iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VI Month of publication: June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35375

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Plant Diseases Detection and Pesticide Suggestion Implementation

Onkar Kunjir[1], Pramod Kadam[2], Prathamesh Sahane[3], Saif Momin[4]
[1, 2, 3, 4]Computer Engineering Department, Sinhgad Academy Of Engineering

Abstract: Plant diseases affect the life of not only farmers but also businesses which are dependent on it. Plant disease detection is a computer vision problem which tries to identify the disease splat is infected using an image of a plant leaf. Different kinds of models have been proposed to tackle this problem. This paper focuses on generating small, lightweight and accurate models with the help of deep learning and transfer learning.
Keywords: Transfer learning, Neural network, Convolutional Neural Network, MobileNet, Deep Learning, Computer Vision

I. INTRODUCTION

For any living organism to survive it requires three most essential things that is food, water and shelter. Out of these requirements, the necessity of food is fulfilled with the help of agriculture. Agriculture also affects day to day life of every human directly or indirectly and it's products are used in many fields as raw material or as base ingredients. Agriculture helped humanity to break though the primitive world and helped to colonize and become more civilized beings. A farmer has to invest a lot of time, money into farming before they can start getting any results. Even after making a large amount of investment farming is very highly dependent on natural factors such as rain, quality of soil, humidity etc. If the conditions are not suitable for plants to grow, all the time and money invested by farmers is pointless. Farming is a field which has lots of problems which requires supervision of an expert and experienced person to tackle which makes it even more difficult for new people to get into farming.

One of such problems is to tackle plant diseases that, if ignored and not treated in time, could destroy whole farms in no time. It takes an experienced person to identify a plant disease before it's spread to other plants. In countries whose economy highly depends upon farming, the majority of the population is engaged in farming activity. Poor disease identification and management can have a damaging effect on the economy.

On the individual scale these effects can be catastrophic because farming is their primary occupation. Plant diseases can be easily stopped if acted quickly but the method of detection is extremely tedious, time consuming and takes a talented person. Many people can't afford professionals just to spot the disease. Due to current technical advancement, simple computation and high connectivity this problem can often be automated and solved using advanced machine learning and deep learning algorithms to avoid any damage as early as possible.

Due to the ease in availability of the internet and computation problems like this can easily be automated and implemented in remote areas with the help of smartphones. Machine learning algorithms like decision trees, artificial neural networks, SVM are proving to be as effective as humans in finding patterns. This can be used to implement a system which can be cost effective, accurate and fast for detecting plant diseases. In the following sections of this document we are going to briefly discuss a few solutions which have been proposed to solve this particular problem of disease detection with high accuracy. This paper describes implementation of a system which can be used to detect plant diseases with high accuracy in real time.

II. MOTIVATION

Even though nowadays many problems are being solved with the help of automation, the agriculture sector is ignored quite often although solving these problems can impact the lives of people directly. Plant diseases seriously damage whole farms within a time of a few days. This problem can be easily solved with the help of computer vision and can be easily automated with help of hardware acceleration to generate results in real time.

Detecting diseases early in the stage not only saves many more plants from being affected by it but also reduces the cost required for pesticides. This could maximize the profit gained by the farmers by improving product quality and reducing cost for maintenance at the same time.

©IJRASET: All Rights are Reserved

III.LITERATURE SURVEY

A. Tomato Plant Disease Classification in Digital Image using Classification Tree

In this paper the author proposed a system to detect and classify diseased tomato leaves. The proposed system makes use of the choice tree algorithm as it's core component to classify the diseased leaves. In the process of training this technique, a dataset of 383 colored images of healthy and infected tomato leaves is employed where each image belongsto at least one of the 6 categories based on the type of disease. Otsu's method is used to segment images using thresholding. Otsu's method is an image processing technique which is used to perform automatic thresholding and divind an image into 2 classes foreground and background. Then each image is randomly sampled 10 times to increase dataset size. Author Tried to make use of different features extracted from the color, shape and texture of the leaf in the image. From pixel value of image author extracted 9 different features

1) Maximum RGB value among all the pixels.
2) Mean of all the pixel values.
3) Standard deviation of all pixel values.

