important of them all. The Indian economy, which is heavily reliant on horticulture, is not an exception. Potato is the most adaptable crop, accounting for roughly 28.9% of total rural harvest creation in India. After maize, wheat, and rice, potatoes are the world's fourth most important rural food crop. With an annual production of 48.5 million tonnes, India is the second-largest potato grower in the world [23]. The Horticultural & Processed Food Products Trade Promotion Authority of Uttar Pradesh produces roughly 30.33 percent of the country's output of potatoes (APEDA). In the textile industry, farina, or potato starch, is used to measure cotton and worsted fabrics. Potassium, vitamins (particularly C and B6), and fibre are all abundant in potatoes. It helps treat conditions including heart disease, coronary sickness, and disease by lowering the overall level of blood cholesterol.

Infections have a detrimental effect on horticulture grounds and plants. The major causes of these ailments are microorganisms, genetic diseases, and enticing specialists such small organisms, growths, and infections. Diseases of the potato leaf are often brought on by growths of microscopic organisms. While delicate deterioration and common scab are bacterial illnesses, late scourge & earlier curse are parasite infections [22]. Our research into developing a robotic strategy that might increase harvest productivity, rancher profit, as well as commitment to the national economy has been prompted by the discovery and diagnosis of these illnesses on such important plants.

Earlier, a number of authorities in the field of image processing techniques suggested using popular image processing methods as K-implies clustering [13] and LBP [14] to distinguish between different leaf diseases. Profound learning models are better component producers because they have more planning skills. In this work, we used a few classifiers to create a deep learning model to identify potato leaf diseases. The essay is divided into three sections: the first section is a presentation, the second is a study of writing, and the third is a conclusion.

II. RELATED WORKS

There are a variety of ways for identifying plant illnesses, and different scientists have recommended several procedures for identifying potato leaf infections in their research. This section gives a rundown of such strategies.

P. Badar et al. [2] applied division study methodology based on The k Mean Clustering [13] to different features of Leafy vegetables photo testing such as color, surfaces, region, and so forth in order to differentiate and rank the illness in the leaf picture with such a 92 percent characterisation precision. U. Kumari et al. [4] divided many aspects of a picture, including variance, connectivity, power, uniformity, average, standard deviation, and volatility, among others, using an image segmentation approach. There are several strategies for finding infections on potato leaves, and numerous researchers have proposed a variety of techniques for doing so. A list of such tactics is provided in this section.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

After that the accents have been eliminated, a Neural Network is frequently used as a classifiers to identify the group illnesses also on leaflets of two varieties, such tomato and cotton. They were able to achieve a categorization accurateness of 92.5 percent using this method.

A. Analysis OF Data – set

The Vegetation Village Dataset may be used to investigate inspirations using Kaggle, an open-source resource [1]. The collection contains over 55,000 images of both healthy and diseased leaves from a range of agricultural crops, including apples, blueberries, cherries, grapes, peaches, peppers, lemons, tomatoes, and potatoes. Each organiser for vegetables that emerge from the ground contains two distinct picture types: grayscale and shaded. Each harvest contains several types of leaf disease, and each form is classified as a different class of illness for categorization purposes.

Each photo in Dataset [1] has both a leaf picture with foundation as well as a leaf photo without a foundation

Figure 1: Typical images of each class

Every class has between 152 and 1000 images. We only used potato photos from three distinct classes, including Bacterial Blight, Late Blight, & solid leaf photographs, for our characterisation challenge. Details on the train-test divide is shown in Table 1.

Table 1 shows that the total number of observations in the training and test datasets for our model.

Label	Category	Number	Training sample	Test sample
1	Early Blight	1000	787	213
2	Late Blight	1000	791	209
3	Healthy	152	122	30
Total		2152	1700	452

B. Proposed Approach

1) Include Extraction Using VGG19

VGG19 is a CNN-based method that was proposed by K. Simonyan and A. Zisserman [7]. This model was developed using the ImageNet dataset, which consists of approximately 15 million identified high-goal pictures arranged into 22000 categories. The ImageNet LargeScale Visual Recognition Challenge (ILSVRC) requires 1.3 million photos for preparation, 50 000 for approval, and 100 000 for testing. This model was developed for the ILSVRC. The sole preprocessing in VGG19 [7] is the removal of each pixel's mean RGB value.

