iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: V Month of publication: May 2023

DOI: https://doi.org/10.22214/ijraset.2023.51611

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Crop Health Monitoring System using Deep Learning

Suraj Kumar B P[1], Ekant[2], Pradeep Mogallapu[3], Vivekananda Kalmat[4], Yashwanth Nalla[5]
[1]Assistant Professor, [2, 3, 4, 5]Student, Department of Computer Science and Engineering Sir M Visvesvaraya Institute of Technology, Bangalore, India and affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

***Abstract:** Food is the basic need of human being which can be fulfilled by agriculture. The agriculture plays an important role in the country's economy. However, yield of crop per acre is very low when compared to the other developing countries. Using fertilizer to provide additional nutrients to help and develop a healthy crop is a crucial aspect of agriculture. Lack of required nutrients and climate changes has given rise to many crop diseases. The very big problem in this context is to identify the crop diseases and spread the knowledge among farmers. The main purpose of this project is to provide operational and user-friendly system that recommends best profitable crop according to that instance of field, fertilizers to use, detects diseases that attacked crop and additionally provides pesticide suggestions to cure the disease. The model created using ML algorithms such as Logistic Regression, Decision Tree, Naive Bayes, Support Vector Machine (SVM), Random Forest allow choosing the most profitable crop and recommends the required fertilizers for the given soil. The leaves have similarities in their texture, shape, size and other visual aspects that are used to identify the disease on the given crop. The project uses Deep Learning Algorithm with Digital Image Processing to detect the disease.*
***Keywords:** Machine Learning, Deep Learning, Fertilizers, Crop disease, Logistic Regression, Decision Tree, Naive Bayes, Support Vector Machine (SVM), Random Forest, Digital Image Processing, ResNet.*

I. INTRODUCTION

As we know Agriculture is backbone of India, it ranks second worldwide in farm outputs. As per 2018, more than 50% of Indians depend on agriculture and contributed 70-80% to the country's Gross Development Product (GDP) [6]. The country produces different crops and food grains such as coffee, jute, groundnut, wheat, rice, pulses, oil seeds, fruits etc. The crop yield depends on the different factors such as climate change, geographics, crop diseases, lack of irrigation, fertilizers and financial components [7]. It is difficult for farmers to decide when and which crop to be planted due to fluctuating market demand. According to the Wikipedia figures India's farmer suicides account for 11.2% of total suicides in India. Crop disease decreases the production and the quality of the yield. Identifying the crop disease is necessary to cure the diseases. The disease attacked to crops can be identified by their leaves, stem, fruit conditions like colour, shape, size.

The objective of this paper is to provide a model that recommends the best profitable crop for a specific region. This model provides the crop selection based on the environmental conditions such as nature of soil, temperature, humidity and pH level to maximize the crop yield. As the nutrients in the soil plays an important role in the growth of the crop the fertilizers are essential to maintain the fertility of the soil. So, the model also provides the information on the fertilizers to be used based on the soil condition for the healthy and efficient crop yield. The rapid growth of different types of diseases and lack of knowledge among the people, identifying these diseases and treatment of these diseases become major challenge. The datasets for healthy and diseased crop leaves are available on the internet which can be used to take the necessary action towards the fertilization of crops to improve the efficiency of yield.

II. RELATED WORK

The dependence of soil type on crop output has been studied in article [1]. Machine learning approaches enable selecting the most productive crop or forecasting the crop yield for a crop that the user has chosen. A few Machine Learning methods, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests (RF), Multivariate Linear Regressions (MLR), and K-Nearest Neighbors (KNN), are used to forecast agricultural productivity. With 94.3% accuracy, the Random Forest model outperformed the others.

In paper [2], proposed a smart farming technique, that will use IoT and Machine Learning technologies in order to determine the nature of the soil and the type and amount of pesticides and fertilizers to be used in an efficient methodology. It uses the K-Nearest Neighbour (KNN), Naive Bayesian model and Linear Supporting Vector Classifier to predict the fertilizers and pesticides and maximize the yields.

In paper [3], a technique to identify the crop species has been discussed. Different measures applied on the pixels of the leaf image like mean, median, standard deviation (SD), skewness and kurtosis are used to prepare the model. It also describes the outcomes of second order statistical texture examination with relationship among pixels (GLCM) on the image to classify the crop.