Author also took features such as area of leaf, euler number, orientation. extent , perimeter, convex area, filled area, eccentricity, length of major axis, equi diameter and length of main axis into consideration, These features are extracted with help of image processing done on output images obtained by applying Otsu's method. For extracting the texture of the imageautor made use of correlation of pixels, contrast, energy and homogeneity for grey scaled images. Above features are used to train decision tree algorithms. Using this process autor was able to predict plant diseases with accuracy of 97.3%. Even though the system is able to predict with high accuracy this system is not suitable for a wide variety of plants. Systems accuracy can be increased with the help of a random forest algorithm which reduces the effect of outliers on the output even more.

B. Leaf Diseases Detection and Fertilizer Suggestion

In this paper a more generalized system is proposed by the author which is able to detect plant diseases of a wide variety of plants accurately. The author also tries to suggest fertilizers based on the plant species and the disease with which it is infected with the help of a database which stores the list of diseases and their fertilizers. This paper can be considered as an improvement over previous paper as it uses a random forest classification algorithm to predict the type of disease plant is suffering from. To train the model, the author splitted image dataset consisting of healthy and infected plant leaf images into a training and testing set randomly which helps to generate a well balanced training and testing set. After Splitting images into training and testing sets, the contrast of each image is enhanced in order to separate the background and foreground. These enhanced images are segmented using a clustering algorithm based upon their pixels which are used to determine the percentage of affected areas. For training purpose instead of using raw image as input for the random forest algorithm following features are handpicked by the authors like mean, standard deviation, skewness, RMS value, variance, smoothness, kurtosis, IDM, contrast, correlation, energy and homogeneity of the pixels of the input image. These 13 features are used in the training process of a random forest algorithm which is an ensemble learning method which uses many randomly generated decision trees to predict the result.This approach reduces the probability of overfitting and the random forest algorithm is not easily affected by outliers compared to the decision tree. Author of this paper goes even one step further to suggest fertilizers based on diseases detected. Pair of diseases and their fertilizers are stored in a database which is used to implement this system. Even Though the results obtained by this method are more reliable and accurate, the strategy of extracting features by hand is inefficient and sometimes humans are not able to recognize patterns as good as algorithms such as neural networks.

C. Plant Diseases Detection and Classification using CNN and LVQ Algorithm

In this paper a more sophisticated and complex algorithm is used to solve the problem of detecting and classifying plant diseases. Algorithm which detects the plant diseases can be divided into 2 units

1) Convolutional Neural Network (CNN)
2) Learning Vector Quantization unit (LVQ)

CNN or Convolutional Neural Network is a neural network algorithm which is specialized to work with image data. In this system output of CNN is given as input to LVQ. Role of CNN in this system is to learn hidden patterns and detect features hidden in the input image and convert them into simpler internal representation with only few data points representing all the information.

LVQ or Learning Vector Quantization is an algorithm that combines competitive learning which is a type of unsupervised learning and supervised learning. LVQ is a 3 layered neural network where the first layer acts as an input layer and last layer acts as output layer and the remaining hidden layer is known as kohonen layer which is partially connected to the output layer.

©IJRASET: All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

IV. DATASET FOR DISEASE CLASSIFICATION

For training our system we used a dataset of plant diseases hosted on https://kaggle.com. Dataset is available for free to use and consists of 70295 training images and 17572 validation images belonging to 38 different classes which includes images of diseased and healthy plant leaf of following plants

A. Apple
B. Blueberry
C. Cherry
D. Corn
E. Grape
F. Orange
G. Peach
H. Pepper
I. Potato
J. Raspberry
K. Soybean
L. Tomato

Fig. 1 shows how the images are distributed in the dataset which affects the performance of the trained model heavily. the difference in number of images for various classes is not very large which removes the need to apply techniques which are required for dataset balancing.