By substituting all of the huge Kernel-sized channels with various 3*3 piece approximation pooling across a time window of 2*2 with a step value of 2, the VGG19 [7] model improved order accuracy in comparison to AlexNet. The model's last layer is the softmax layer. Each one of the model's hidden layers has been provided with nonlinearity with the aid of a ReLU work [7]. channels that follow a set hierarchy. In order to use the straight change, the model additionally contained the 1*1 channels. For the purpose of maintaining the spatial aim, the 1 pixels cushioning is now complete. Spatial pooling is expressed using five convolutional layers.

Inside the top 5 approvals and test blunder, the test's victor, GoogleLeNet, used to have a 6.7 percent mistake, while VGG19 had such a 6.8 percent error.

2) Engineering

One type of artificial brain organization named CNN Fully Convolutional Organization is often employed for tasks including institution, image analysis, divisions, and other things. The channel must be moved around the image during convolution in order to become familiar with a few important details of the input picture. Since we know that the image is a lattice filled with the arithmetic value shown by I in figure 2, we may assume that it is just a lattice loaded with some theoretical value. Thus, the basic information or elements are learned at different phases with the help of these channels, designated as K in picture 2, that are common purpose or goal throughout the information picture.

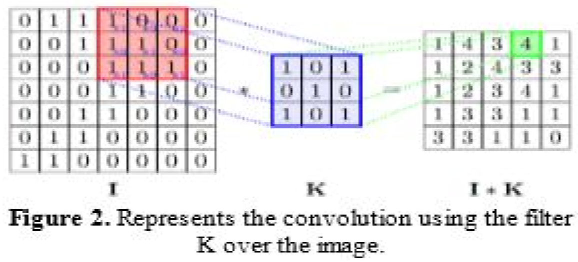


Figure 2. Addresses the convolution utilizing the channel K over the picture

Convolutional layers, pooling layers, totally associated layers, and a standardisation layer make up the CNN's secret layer. where the enactment task is taken care of by the outcome of the secret layers Enactment work adds to the organization's ability to learn something complicated and convoluted structure information and address non-straight complex erratic beneficial mappings across data sources and results, making it all the more astounding.

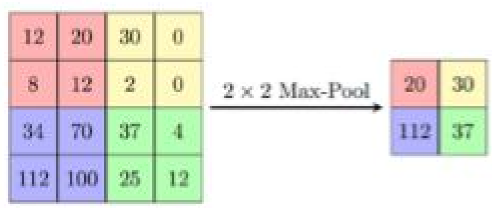


Figure 3 shows a reduction in the perspective from 4*4 to 2*2 and the greatest pooled from the 2*2 window.

Normal pooling involves taking the mean of the possible values present in that framework. Layers that are fully linked take into account the fact that each neuron in the following layer will receive loads from the one above. The standardisation layer was utilised to standardise the enactments from the preceding layer while maintaining the mean actuation close to 0

3) Framework

Convolutional Network Brain Network are already often used for a number of purposes in the medical, agricultural, and other fields. In our situation, we used a pre-trained model like VGG19 to focus on the information that was pertinent. By doing so, we can modify prior gains and keep from starting without any planning. Pretrained models that are readily available including VGG16 [7], as well as InceptionV3 [11]. The missed elements are now accepted as contributions by a number of classifications, notably SVM [8], Knn [10], K - nearest neighbors [12], as well as Logistic Relapse [9]. By incorporating the highlights into the pretrained models, we were able to extract the highlights from the images. The removed parts are now being used as contributions by a number of classifiers, including SVM [8], Neural Network [10], KNN [12], and Logistic Relapse [9].

(A) = Prob(B=1 | A; (1)

The likelihood that an event A will result in an additional event B=1 defined by factor "as mentioned in position 1 over.

Prob (B=1|A;) + Prob (B=0|A;) = 1 (2)

Prob(B=0|A;$\boldsymbol{\alpha}$)=1

-Prob(B=1|A;$\boldsymbol{\alpha}$)

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

An expense work is used to entirely convert probabilistic results to plain outcomes, as indicated in condition (3) below.

((A),B) = - B*log(A), - (1-B)*log(A),)_ (3)

When B = 1, a (1-B) term becomes 0 and just log((A)) is introduced. In this case, log(1-(A)) only will be introduced and the (B) part will be zero.