Neural networks are employed in paper [4] to identify the condition of crop. If a leaf is contaminated, more processing is carried out to determine the illness. SVM and a genetic algorithm are used to minimise loss and determine the kind of illness. This study presents a genetic algorithm-based technique for loss function optimization that is analogous to the principle of natural selection, which holds only for robust parameters.

Image capture, picture pre-processing, image segmentation, feature extraction, and illness classification are all covered in article [5]. A routine monitoring strategy for plant disease identification becomes crucial to raising productivity and produce quality of yield. Digital image processing is a potent tool for identifying challenging symptoms much more quickly than people. Different approaches, including Support Vector Machine, K-Nearest Neighbors, and other classifier methods, have been discussed for categorising plant diseases. The efficiency of plant disease detection might be increased by using hybrid algorithms that combine genetic algorithms, ant colony, cuckoo, and particle swarm optimization with SVM, ANN, and KNN..

In paper [13], analysis of crop yield prediction has been performed using data mining technique. Association rule-based technique has been implemented in different areas of Tamil Nadu.

In paper [14], k-means clustering has been used for clustering and identification of plant pest has been performed using correspondence filtering. Experimental results show that the proposed method is a useful method, that provides accurate detection of various plant pests using shape, size, position and orientation which results in computationally efficient algorithm.

In paper [15], a review of different models such as Generalized Regression Neural Network, Back-Propagation Neural Network and Feed-Forward Neural Network has been discussed with their role in detection of leaf diseases in agriculture. Neural Network models are good for hyperspectral data processing.

In paper [16], a deep convolutional neural network is used to identify the diseases on rice crop. The architecture of CNN is made up of different layers to have a better performance. In future, authors of the paper wants to implement their model for fault detection.

In paper [17], DenseNet-121 is used for identification of apple leaf disease. An approach of multi-label classification is used for identification of diseases more efficiently. Healthy Apple, Normal Cedar Apple Rust and Serious Cedar Apple Rust, Normal Apple Scab, Serious Apple Scab, Apple Gray Spot are the different diseases that are identified.

In paper [18], modified LeNet a deep learning CNN model has been used to classify the leaf disease for maize. The convolutional, pooling and fully connected layers are designed in such a way that it improves the performance of the model.

In paper [19], a deep learning model using CNN has been implemented to identify the sugarcane disease. Pre-processing, extraction of features and classification has been performed. They want to implement different learning rates that can evaluate their method.

The illnesses Early Blight and Powdery Mildew of Tomato Leaves were found in paper [14]. They used a variety of approaches for picture pre-processing, including smoothness, noise reduction, image scaling, image isolation, and backdrop removal. Both feature extraction for feature vectors and classification used the Gabor Wavelet Transformation. SVM also used Cauchy, Laplacian, and Invmult Kernels for output judgement and training on illness detection. According to the author, this strategy might produce 99.5% accuracy.

III. METHODOLOGY

In this paper the implementation of the system that recommends the best yield crop, predicting the required fertilizers based on the soil and crop data and detecting the disease attacked to the crop is done in different phases such as: obtaining the relevant dataset, pre-processing the data, training the ML model for predicting the best profitable crop, use of fertilizers and detection of disease.

A. Acquiring the Datasets

The most important component of any ML-based application is gathering the dataset. Acquiring a dataset with sufficient data on which to create a precise anticipated model is crucial. The rainfall, climate, and fertiliser datasets are combined to provide the dataset for crop recommendations and fertiliser suggestions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

The ability to view photographs of both healthy and ill crop leaves is made possible by a number of open-source datasets. There are 54303 photos of healthy and sick crop leaves in the Plant-Village collection, which has been divided into 38 groups by species and illnesses [8]. Every dataset has to be classified into two sections, Train dataset and Test dataset.

1) *Train dataset:* This covers the major part of the dataset that is used to train the learning ML model.
2) *Validate dataset*: The subset of data used to assess a model's fit to a training dataset while adjusting model hyperparameters.
3) *Test dataset*: This part of the dataset is to check the unbiased evaluation of accuracy in the model.