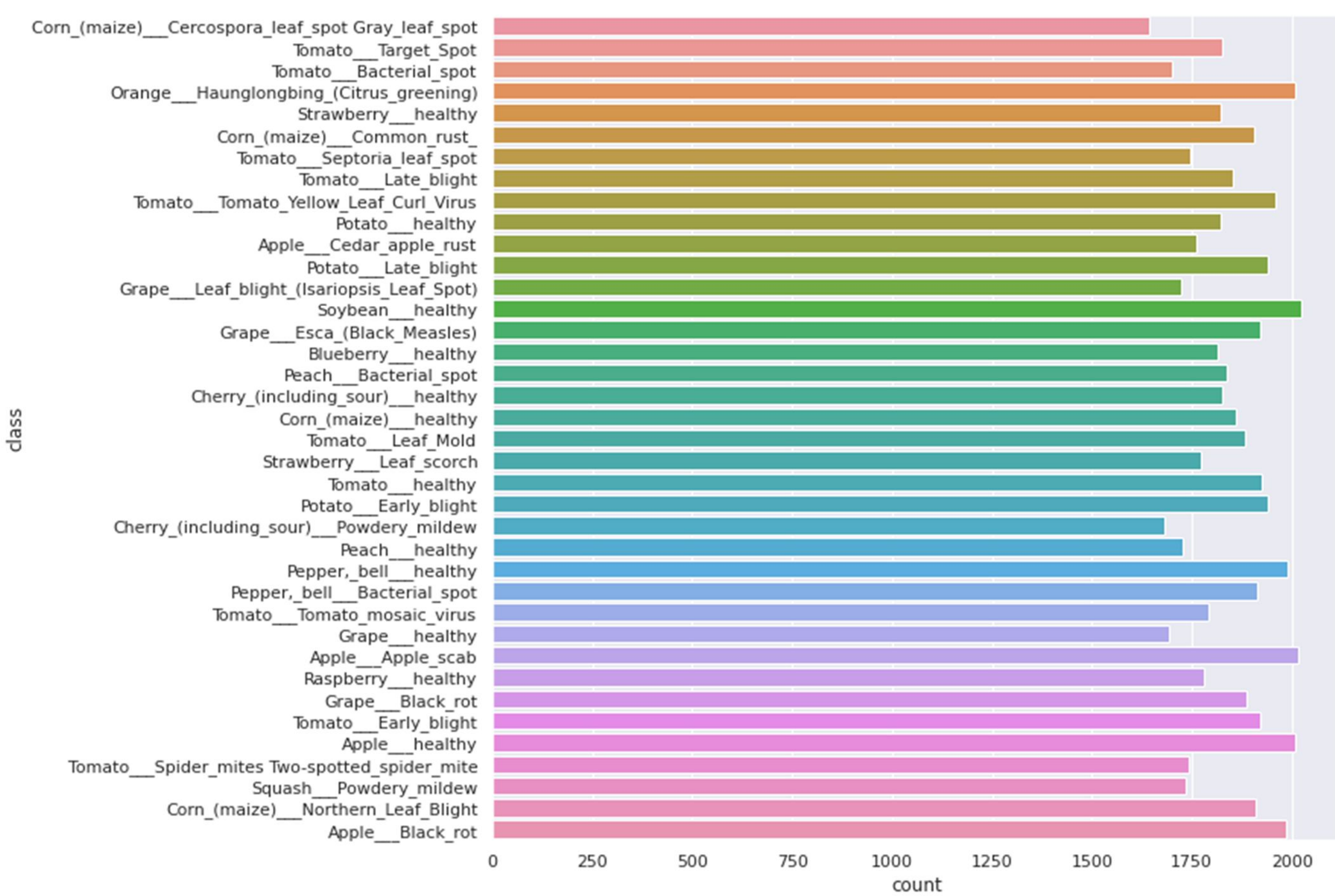


Fig. 1 Distribution of images in dataset.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Fig. 2 shows sample images of plant leaves which are present in the dataset and are used in the training process of the model.

Fig. 2 Images from dataset.

To use images from this dataset for training of model the images are preprocessed. Each image is rescaled to the size 224*224 which reduces overall data to be processed while having little to no performance loss, after this images are normalized by changing pixel value range from 0-255 to 0-1 this helps to improve learning speed of neural network model as neural network tend to perform better when all the inputs are normalized and are having same range.

V. CLASSIFICATION MODEL

To solve the problem of plant diseases detection and classification we are making use of transfer learning. Transfer learning technique is used for neural network algorithms to make one neural network learn a task which it was not originally trained for, it reduces training time significantly and also helps to attain high accuracy. In transfer learning the output layer of base models is replaced with a new output layer along with one or more hidden layers. All the layers except these newly added layers are freezed and weights of these layers are not updated in the training process. Base model increases the accuracy from the beginning of the training process reducing the training time required.

For this problem we are using the MobileNet model as a base model. The MobileNet model was originally built to be used on low powered devices which have limited computing resources such as smartphones and IOT devices which is perfect for our requirement as we want to make the system less resource heavy so it can be used for automation using IOT.

For transfer learning we are using imagenet pretrained MobileNet weights and replacing it's output layer with GlobalAveragePooling layer, dropout layer and one output layer consisting of 38 nodes each representing one class from the dataset. The GlobalAveragePooling layer is generally used to reduce the number of output nodes coming from convolutional neural networks by averaging the output values in one dimension; this helps to reduce the number of parameters required in the training process. Dropout layer is used to drop output of any neron in a particular layer which reduces the chances of overfitting the dataset and improves reliability of the model. In this system we are using a dropout layer which has a 20% chance of dropping any neurons output from the previous layer.

©IJRASET: All Rights Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Fig. 3 shows an overview of the classification model used in the system.

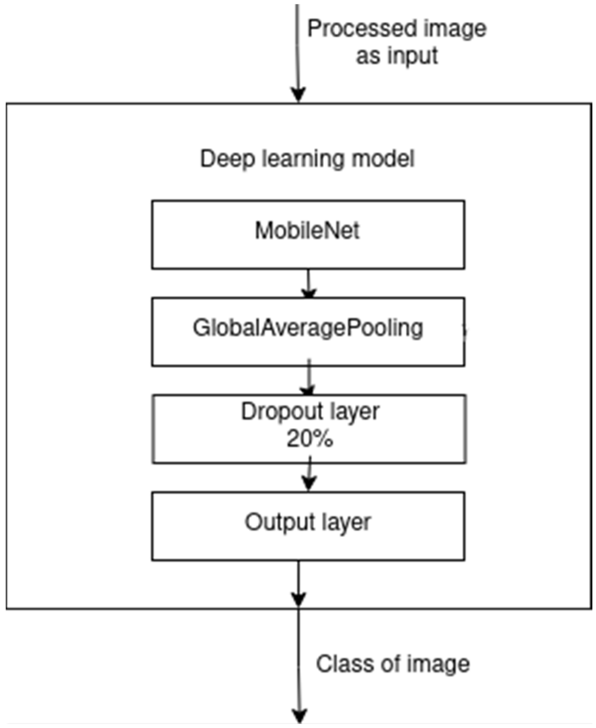


Fig. 3 Classification model overview.

Fig. 4 and Fig. 5 shows the training time accuracy and the loss function of the model. In these figures the line in blue color represents accuracy and loss recorded on training dataset while line in red color represents accuracy and loss for images in validation dataset.

The classification model was trained for 10 epochs after which we stopped seeing significant changes in accuracy and loss on both training and validation dataset.

After the final epoch model was trained with accuracy of 97.3% on training dataset and accuracy of 97.15% on validation dataset which confirms that model is not overfitted on training dataset.

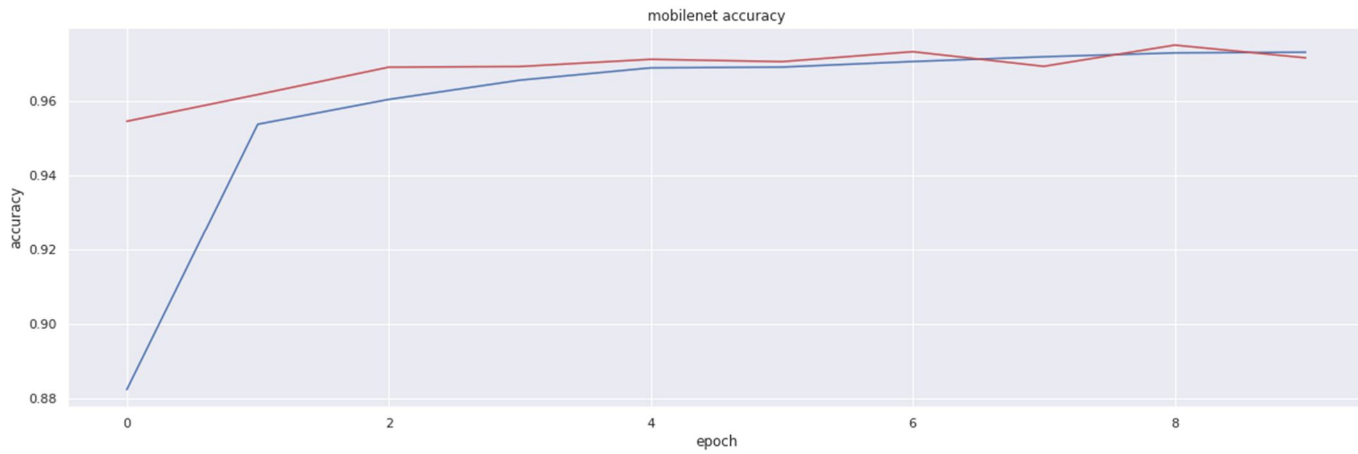


Fig.4 Accuracy of model

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

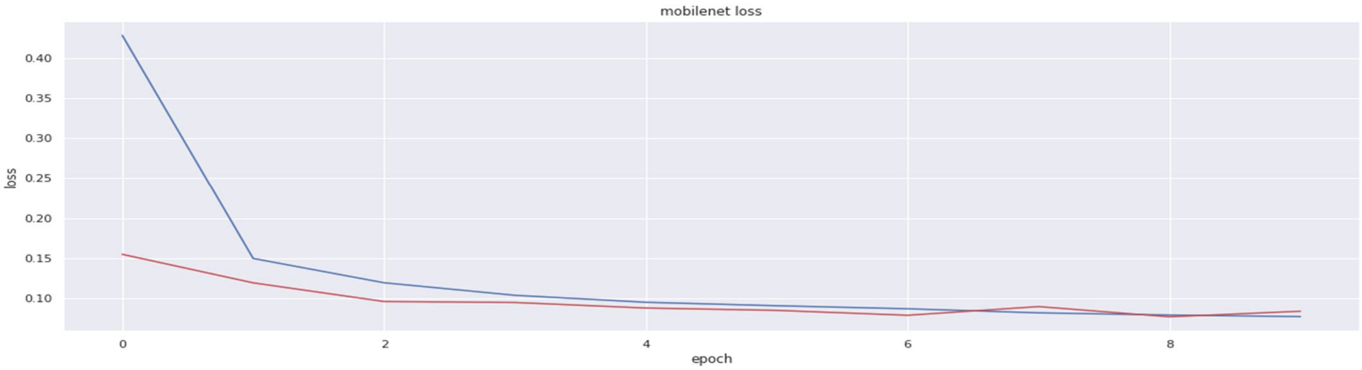


Fig.5 Loss of model..

VI. SYSTEM ARCHITECTURE

While designing system our main focus was on simplicity, ease of use and accessibility that is use should be able to access system without any platform dependencies, hence the system is designed as an web application which requires little to no setup on client side to be able to use the application and being an web application makes it platform independent without any extra efforts.

The system is heavily inspired by microservice architecture. In microservice architecture instead of building one big system it is divided into a number of individual components which are loosely coupled which makes the system more robust, flexible and scalable. our system is divided into 3 main components

A. Frontend application.
B. Backend server for performing classification.
C. Database server.

Each application is hosted on a different server which allows it to be more maintainable as if one component stopped working other components can work without being affected.

For communicating between various servers we are making use of REST api which is a highly scalable api and works on HTTP and HTTPS networking protocols.

Python Flask framework is used for building api for the system. Flask is a web framework used in the Python programming language for building REST api. Flask is very flexible and allows you to create api in minimal lines of code.

For the database we are making use of MongoDB which is a NoSQL database. MongoDB stores data in the form of JSON documents which makes it easy to use with REST api and data obtained from the database requires little to no preprocessing before sending data. Information related to diseases is stored in the database as it is expected that some information might be missing hence making use of NoSQL databases is more suitable than using an SQL database.

The database is also being used to store user information which is authenticated every time any user makes a request, this user information can be used to mine more information about plant diseases in future.

Fig. 6 shows overall system architecture and how each component communicates with each other to generate results.

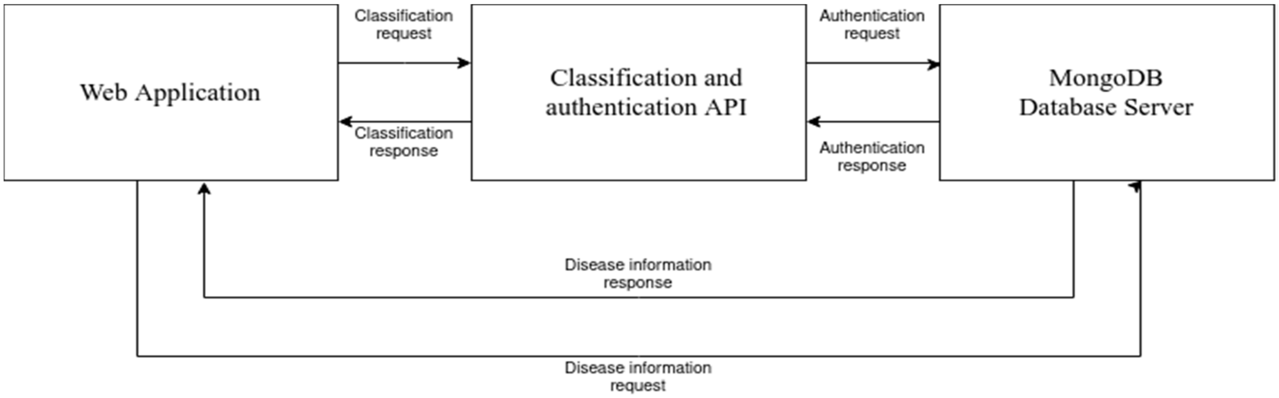


Fig.6 System architecture.

©IJRASET: All Rights are Reserved

VII. FUTURE SCOPE

The system we built and which were explained in the literature survey have one thing in common: they support only a fixed number of plant types only and to be able to classify a new diseases model needs to be re-trained which is a time consuming process. This could be tackled with the help of the Siamese network which is trained indirectly for classification tasks which removes the necessity of retraining every time new diseases need to be introduced.

Also the data collected from user queries can be mined and useful information like which location having what kinds of diseases common can be obtained which could further help to tackle this problem even before it happens.

VIII. CONCLUSION

Aim for this paper was to implement a lightweight, cross platform system which could be used to detect plant diseases in real time, and according to our finding system which we have built meets all the requirements. System is able to produce accurate results in real time even without hardware acceleration from the GPU. This system can be very beneficial for farmers as well as people who are getting started in gardening without any prior knowledge.

REFERENCES

[1] Sabrol, K. Satish, Tomato Plant Disease Classification in DigitalImages using Classification Tree, International Conference on Commu-nication and Signal Processing, April 6-8, 2016, India

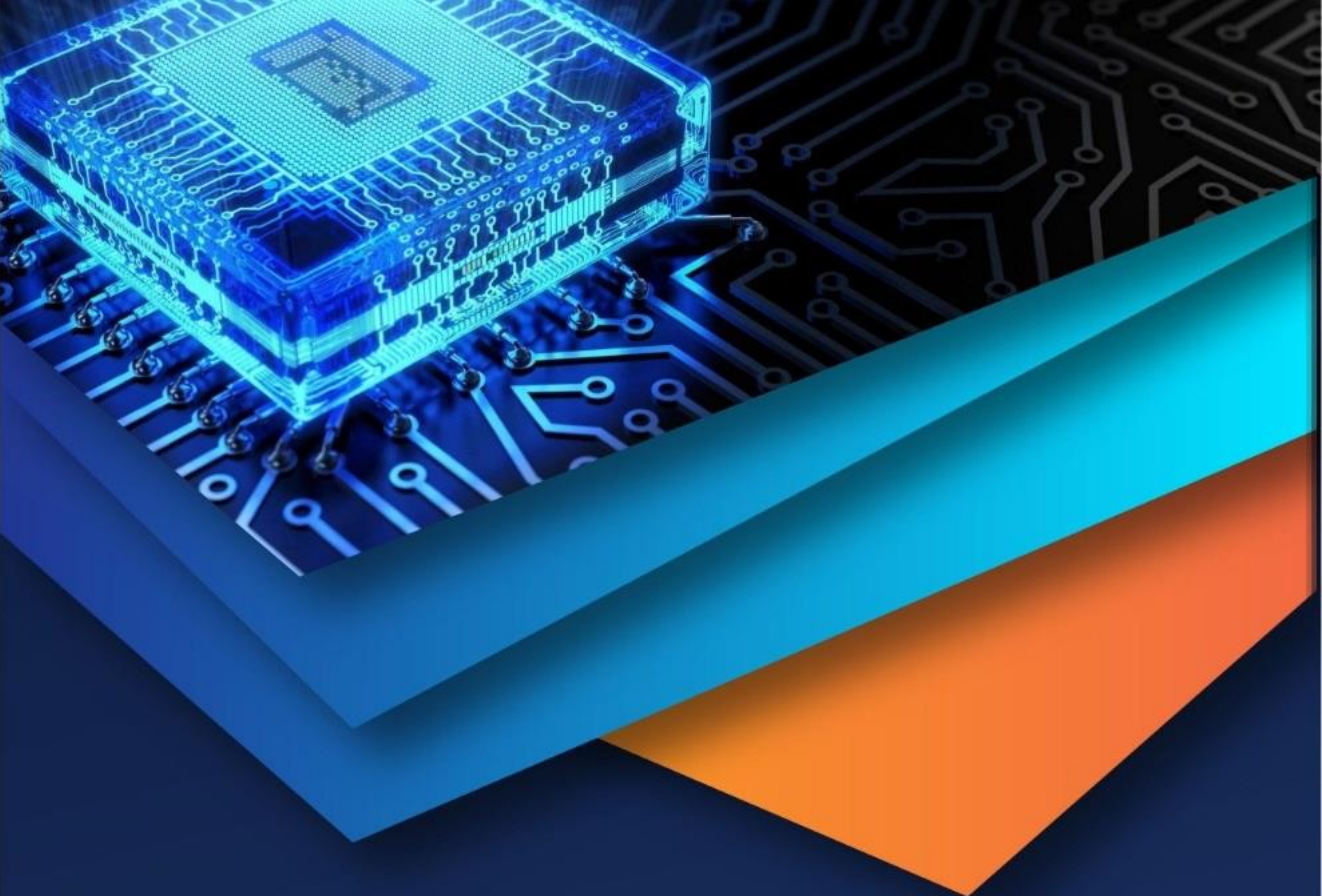
[2] Indumathi.R, Saagari.N, Thejuswini.V, Swarnareka.R, LEAF DISEASE DETECTION AND FERTILIZER SUGGESTION, Proceeding of international conference on systems computation automation and networking2019

[3] Md. Arifur Rahman, Md. Mukitul Islam, G M Shahir Mahdee, Md. WasiUl Kabir, Improved Segmentation Approach for Plant DiseaseDetection,1st International Conference on Advances in Science, Engineering andRobotics Technology 2019 (ICASERT 2019)

[4] Melike Sardogan, Adem Tuncer, Yunus Ozen, Plant Leaf DiseaseDetection and Classification based on CNN with LVQ Algorithm

[5] Sumathi Bhimavarapu, Vinitha Panicker J, Analysis and Characterization Of Plant Diseases using Transfer Learning, International Conference on Communication and Signal Processing, July 28 - 30, 2020, India

[6] Melike Sardogan, Adem Tuncer, Yunus Oze,n Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm, 3rd International Conference on Computer Science and Engineering



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)