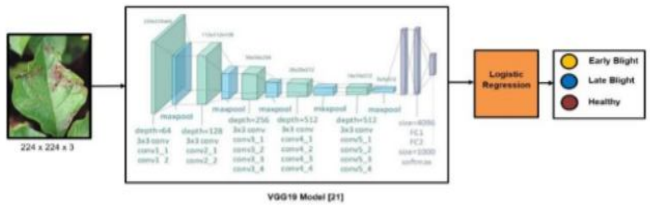


Figure 4 : Proposed Model

III. RESULTS AND DISCUSSION

This suggested model uses a plant town collection of 2152 images of potatoes leaf, 1000 images of alternaria blight, 1000 images of late scourge, and 152 sound-enabled pictures of potato leaves. The preparation component of the data set has 1700 photographs (70 percent), while the test segment contains 452 photos (30 percent). For highlight extraction, a number of previously made models, particularly inceptionV3 [11], & VGG19 [7], were used. Table 3 shows that VGG19 produced the best results. Numerous classifiers, such as KNN [12], SVM [8], machine learning [10], and computed relapse, are used for organisation. The most current approach, known as "Logistic Regression," has a 97.7 percent accuracy rate in characterising data. AUC (Area Underneath the Curve), CA (Accuracy), Accurateness, Recall, and F1Score execution limitations are derived to assess the model's efficacy, as shown in table 3.

Table 2 correlation report with different models

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

Table 3 displays the VGG16, V19, and InceptionV3-based AUC, CA (Language Model), Exact, Retrieval, and F1-Score metrics.

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

Pictures 5, 6, and 7 display the ROC charts for a number of classes, include Early Blight, Wholesome, and Late Blight.

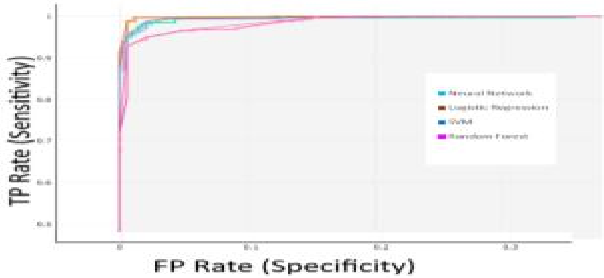


Figure 5. A ROC diagram for the early blight class using the VGG19
Figure 6 shows the Healthy class's logistic regression as well as VGG19 ROC Plot.
Figure 7. Late Blight Class ROC Plot for VGG19 Using Logistic Regression.