B. Pre-Processing the Data

Data pre-processing is the process of modifying the raw data to make it appropriate for the necessary ML application. Finding the data in the proper format is challenging since it often includes missing numbers, noise, outliers, and inconsistent data. These cannot be used to train ML models directly. Pre-processing is needed to organise the data into a structured manner and clean up the data by eliminating noise and missing information. It improves the model's precision, dependability, consistency, and effectiveness.

Image processing: This phase of processing the data provided to the ML model is crucial. By concentrating on the subject of interest, it shortens the model's reaction time. The photos in the collection must be treated to address issues with lighting, background, and noise. To extract the necessary portion of the image, segmentation is applied to the original image. Thus, segmentation produces photos with a black backdrop, and processing images in grayscale prevents uneven lighting situations.

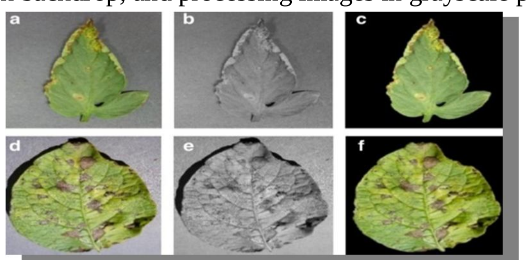


Fig. 1 (a) Original image of Leaf 1, (b) Greyscale image of Leaf 1, (c) Segmented image of Leaf 1, (d) Original image of Leaf 2, (e) Greyscale image of Leaf 2, (f) Segmented image of Leaf 2 [10]

C. Crop Recommendation and Fertilizer Suggestion

According to the data for the specified location, the model forecasts the most lucrative crop. By increasing yields and effectively using resources, combining agriculture with machine learning will benefit the agriculture industry. The prior data is quite valuable for assessing present performance and improving yield output. Aspects of the data include state, city, season, amount of rainfall, type of crop, etc. Other datasets include features for soil data such as pH value, nitrogen, phosphorus, and potassium. The primary data set is created by extracting and combining these soil attribute columns. The data sets have already been processed. Prior to running the algorithms, the labels are produced using the category characteristics.

The different supervised ML algorithms such as Logistic Regression, Support Vector Machine (SVM), Random Forest, Naive Bayes and Decision Tree are used in this model to predict the best suitable crop and to suggest the fertilizers to get the best yield possible. The Random Forest and Naive Bayes algorithm gives the best accuracy in this case.

1) *Random Forest*: It is a very famous supervised machine learning algorithm used for both classification and regression applications. It is an ensemble learning concept that integrates different classifiers to solve a given complex problem and improve the efficiency of the model.

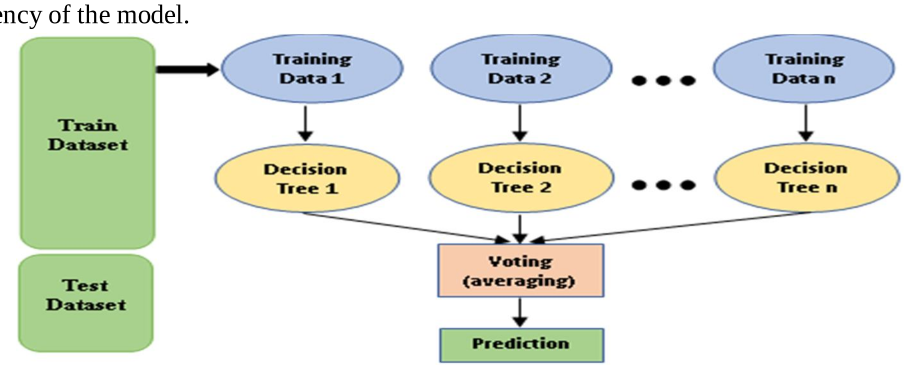


Fig. 2 Working of Random Forest Algorithm [11]

2) *Support Vector Machine*: This is also a famous supervised machine learning algorithm that is primarily used for classification problems. The algorithm creates the best line or decision boundary known as hyperplane that can segregate n-dimensional spaces into classes so that we can easily include the new entry to the correct class.

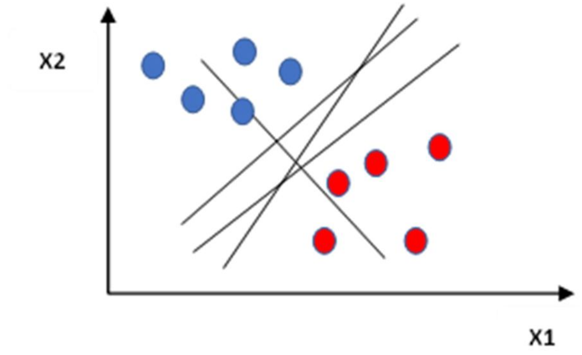


Fig. 3 Linearly separable data points [12]

3) *Naive Bayes:* This is a supervised learning algorithm based on Bayes theorem that is used for solving classification problems. It is simple and most effective algorithm that helps in building machine learning models quicker that can make faster predictions. It is useful especially for very large datasets.

The Bayes theorem formula to calculate the probability $P(A|B)$,

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Where, $P(A|B)$ is the posterior probability, $P(B|A)$ is the likelihood probability, P(A) is the prior probability and P(B) is marginal probability.

D. Disease Detection and Pesticide Suggestion

The main motive of this model is to classify between the healthy and disease attacked crop leaves. Additionally, we have to find the disease attacked to the crop and based on it the suggestion has to be given as output. There are various pre-trained models like Xception, VGG, ResNet, Inception and others that could be used to implement our deep learning model. ResNet is used for implementation of this model.

In Residual Networks, we use a technique called skip connections that connects activations of some layer to upcoming layers by skipping some intermediate layers to avoid overfitting. That forms a residual block when stacked together constitutes ResNets. It also helps in preventing vanishing/exploding gradient problem that makes the gradient to become zero or too high.

$$F(x) := H(x) - x$$
$$H(x) := F(x) + x$$

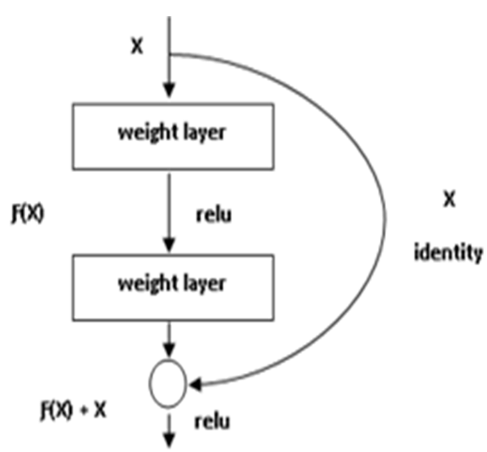


Fig. 4 Residual learning: a building block [9]

The benefit of including this type of skip connections is that if any layer that will impact the performance of the architecture will be skipped. This yield in training the very deep neural network without any problem of vanishing/exploding gradient.

E. Working

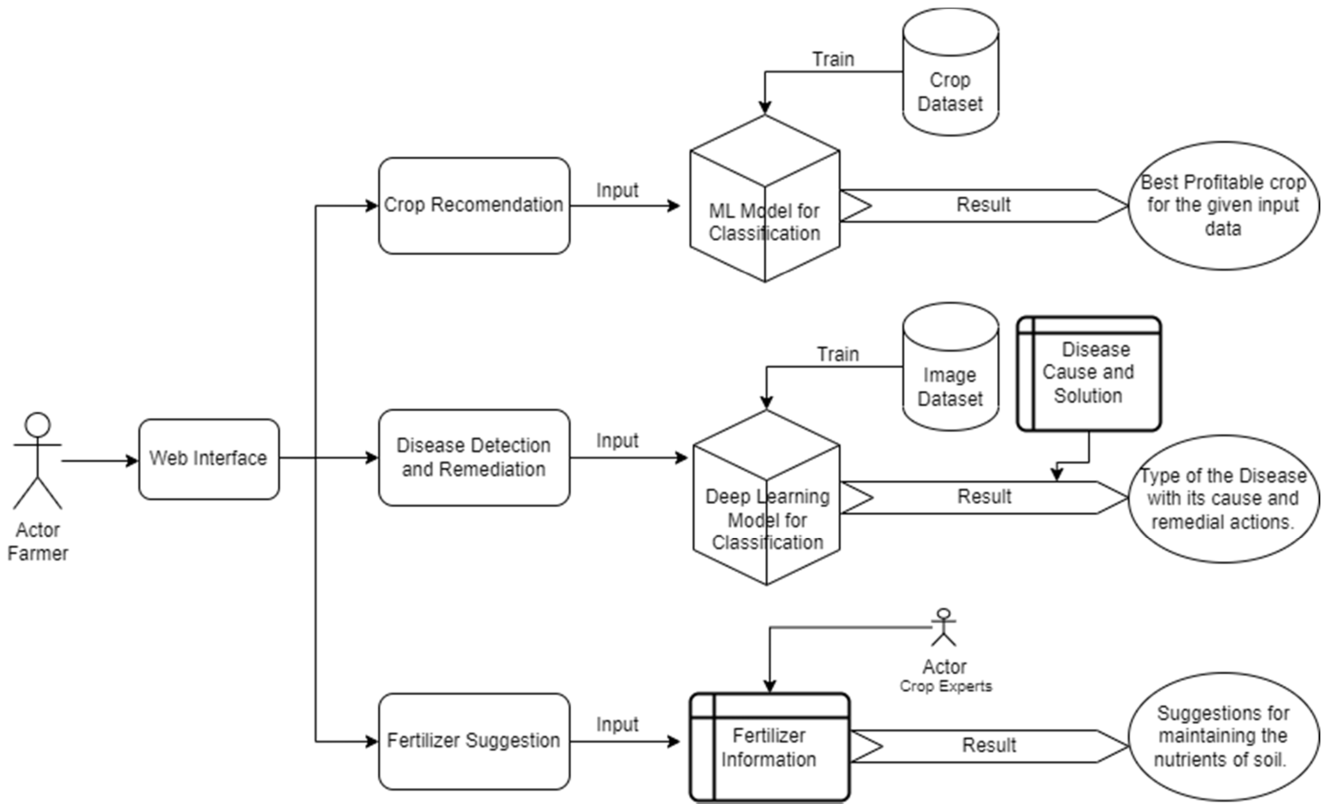


Fig. 3 Work Flow Diagram of the proposed Project

The main actors of this project are the farmers, crop experts and Machine Learning engineers. When farmers enter environmental information like n, p, k, pH, humidity, and temperature, they obtain recommendations for crops and fertiliser.

In order to identify sickness, they also submit photos of leaves for disease detection. Based on the environmental parameters and the area where the crops are cultivated, crop specialists provide their knowledge in choosing the suitable crops and fertilisers.

Our trained models are given with the input to perform the classification operation to provide the appropriate result for crop recommendation and disease detection. The data collected by the crop experts is used to provide the fertilizer information to maintain the soil nutrients and to give the suggestion for the disease attacked to the crop.

IV. IMPLEMENTATION

The tech stack used in the project contains python libraries like NumPy, Pandas and MatPlotLib for dataset manipulation and visualization. PyTorch, Scikit-learn, Pickle are used for model implementation and serialization. Flask framework is used as backend of the project. The HTML templates with CSS stylings are used to respond to the requests.

Accuracy for different Machine Learning classification algorithms used for crop recommendation.

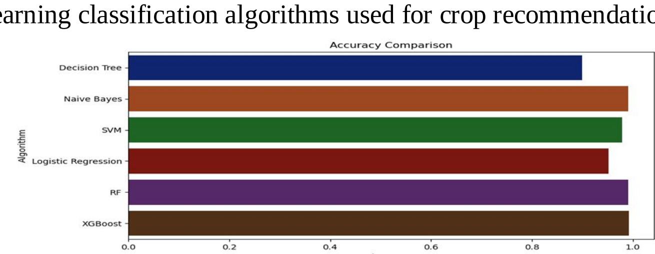


Fig.4 Accuracy comparison of different ML Algorithms

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Code for implementing the Random Forest Algorithm.

```
from sklearn.ensemble import RandomForestClassifier

RF = RandomForestClassifier(n_estimators=20, random_state=0)
RF.fit(Xtrain,Ytrain)

predicted_values = RF.predict(Xtest)

x = metrics.accuracy_score(Ytest, predicted_values)
acc.append(x)
model.append('RF')
print("RF's Accuracy is: ", x)

print(classification_report(Ytest,predicted_values))
```

```
RF's Accuracy is:  0.990909090909091
```

Fig. 5 Jupyter Cell

Code for implementing the ResNet18 Deep Learning CNN model.

```
import torch
import torch.nn as nn
import torch.nn.functional as F

class ResNet(nn.Module):
    def __init__(self, num_classes=39):
        super().__init__()
        self.resnet = nn.Sequential(
            nn.Conv2d(3, 64, kernel_size=7, stride=2, padding=3, bias=False),
            nn.BatchNorm2d(64),
            nn.ReLU(inplace=True),
            nn.MaxPool2d(kernel_size=3, stride=2, padding=1),
            nn.Sequential(*list(torch.hub.load('pytorch/vision:v0.9.0', 'resnet18', pretrained=True).children())[4:-1]),
            nn.AdaptiveAvgPool2d((1, 1)),
            nn.Flatten(),
            nn.Linear(512, num_classes)
        )

    def forward(self, x):
        return self.resnet(x)
```

Fig. 6 Python script for ResNet

V. EXPERIMENTAL OUTPUT

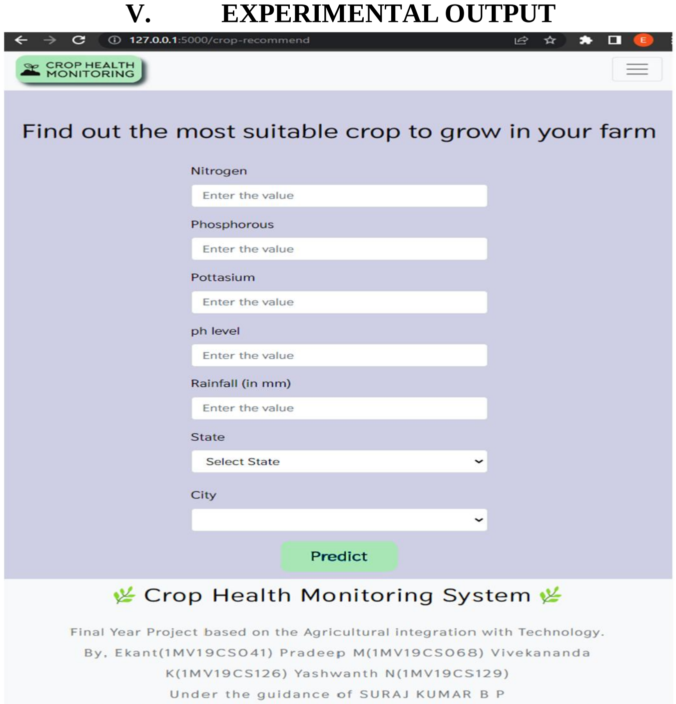


Fig. 7 Web interface to accept input from farmers

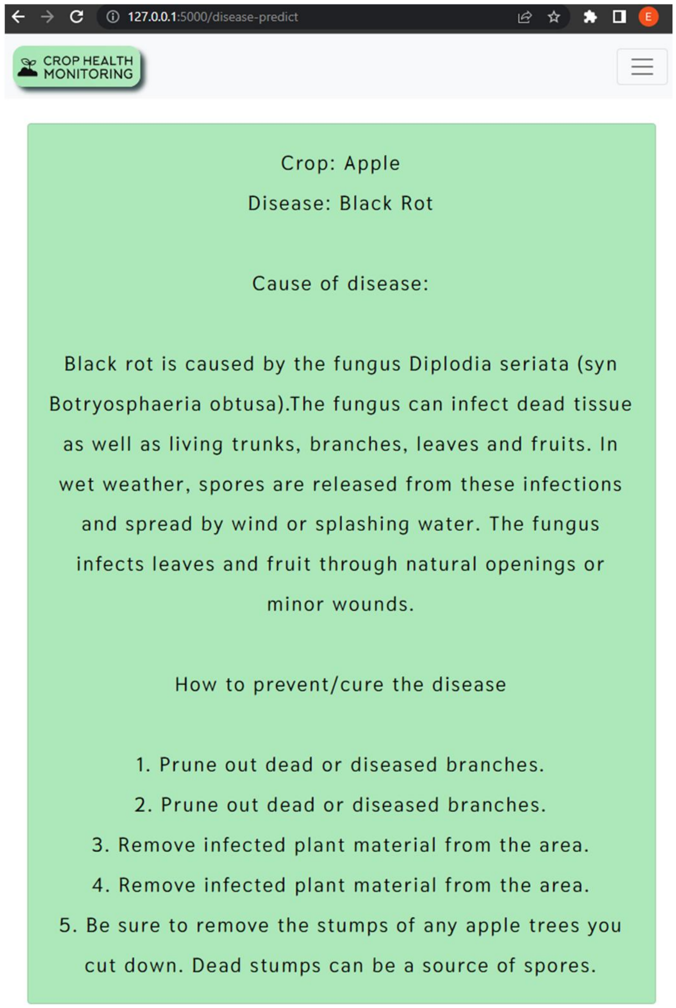


Fig. 8 Output of disease detection feature

VI. CONCLUSION AND FUTURE WORK

Furthermore, for effective crop cultivation and loss reduction, it is crucial to have precise disease detection and categorization of plant leaf images. It can be observed from the review that several categorization techniques were employed to forecast the most valuable crop, suggest fertilizer, and identify and detect agricultural diseases. This project acts as an assistive hub for the farmers to get the best profitable crop to grow, suggestions to maintain the soil nutrients required for the crop and detection of disease on the crop by getting the leaf as input and providing the cause of the disease with the solution to get rid of it.

The efficiency would rise in the future as hybrid algorithms developed. Farmers may utilise web and mobile applications that will be developed with corrective remedies to find illness in any type of flower, stem, fruit, or leaf including nutrient inadequacy. Without requiring any human intervention, the procedure may be automated utilising real-time data with the help of Internet of Things. In comparison to machine learning, deep learning is the most potent technology that offers promising characteristics for the accurate identification and categorization of plant diseases.

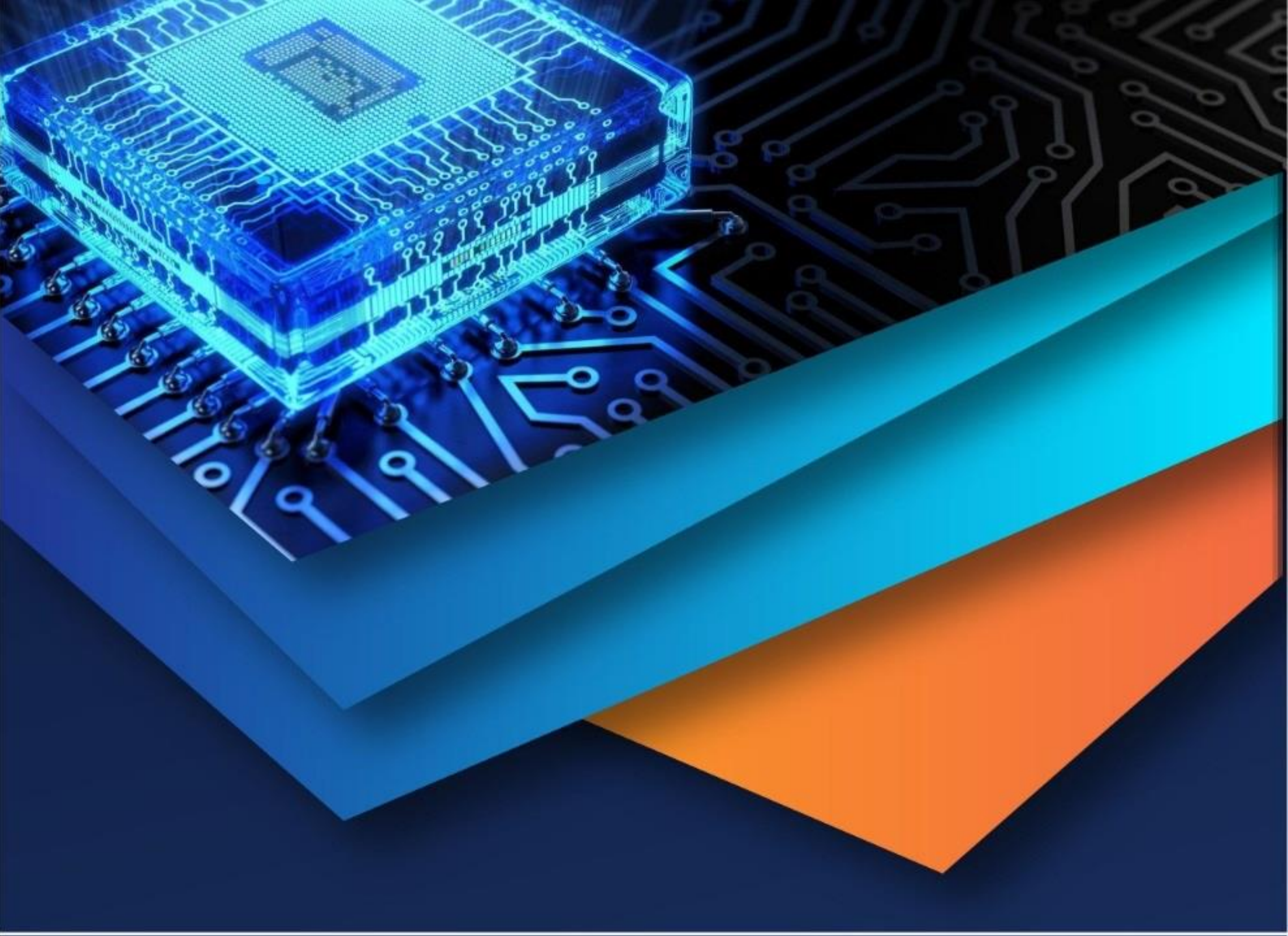
REFERENCES

[1] S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and K. SHAURYA, "Crop Recommender System Using Machine Learning Approach," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351.

[2] L. Kanuru, A. K. Tyagi, A. S. U, T. F. Fernandez, N. Sreenath and S. Mishra, "Prediction of Pesticides and Fertilizers using Machine Learning and Internet of Things," 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9402536.

[3] Kadir, A., "A Model of Plant Identification System Using GLCM, Lacunarity and Shen Features," Research Journal of Pharmaceutical, Biological, and Chemical Sciences Vol.5(2) 2014.

[4] Naik, M.R., Sivappagari, C., "Plant Leaf and Disease Detection by Using HSV Features and SVM," IJESC, Volume 6 Issue No.12, 2016.
[5] Shrivastava, Gyanesh. (2021). "Review on Emerging Trends in Detection of Plant Diseases using Image Processing with Machine Learning". International Journal of Computer Applications. 174. 10.5120/ijca2021920990.
[6] Wikipedia url:https://www.wikipedia.org/wiki/agriculture_in_india
[7] Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HC, Herrero M, Howitt RE, Janssen S, Keating BA, "Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural systems", 2017 Jul 1;155:269-88.
[8] Plant_Village dataset url:https://tensorflow.org/datasets/catalog/plant_village
[9] He, Kaiming et al. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 770-778.
[10] https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full
[11] https://www.javatpoint.com/machine-learning-random-forest-algorithm
[12] https://www.geeksforgeeks.org/support-vector-machine-algorithm/
[13] Manjula E, Djodiltachoumy S, "A model for prediction of crop yield" International Journal of Computational Intelligence and Informatics, 2017 Mar;6(4):2349-6363.
[14] F. Fina, P. Birch, R. Young, J. Obu, B. Faithpraise and C. Chatwin, "Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters", Int. J. Adv. Biotechnol. Res., vol. 4, no. 2, pp. 189-199, Jul. 2013.
[15] Golhani, K., Balasundram, S.K., Vadamalai, G., Pradhan, B.: A review of neural networks in plant disease detection using hyperspectral data. Information Processing in Agriculture, 5(3), 354–371. (2018)
[16] Lu, Y., Yi, S., Zeng, N., Liu, Y., Zhang, Y, "Identification of rice diseases using deep convolutional neural networks". Neurocomputing, 267, 378–384. (2017).
[17] Zhong, Y., Zhao, M, "Research on deep learning in apple leaf disease recognition", Elsevier, p1-6. (2020)
[18] Ahila Priyadharshini, R., Arivazhagan, S., Arun, M., Mirnalini, A.: Maize leaf disease classification using deep convolutional neural networks. Neural Computing and Applications. P1-9. (2019)
[19] Militante, S.V., Gerardo, B.D., Medina, R.P, "Sugarcane Disease Recognition using Deep Learning". 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). P1-4. (2019)
[20] U. Mokhtar, A.S. Mona, Alit, A. E. Hassenian and H. Hefny, "Tomato leaves diseases detection approach based on support vector machines", IEEE pp. 978-1-5090-0275-7/15, 2015.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY