PLANT DISEASE PREDICTION USING CNN

Manoj Praphakar
Assistant Professor
Sri Shakthi Institute of Engineering and Technology
Coimbatore,India

Jenisha R
AssistantProfessor
DepartmentofArtificial Intellegence and Machine Learning
Sri Shakthi Institute of Engineering andTechnology
Coimbatore,India

Dhanush B
DepartmentofArtificial Intellegence and Machine Learning
Sri Shakthi Institute of Engineering andTechnology
Coimbatore,India

Dhanush K
DepartmentofArtificial Intelligence and Machine Learning,
Sri Shakthi Institute of Engineering and Technology
Coimbatore,India

Varshik Daniel L
DepartmentofArtificial Intelligence and Machine Learning
Sri Shakthi Institute of Engineering andTechnology
Coimbatore,India

Rathivarma C
DepartmentofArtificial Intellegence and Machine Learning
Sri Shakthi Institute of Engineering andTechnology
Coimbatore,India

Abstract—*Farmers must apply the appropriate insecticides for their crops. Too many pesticides are harmful to crops and farmland. Getting expert advice will help you avoid misusing chemicals on plants. Plants have been the focus of many researchers to aid farmers and others involved in agriculture. When a disease is visible to the naked eye, it is straightforward to detect. The illness may be discovered and treated early if the farmer has sufficient information and monitors the crops on a regular basis. However, this phase only exists when the disease is extreme or crop output is low. Then there are the different innovations. Farmers will benefit from the introduction of automated disease detection tools. This approach yields outcomes that are suitable for both little and large-scale agricultural cultivation. Importantly, the results are precise, and the disorders aredetected in a very short amount of time. These technologies rely heavily on deep learning and neural networks to function. Deep Convolutional Neural Network is utilized in this study to identify infected and healthy leaves, as well as to detect illness in afflicted plants. The CNN model is designed to suit both healthy and sick leaves; photos are used to train the model, and the output is determined by the input leaf.*

Keywords-Disease, CNN, sick Leaves

I. INTRODUCTION

Modern technologies have given human society the ability to produce enough food to meet the demand of more than 7 billion people. However, food security remains threatened by a number of factors including climate change (Tai et al., 2014), the decline in pollinators (Report of the Plenary of the Intergovernmental Science-PolicyPlatform on Biodiversity Ecosystem and Services on the work of its fourth session, 2016), plant diseases (Strange and Scott, 2005), and others. Plant diseases are not only a threat to food security at the global scale, but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. In the developing world, more than 80 percent of the agricultural production is generated by smallholder farmers (UNEP, 2013), and reports of yield loss of more than 50% due to pests and diseases are common (Harvey et al., 2014). Furthermore, the largest fraction of hungry people (50%) live in smallholder farming households

(Sanchez and Swaminathan, 2005), making smallholder farmers a group that's particularly vulnerable to pathogen-derived disruptions in food supply.

Various efforts have been developed to prevent crop loss due to diseases. Historical approaches of widespread application of pesticides have in the past decade increasingly been supplemented by integrated pest management (IPM) approaches (Ehler, 2006). Independent of the approach, identifying a disease correctly when it first appears is a crucial step for efficient disease management Historically, disease identification has been supported by agricultural extension organizations or other institutions, such as local plant clinics. In more recent times, such efforts have additionally been supported by providing information for disease diagnosis online, leveraging the increasing Internet penetration worldwide. Even more recently, tools based on mobile phones have proliferated, taking advantage of the historically unparalleled rapid uptake of mobile phone technology in all parts of the world (ITU, 2015).

Smartphones in particular offer very novel approaches to help identify diseases because of their computing power, highresolution displays, and extensive built-in sets of accessories, such as advanced HD cameras. It is widely estimated that there will be between 5 and 6 billion smartphones on the globe by 2020. At the end of 2015, already 69% of the world's population had access to mobile broadband coverage, and mobile broadband penetration reached 47% in 2015, a 12-fold increase since 2007 (ITU, 2015). The combined factors of widespread smartphone penetration, HD cameras, and high performance processors in mobile devices lead to a situation where disease diagnosis based on automated image recognition, if technically feasible, can be made available at an unprecedented scale. Here, we demonstrate the technical feasibility using a deep learning approach utilizing 54,306 images of 14 crop species with 26 diseases (or healthy) made openly available through the project PlantVillage (Hughes and Salathé, 2015).

In order to develop accurate image classifiers for the purposes of plant disease diagnosis, we needed a large, verified dataset of images of diseased and healthy plants. Until very recently, such a dataset did not exist, and even smaller datasets were not freely available. To address this problem, the PlantVillage project has begun collecting tens of thousands of images of healthy and diseased crop plants (Hughes and Salathé, 2015), and has made them openly and freely available. Here, we report on the classification of 26 diseases in 14 crop species using 54,306 images with a convolutional neural network.

We measure the performance of our models based on their ability to predict the correct crop-diseases pair, given 38 possible classes. The best performing model achieves a mean F1 score of 0.9934 (overall accuracy of 99.35%), hence demonstrating the technical feasibility of our approach. Our results are a first step toward a smartphone-assisted plant disease diagnosis system.

II.LITERATURE REVIEW

K.Muthukannan and colleagues discovered spot infections in leaves and categorized them according to the diseased leaf categories using various machine learning algorithms. LVQ - Learning Vector Quantization, FFNN - Feed Forward Neural Network, and RBFN - Radial Basis Function Networks were utilized to diagnose diseased plant leaves by analyzing the collection of form and texture data from the afflicted leaf picture. The simulation showed that the proposed system is effective. With the support of this work, a machine learning-based system for improving crop quality in the Indian economy can be developed.

The study of plant leaf disease detection by Malvika Ranjan and colleagues starts with image capturing. Color data, such as HSV features, are retrieved from the segmentation results, and an artificial neural network (ANN) is then trained by selecting feature values that can effectively discriminate between healthy and sick samples. Using a combination of image data processing methods and ann, the current study suggests a method for identifying cotton leaf illnesses early and reliably.

A Deep-Learning-Based Detection for Real-Time Recognition of Tomato Plant Pest and Diseases Alvaro Fuentes and colleagues look at three types of detectors: the Faster Region-based Cnns (Faster R-CNN), the Area Convolutional Neural Network (R-FCN), and the Single Action Multibox Detector (SSD), all of which are referred to as "deep learning meta-architectures" in this paper. We use "deep feature extractors" like VGG net and Residual Network to merge every one of these meta-architectures (ResNet). We show how deep morpho and feature extractors perform, and we also suggest a way for locally and globally category labeling and feature extraction to improve accuracy and reduce false positives throughout training. We train and test our systems end-to-end on our large Tomato Diseases and Pests Dataset, which contains challenging images of diseases and pests, including several inter-and extra-class variations, such as infection status and location in the plant.

This paper outlines a method for accurately identifying apple leaf diseases. Building enough unhealthy photos and unique architecture of a deep CNN based on AlexNet are required to identify apple leaf infections. Using a database of 13,689 pictures of sick apple leaves, the suggested deep CNN model is meant to detect four common apple leaf disorders. The total accuracy of the suggested illness detection model is 97.62 percent. When compared to the AlexNet model, the parameters of the suggested model were reduced by 51,206,928 and the model's accuracy was enhanced by 10.83 percent with produced pathological pictures. According to this research, the deep learning model for disease management may be more accurate and have a faster convergence rate, therefore enhancing disease control.

Prasanna Mohanty and colleagues developed a deep convolutional neural network using deep learning to detect 14 different crops and 26 illnesses. On a held-out test set, the training set model obtained an accuracy of 99.35 percent, illustrating the practicality of this strategy. The model still obtains a 31.4 percent accuracy when tested on a collection of photographs acquired from reputable web sources - i.e. images shot under settings distinct from those used for training. While this accuracy is substantially greater than the one based on random selection 2.6%, a larger collection of training data is required to increase overall accuracy.

To diagnose plant leaf illnesses, Ashwin Dhakal and colleagues created a model that includes feature extraction, segmentation, and classification of collected leaf patterns. Yellow Leaf Curl Virus, Bacterial Spot, Late Blight, and Healthy Leaf are the four classifier labels employed. With 20 epochs, the retrieved characteristics are fitted into the neural network. Various neural networkbased topologies are used, with the greatest accuracy of 98.59 percent in predicting plant disease.

S. Khirade and colleagues used digital image processing algorithms and BPNN - backpropagation neural networks to solve the problems of detection of plant diseases in 2015. Different techniques for identifying plant disease using photographs of leaves have been developed by the authors. To segment the contaminated section of the leaf, they used Otsu's thresholding, followed by border detection and spot detection algorithms. They then extracted properties such as colour, texture, morphology, edges, and so on in order to classify plant diseases. The BPNN algorithm is used to classify or identify plant diseases.

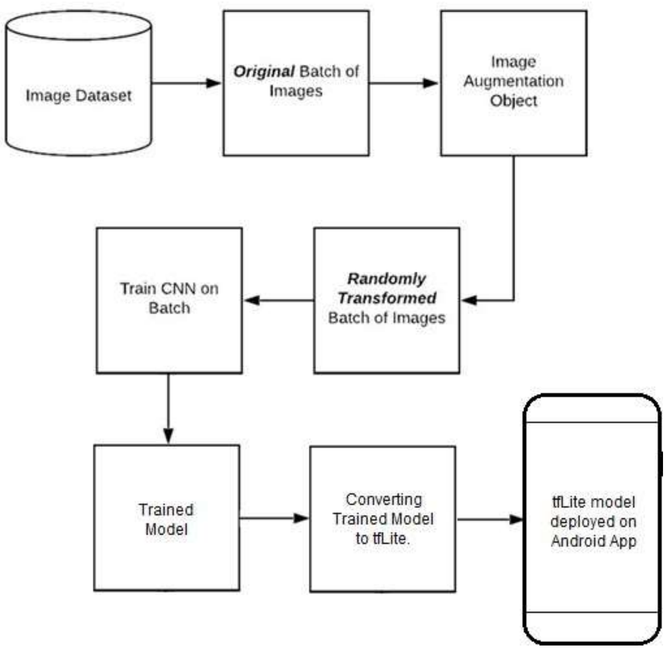
In 2017, Peyman Moghadam and colleagues proved the use of hyperspectral imaging in the diagnosis of plant diseases. In this research, the VNIR - visible and nearinfrared and SWIR - short-wave infrared spectrums were employed. For leaf segmentation, the authors employed the k-means clustering approach in the spectral domain. To remove the grid from hyperspectral pictures, they suggested a unique grid removal technique. The accuracy of vegetation indices in the VNIR spectral range was 83 percent, and full-spectrum accuracy was 93 percent. Despite the fact that the suggested technique achieved improved accuracy, it necessitates the use of a hyperspectral camera with 324 spectral bands, making the solution prohibitively expensive.

Sharath D. M. and colleagues created a Bacterial Blight detection method for Pomegranate plants in 2019 utilizing variables including colour, mean, homogeneity, SD, variance, correlation, entropy, and edges. Grab cut segmentation was used by the authors to segment the image's region of interest. The edges of the photos were extracted using the Canny edge detector. The authors have succeeded in developing a system that can forecast the degree of infection in the fruit.

The convolutional neural network was used by Garima Shrestha and colleagues to identify plant disease in 2020. With an accuracy of 88.80 percent, the authors were able to classify 12 plant diseases. Experimentation was carried out by using a collection of 3000 high-resolution RGB photographs. The convolutional layer and pooling layer have 3 blocks in this network. Eventually, the network becomes very expensive as a result of this. Additionally, the model's F1 score is 0.12, which is extremely poor due to the significant amount of erroneous negative predictions.

III.PROPOSED SYSTEM

We are building a neural network model for image classification. this model will be deployed on the android application for live detection of plant leaf disease through an android phone's camera. The recognition and classification procedures are depicted in the below figure



(1) The first step is to collect data. We are using the PlantVillage Dataset, which is widely available. This dataset was released by crowdAI.

(2) Pre-processing and Augmentation of the collected dataset is done using pre-processing and Image-data generator API by Keras.

(3) Building CNN(Convolutional Neural Network) Model (Vgg-19 architecture) for classification of various plant diseases.

(4) Developed model will be deployed on the Android Application with help of TensorFlow lite

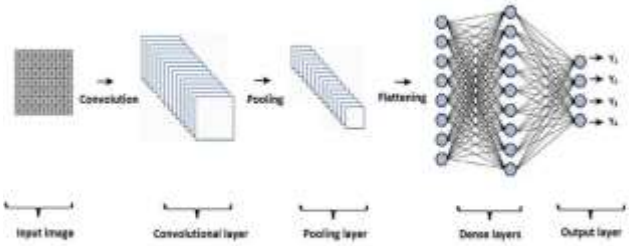
IV.DATASET DESCRIPTION

We analyze 54,306 images of plant leaves, which have a spread of 38 class labels assigned to them. Each class label is a cropdisease pair, and we make an attempt to predict the crop-disease pair given just the image of the plant leaf. Figure shows one example each from every crop-disease pair from the PlantVillage dataset. In all the approaches described in this paper, we resize the images to 256 × 256 pixels, and we perform both the model optimization and predictions on these downscaled images.

Across all our experiments, we use three different versions of the whole PlantVillage dataset. We start with the PlantVillage dataset as it is, in color; then we experiment with a gray-scaled version of the PlantVillage dataset, and finally we run all the experiments on a version of the PlantVillage dataset where the leaves were segmented, hence removing all the extra background information which might have the potential to introduce some inherent bias in the dataset due to the regularized process of data collection in case of PlantVillage dataset. Segmentation was automated by the means of a script tuned to perform well on our particular dataset. We chose a technique based on a set of masks generated by analysis of the color, lightness and saturation components of different parts of the images in several color spaces (Lab and HSB). One of the steps of that processing also allowed us to easily fix color casts, which happened to be very strong in some of the subsets of the dataset, thus removing another potential bias.

This set of experiments was designed to understand if the neural network actually learns the "notion" of plant diseases, or if it is just learning the inherent biases in the dataset. Figure 2 shows the different versions of the same leaf for a randomly selected set of leaves

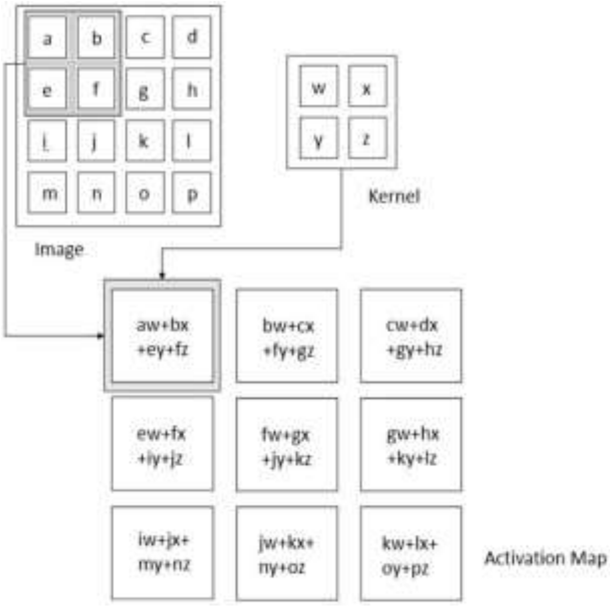
V.CONVOLUTIONALNEURAL NETWORK ARCHITECTURE



AConvolutionalNeuralNetworkhasthreelayers:aconvolutionallayer,apoolinglayer,andafullyconnectedlayer.Figure showsall layerstogether.

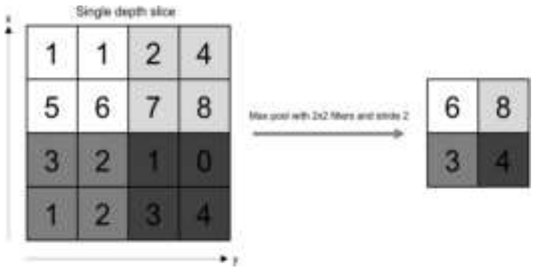
5.1.Convolutional Layer

Convolutionallayer:producesanactivationmapbyscanningthepicturesseveralpixelsatatimeusingafilter.Figshowstheinternalworkingoftheconvolutionlayer.



5.2.Pooling Layer

Pooling layer: reduces the amount of data created by theconvolutional layer so that it is stored more efficiently.Fig showstheinternalworkingofthepoolinglayer.

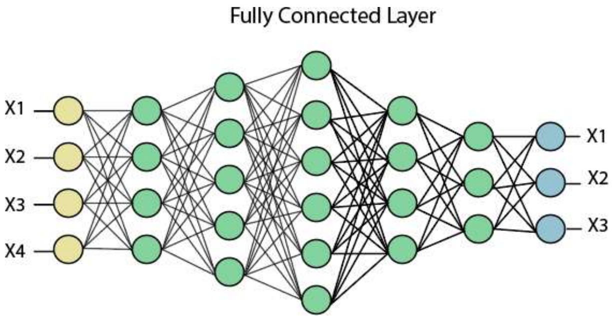


5.3. Fully Connected Layer

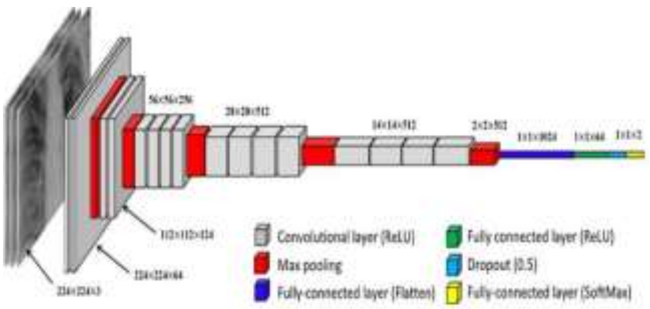
Fully connected input layer–The preceding layers'outputis"flattened"andturnedintoasinglevectorwhich isusedasaninputfor thenextstage.

The first fully connected layer – adds weights to theinputs from the feature analysis to anticipate the properlabel.

Fullyconnectedoutputlayer–offerstheprobabilityforeachlabel inthe end. Figshowstheinternalworkingoffullyconnectedlayer



VGG19 is a sophisticated CNN with pre-trained layersandathoroughgraspofhowanimageisdefinedintermsof form, color, and structure. VGG19 is a deep neuralnetwork that has been trained on millions of photos withchallengingclassification problems.



VI.APPROACH

Weevaluatetheapplicabilityofdeepconvolutionalneur alnetworks for the classification problem described above. Wefocusontwopopulararchitectures,namelyAlexNet(Krizhe vskyet al., 2012), and GoogLeNet (Szegedy et al., 2015), which weredesigned in the context of the "Large Scale Visual RecognitionChallenge"(ILSVRC)(Russakovskyetal.,2015)f ortheImageNetdataset(Dengetal.,2009).

We analyze the performance of both these architectures onthe PlantVillage dataset by training the model from scratch inone case, and then by adapting already trained models (trainedon the ImageNet dataset) using transfer learning. In case oftransferlearning,were-initializetheweightsoflayerfc8incaseof AlexNet, and of the loss {1,2,3}/classifier layers in case ofGoogLeNet. Then, when training the model, we do not limit thelearning of any of the layers, as is sometimes done for transferlearning. In other words, the key difference between these twolearning approaches (transfer vs. training from scratch) is in theinitial state of weights of a few layers, which lets the transferlearningapproachexploitthelargeamountofvisualkno wledgealreadylearnedbythepre-trainedAlexNetandGoogleNetmodelsextractedfromImageNe t(Russakovskyetal.,2015).

Tosummarize,wehaveatotalof60experimentalconfigu rations,whichvaryonthefollowingparameters:

1. **Choiceofdeeplearningarchitecture:**
   AlexNet,
   GoogLeN
   et.
2. **Choiceoftrainingmechanism:**
   Transfer
   Learning,Trainingfr
   omScratch.
3. **Choiceofdatasettype:**
   Color,Gr
   ayscale,
   LeafSegmented.
4. **Choiceoftraining-testingsetdistribution:**
   Train:80%,Test:20%,
   Train:60%,Test:40%,
   Train:50%,Test:50%,
   Train:40%,Test:60%,
   Train:20%,Test:80%.

Toenableafaircomparisonbetweentheresultsofalltheexp erimentalconfigurations,wealsotriedtostandardizethehyper-parametersacrossalltheexperiments,andweusedthe followinghyper-parametersinalloftheexperiments:

- Solvertype:StochasticGradientDescent,
- Baselearningrate:0.005,
- Learningratepolicy:Step(decreasesbyafactorof10every30/3e pochs),
- Momentum:0.9,
- Weightdecay:0.0005,
- Gamma:0.1,
- Batchsize:24(incaseofGoogLeNet),100(incaseofAlexNet).

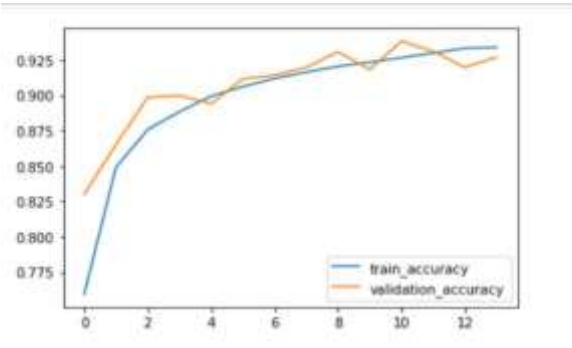
Alltheaboveexperimentswereconductedusingourownforko fCaffe(Jiaetal.,2014),whichisafast,opensourceframeworkf ordeeplearning.Thebasicresults,suchastheoverallaccuracy canalsobereplicatedusingastandardinstanceofcaffe

VII.MEASURMENT OF PERFORMANCE

To get a sense of how our approaches will perform on new unseendata, and also to keep a track of if any of our approaches areoverfitting,werunallourexperimentsacrossawholerangeof train-test set splits, namely 80–20 (80% of the whole datasetusedfortraining,and20%fortesting),60–40(60%ofthewholedatasetusedfortraining,and40%fortesting),50–50(50%of thewholedatasetusedfortraining,and50%fortesting),40–60(40%ofthewholedatasetusedfortraining,and60%fortesting)and finally 20–80 (20% of the whole dataset used for training,and 80% for testing). It must be noted that in many cases, thePlantVillagedatasethasmultipleimagesofthesameleaf(takenfrom different orientations), and we have the mappings of suchcasesfor41,112imagesoutofthe54,306images;andduring all thesetest-trainsplits,wemakesurealltheimagesofthesameleafgoes either in the training set or the testing set. Further, for everyexperiment,wecomputethemeanprecision,meanrecall,meanF$_1$ score, along with the overall accuracy over the whole period oftraining at regular intervals (at the end of every epoch). We usethefinalmeanF$_1$scoreforthecomparisonofresultsacrossallofthedifferentexperimentalconfigurations.

VIII.RESULT

A95.6%accuracyratewasachievedusingearlystoppingwhile Training the model on 50 epochs. Figure 7 depictsthe visualization of training and validation accuracy. Theresult of detecting and recognizing a strawberry plant isshowninFigure8.Ontheleft,ahealthyplantleaf,andontheright,asickinfectedplant.Theresultofdetectingandrecognizing a potato plant is shown in Figure 9. On theleft, a healthy plant leaf, and on the right, a sick infectedplant.



IX.CONCLUSION AND FUTURE WORK

Wearesuccessfulincreatingdiseaseclassificationtechniques used for plant leaf disease detection. A deeplearning model that can be used for automatic detectionandclassificationofplantleafdiseasesiscreated.Tomato,strawberry,soybean,raspberry,potato,corn,Pepperbell,peach,orange,grape,cherry,blueberry,appleare 13 species on which the proposed model is tested. 38classesofplantsweretakenforidentificationthroughthiswork. We were successfully able to work with the imagedata generator API by Keras. Through this, we were abletodoimage-processingtasks.Wewerealsoabletocreatethevgg-19modelwhichisanadvancedconvolutionmodel and train the model with the data for prediction.Thepredictionisdonebyourmodelisalmostcorrect.Wehavesuccessfullydeployedthesemodelontheandroid app and are trying to increase the accuracy of the androidapp aswell asthemodel.

X.REFERENCE

[1] On the use of depth camera for 3D phenotyping of entire plants.

[2] ImageNet: A large-scale hierarchical image database

[3] Integrated pest management (IPM)

[4] The PASCAL Visual Object Classes (VOC) Challenge

[5] Deep Residual Learning for Image Recognition

[6]Using deep learning for Image based palnt disease prediction.

XI.AUTHORS

First Author- Mr.ManojPraphakar T ,Assistant Professor, Sri Shakthi Institute of Engineering andTechnology,

Second Author- Ms.R.Jenisha, Assistant professor, Sri Shakthi Institute of Technology,

Third Author- Mr.B.Dhanush, Student ,Sri Shakthi Institute of Engineering and Technology,

Fourth Author- Mr.K.Dhanush, Student, Sri Shakthi Institute of Engineering and Technology,

Fifth Author- Mr.L.Varshik Daniel, Student, Sri Shakthi Institute of Engineering and Technology,

Sixth Author- Mr.C.Rathivarma, Student, Sri Shakthi Institute of Engineering and Technology,

Correspondence Author- Mr.L.Varshik Daniel, Student, Sri Shakthi Institute of Engineering and Technology,