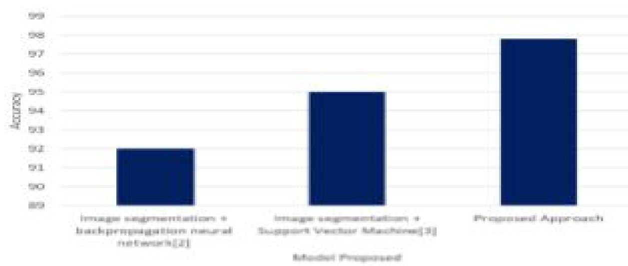


Figure 8: Correct categorisation

IV. CONCLUSION

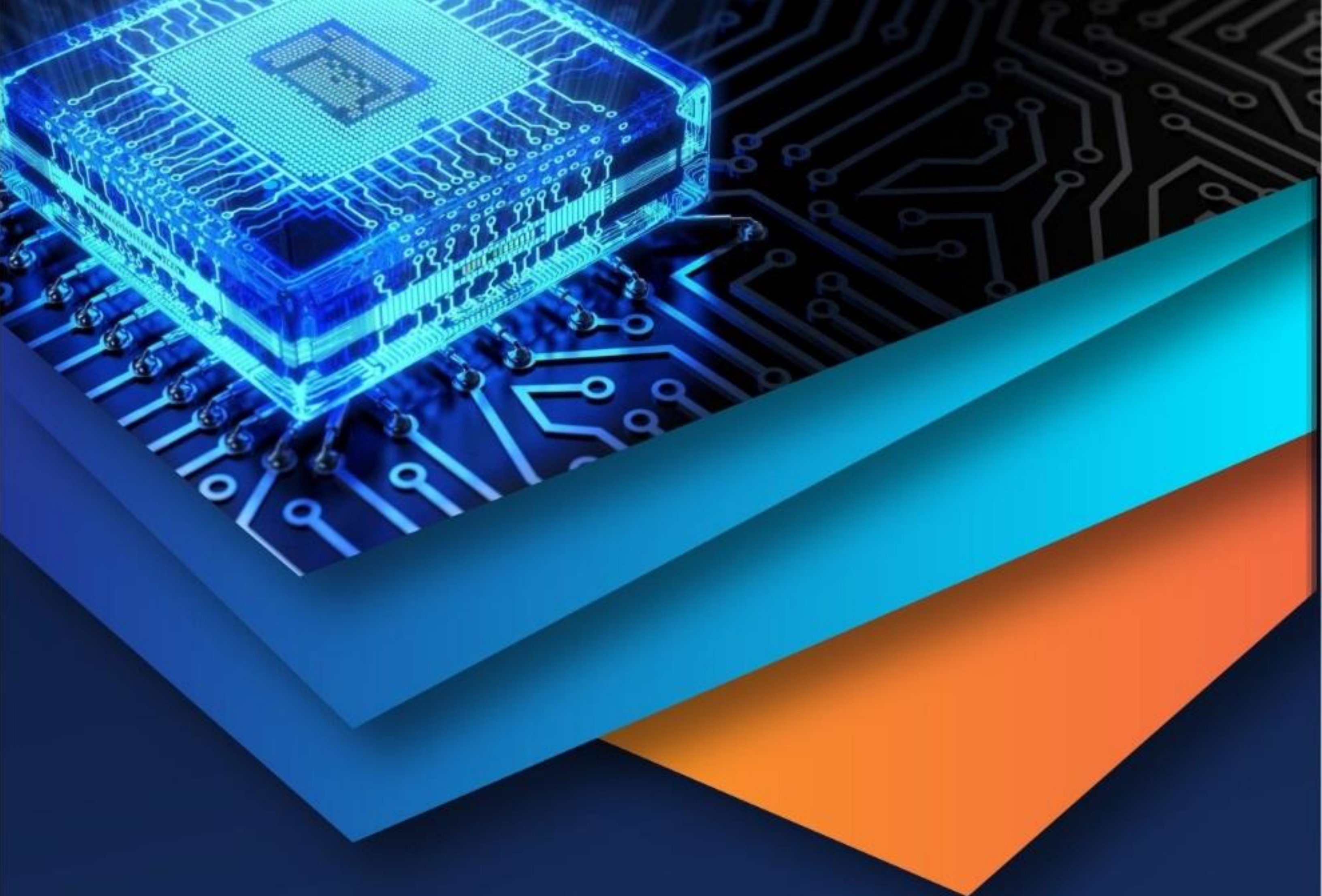
We created a mechanised framework employing the idea of motion learning to analyse and categorise illnesses in potatoes leaves like initial curse, later scourge, and also sound using merely an original arrangement. We were able to achieve a classification precision of 97.8 percent over the test dataset and improvements of 5.8 percent and 2.8 percent with [2] as well as [3] separately. Ranchers can benefit from our method's early illness detection and higher harvest yields. The arrangement precision of the suggested strategy in comparison to several executions is shown in Figure 8.

V. FUTURE SCOPE

It is now essential to detect a problem on the inside of a plant when it is in the budding stage in order to optimise the performance and accuracy of the yield. It would be very helpful if we could install this system on a cell phone, enabling ranchers to take a photo of a leaf and submit it to the server because disease discovery requires a high degree of skill. As a consequence, the server will identify the virus kind, place an order, and send the findings, along with suggested treatments, to the phone.

REFERENCES

[1] Abdallah Ali (2019, September) PlantVillage Dataset, Version 1, Retrieved 22 February 2020, https://www.kaggle.com/xabdallahali/plantvillage-dataset.
[2] Athanikar, . Girish and Ms. PritiBadar. "Potato Leaf Diseases Detection and Classification System Mr." (2016).
[3] M. Islam, Anh Dinh, K. Wahid and P. Bhowmik, "Recognition of potato infections utilizing picture division and multiclass support vector machine," 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, 2017, pp. 1-4.
[4] C. U. Kumari, S. Jeevan Prasad and G. Mounika, "Leaf Disease Detection: Feature Extraction with K-implies bunching and Classification with ANN," 2019 third International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 1095-1098.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY