International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Plant Leaf Disease Detection using CNN

D.Iruthaya Antony Prethika, P.G, Student,
Department of Computer Science and Engineering, Thoothukudi
V. Revathy, Assistant Professor,
Department of Computer Science and Engineering, Thoothukudi
Address
iruthayaantonyprethika2072000@gmail.com

Abstract— **Plants and crops that are infected by pests have an impact on the country's agricultural production. Usually, farmers or professionals keep a close eye on the plants in order to discover and identify diseases. However, this procedure is frequently time-consuming costly, andimprecise. Plant disease detection can be done by looking for a spot on the diseased plant's leaves. The goal of this paper is to create a Disease Recognition Model that is supported by leaf image classification. To detect plant diseases, we are utilizing image processing with a Convolution neural network (CNN). A convolutional neural network(CNN) is a form of artificial neural network that is specifically intended to process pixel input and is used in image recognition.Farmers do not expertise in leaf disease so they produce less production. Plant leaf diseases detection is the important because profit and loss are depends on production.**

Index Terms: Image processing,Crops,Support vector Machine,Plant disease,Classification.

Keywords— **Heart Disease, Diabetes, Machine Learning.**

I. Introduction

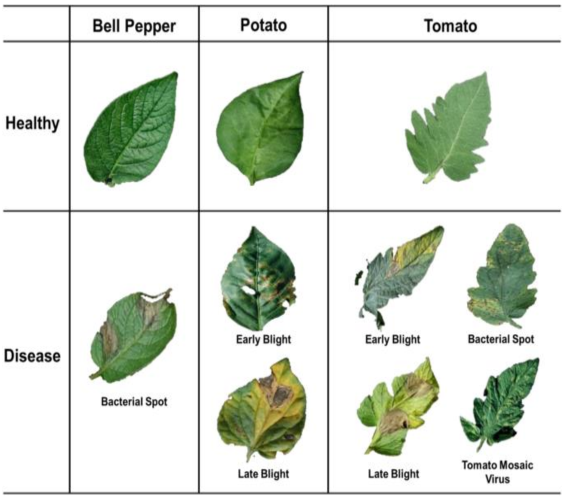
Agricultural production is a very old means of obtaining food. It is a vital source of income for people all around the world. No one can exist in our world without food. Plants are crucial not only for humans, but also for animals who rely on them for food, oxygen, and other necessities. The government and experts are takingsignificant initiatives to enhance food production, and they are working successfully in the real world.environment, Plants and crops that are infected by pests have an impact on the country's agricultural production.Usually, farmers or professionals keep a close eye on the plants in order to discover and identify diseases.imprecise. Plant disease detection can be done by looking for a spot on the diseased plant's leaves. The goal of this paper is to create a Disease Recognition Model that is supported by leaf image classification. To detect plant diseases, we are utilizing image processing with a Convolution neural network (CNN). A convolutional neural network(CNN) is a form of artificial neural network that is specifically intended to process pixel input and is used in image recognition.Farmers do not expertise in leaf disease so they produce less production. Plant leaf diseases detection is the important because profit and loss are depends on production. CNN is the solution for leaf disease detection and classification. Main aim of this research is to detect the apple, grape, corn, potato and tomato plants leaf diseases. Plant leaf diseases are monitoring of large fields of crops disease detection, and thus automatically detected the some feature of diseases as per that provide medical treatment. Proposed Deep CNN model has been compared with popular transfer learning approach such as VGG16. Plant leaf disease detection has wide range of applications available in various fields such as Biological Research and in Agriculture Institute. Plant leaf disease detection is the one of the required research topic as itmay prove benefits in monitoring large fields of crops, and thus automatically detect the symptoms ofdiseases as soon as they appear onplant leaves.

II. Data Collection and Processing

To detect plant leaf diseases, the first step is to collect a dataset of images with both healthy and diseased leaves that accurately represent real-world scenarios. The collected images are preprocessed by removing noise and irrelevant information through techniques such as normalization and data augmentation.

analytics we follow the below steps :

1. Data Collection
2. Preprocessing
3. Model selection
4. Data Visualization
5. Data Modeling
6. Data Training
7. Data Testing
8. Data Prediction
9. Evaluation
10. Optimization.



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

III. Design Implementation and testing

Image processing algorithms, particularly deep convolutional neural networks (CNNs), have gained popularity as an effective approach to detect and diagnose diseases in plants. These algorithms analyze leaf images and use machine learning techniques to identify and classify diseases accurately.

A. Module Building

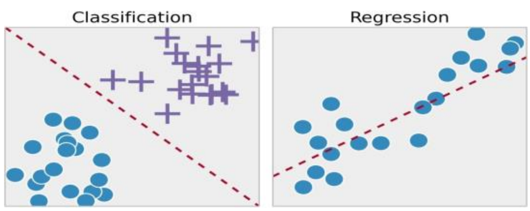
There is a very wide range of machine learning algorithms to choose from, most of which are available in the python library Scikit-learn. However, most of the implementations of these algorithms do not accept sparse matrices as inputs, and since we have a large number of nominal features coming from our n-grams features it is imperative that we encode our features in a sparse matrix. Out of the algorithms that do support sparse matrices in Scikit-learn, I ended up trying naive Bayes, logistic regression and support vector machine (SVM) with a linear kernel. I got the best results in cross validation using SVM with aneuclidean regularization coefficient of 0.1.

- **Supervised learning**
- **Unsupervised learning**
- **Semi Supervised learning**
- **Reinforcement**

.

1)Supervised learning

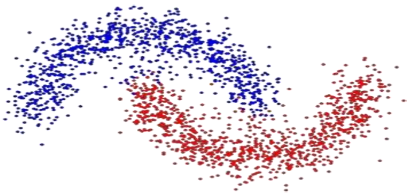
Supervised learning is the task of inferring a function from labeled training data. By fitting to the labeled training set, we want to find the most optimal model parameters to predict unknown labels on other objects (test set). If the label is a real number, we call the task regression. If the label is from the limited number of values, where these values are unordered, then it's classification.



2)UnSupervised learning

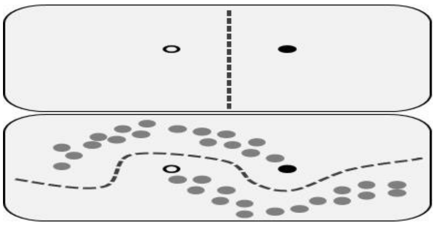
In unsupervised learning we have less information about objects, in particular, the train set is unlabeled. What is our goal now? It's possible to observe some similarities between groups of objects and include them in appropriate clusters. Some objects can differ hugely from all clusters, in this way we assume these objects to be anomalies.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930



3)Semi-Supervised learning

Semi-supervised learning tasks include both problems we described earlier: they use labeled and unlabeled data. That is a great opportunity for those who can't afford labeling their data. The method allows us to significantly improve accuracy, because we can use unlabeled data in the train set with a small amount of labeled data.



4)Reinforcement learning

Reinforcement learning is not like any of our previous tasks because we don't have labeled or unlabeled datasets here. RL is an area of machine learning concerned with how software agents ought to take actions in some environment to maximize some notion of cumulative reward.

1.Data Collection

Data collection is the gathering of information from multiple sources, and data analytics is the processing of that information in order to derive usable insights. To process and gain insights from data acquired from various sources and methods, specific data analysis methodologies and tools are required.

Training Dataset:

To detect plant leaf diseases, the first step is to collect a dataset of images with both healthy and diseased leaves that accurately represent real-world scenarios. The collected images are preprocessed by removing noise and irrelevant information through techniques such as normalization and data augmentation.The dataset used consisted of over 20,000 images, and CNN was used for classification, including 12. classes for diseased leaves and 3 classes for healthy leaves. The model achieved an accuracy of 98.29% for. training and 98.029% for the testing dataset

The data's in the file are:

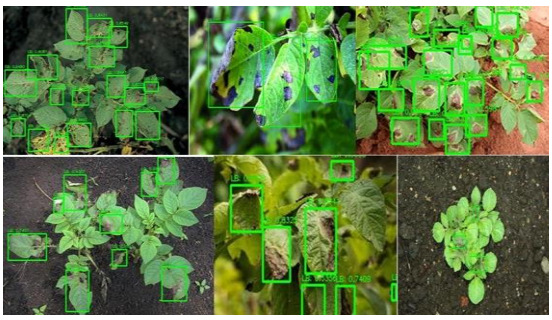
```
/content/drive/My Drive/leafdisease dataset/Apple___Apple_scab
/content/drive/My Drive/leafdisease dataset/Apple___Black_rot
/content/drive/My Drive/leafdisease dataset/Apple___Cedar_apple_rust
/content/drive/My Drive/leafdisease dataset/Apple___healthy
/content/drive/My Drive/leafdisease dataset/Grape___Black_rot
/content/drive/My Drive/leafdisease dataset/Grape___Esca_(Black_Measles)
/content/drive/My Drive/leafdisease dataset/Grape___healthy
/content/drive/My Drive/leafdisease dataset/Grape___Leaf_blight_(Isariopsis_Leaf_Spot)
/content/drive/My Drive/leafdisease dataset/Corn___Cercospora_leaf_spot Gray_leaf_spot
/content/drive/My Drive/leafdisease dataset/Corn___Common_rust
/content/drive/My Drive/leafdisease dataset/Corn___healthy
/content/drive/My Drive/leafdisease dataset/Corn___Northern_Leaf_Blight
/content/drive/My Drive/leafdisease dataset/Potato___Early_blight
/content/drive/My Drive/leafdisease dataset/Potato___healthy
/content/drive/My Drive/leafdisease dataset/Potato___Late_blight
/content/drive/My Drive/leafdisease dataset/Tomato___Bacterial_spot
/content/drive/My Drive/leafdisease dataset/Tomato___Early_blight
/content/drive/My Drive/leafdisease dataset/Tomato___healthy
/content/drive/My Drive/leafdisease dataset/Tomato___Late_blight
/content/drive/My Drive/leafdisease dataset/Tomato___Leaf_Mold
/content/drive/My Drive/leafdisease dataset/Tomato___Septoria_leaf_spot
/content/drive/My Drive/leafdisease dataset/Tomato___Spider_mites Two-spotted_spider_mite
/content/drive/My Drive/leafdisease dataset/Tomato___Target_Spot
/content/drive/My Drive/leafdisease dataset/Tomato___Tomato_mosaic_virus
X_data shape: (24000, 256, 256, 3)
y_data shape: (24000,)
```

Testing Dataset:

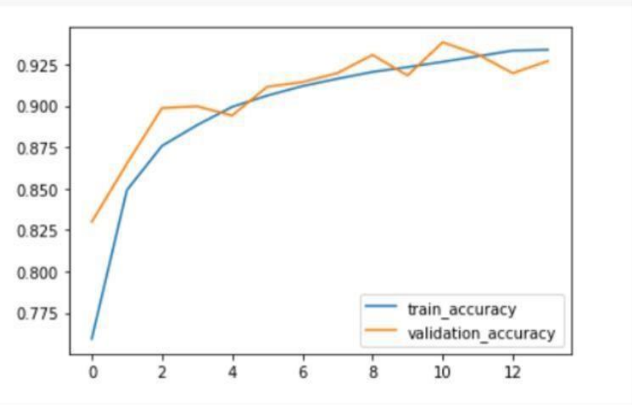
To detect plant leaf diseases, the first step is to collect a dataset of images with both healthy and diseased leaves that accurately represent real-world scenarios The collected images are preprocessed by removing noise and irrelevant information through techniques such as normalization and data augmentation**.**

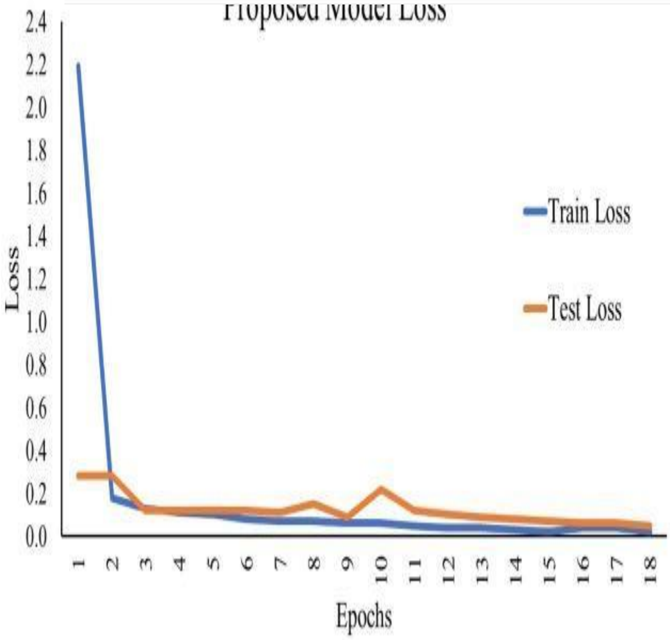
2.Disease leaf Dataset:

The first step is to perform data cleaning. If there's any incorrect, redundant, irrelevant or partially formatted data that will be removed or changed, as these data are not beneficial nor necessary because it might produce false results. Data cleaning is the process of eliminating or changing data that is erroneous, incomplete, irrelevant, redundant, or incorrectly formatted in order to prepare it for analysis. When it comes to data analysis, this data is usually not necessary or beneficial because it can slow down the process or produce false results. To reduce this we first find months for the dates given in the dataset. Then the wind speed is rewritten for 0.5 intervals. For example: wind speeds between 3.25 and 3.75 turns 3.5,wind speeds between 3.75 and 4.25 turns 4.0.The wind direction is rewritten for 30 intervals. For example: wind directions between 15 and 45 turns 30, wind speeds between 45 and 75 turns 60



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930





3.Implementation work:

Apple, grape, potato, and tomato plant leaves which are categorized total 24 types of labels apple label namely: Apple scab, Black rot, Apple rust, and healthy. Corn label namely: Corn Cercospora Gray spot, Corn rust, Corn healthy, Corn Blight [11][13]. Grape label namely: Black rot, Esca, healthy, and Leaf blight. Potato label namely: Early blight, healthy, and Late blight. Tomato label namely: bacterial spot, early blight, healthy, late blight, leaf mold, septoria leaf spot, spider mite, target sport, mosaic virus, and yellow leaf curl virus[11][13].The dataset consist of 31,119 images of apple, corn, grape, potato and tomato, out of 31,119 images 24000 images are used. all Images are resized into 256 x 256,that images divided into two parts training and testing dataset, the whole range of the train test split using 80-20 (80%of the

Classes	no. of images	used images
Apple_scab	1000	1000
Apple_black_rot	1000	1000
Apple_cedar_apple_rust	1000	1000
Apple_healthy	1645	1000
Corn_gray_leaf_spot	1000	1000
Corn_common_rust	1192	1000
Corn_northern_leaf_blight	1000	1000
Corn_healthy	1162	1000
Grape_black_rot	1180	1000
Grape_black_measles	1383	1000
Grape_leaf_blight	1076	1000
Grape_healthy	1000	1000
Potato_early_blight	1000	1000
Potato_healthy	1000	1000
Potato_late_blight	1000	1000
Tomato_bacterial_spot	2127	1000
Tomato_early_blight	1591	1000
Tomato_healthy	1909	1000
Tomato_late_blight	1000	1000
Tomato_leaf_mold	1000	1000
Tomato_septoria_leaf_spot	1707	1000
Tomato_spider_mites_two-spotted_spider_mite	1676	1000
Tomato_target_spot	1404	1000
Tomato_mosaic_virus	1000	1000
Total images	30052	24000

Implementation

whole dataset used for the training and 20% for the testing)[11][13]. Then train CNN model

4.Data Visualization

The graphical depiction of information and data is known as data visualization. Data visualization tools make it easy to examine and comprehend trends, outliers, and patterns in data by employing visual elements like charts, graphs, and maps. Here we data analysis becauseit helps us to spot developing patterns

Training Vs Test model Model

Proposed workflow

Convolutional neural networks (CNN) can be used for the computational model creation that works on the unstructured image inputs and converts to output labels of corresponding classification.complexity of the extracted features increases. The size of the filter is fixed to 5 × 5 whereas number of filters is increased progressively as we move from one block to another. The number of filters is

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

20 in the first convolution block while it is increased to 50 in the second and 80 in the third..

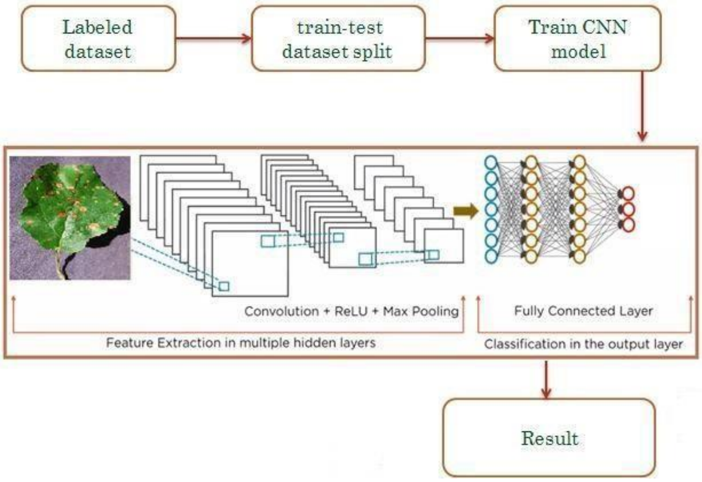


Figure 4. Work flow

IV.CONCLUSIONS

The project demonstrated the effectiveness of deep learning techniques in detecting plant leaf diseases.The project achieved high accuracy rates in classifying the different types of plant leaf diseases using a Convolutional Neural Network (CNN) architecture.

A 95.6% accuracy rate was achieved using early stopping while Training the model on 50 epochs. Figure 7 depicts the visualization of training and validation accuracy. The result of detecting and recognizing a strawberry plant is shown in Figure 8. On the left, a healthy plant leaf, and on the right, a sick infected plant. The result of detecting and recognizing a potato plant is shown in. On the left, a healthy plant leaf, and on the right, a sick infected plant.

V.REFERENCES

[1]. Liu, Bin, "Identification of apple leaf diseases based on deep convolutional neural networks

[2]. Jeon, Wang-Su, and Sang-Yong Rhee. "Plant leaf recognition using a convolution neural network." International Journal of Fuzzy Logic and Intelligent Systems 17.1 (2017): 26-34.

[3]. Amara, Jihen, Bassem Bouaziz, and Alsayed Algergawy. "A Deep Learning-based Approach for Banana Leaf Diseases Classification." BTW (Workshops). 2017.

[4]. Lee, Sue Han, et al. "How deep learning extracts and learns leaf features for plant classification." Pattern Recognition 71 (2017): 1-13.

[5]. Sladojevic, Srdjan, et al. "Deep neural networks-based recognition of plant diseases by leaf image classification." Computational intelligence and neuroscience 2016 (2016).

[6]. Lee, Sue Han, et al. "Plant Identification System based on a Convolutional Neural Network for the LifeClef 2016 Plant Classification Task." CLEF (Working Notes). 2016.

[7]. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference.

[8].K.Padmavathi, and K.Thangadurai,"Implementation of RGB and Gray scale images in plant leaves disease detection –comparative study,"

[9]. Kiran R. Gavhale, and U. Gawande, "An Overview of the Research on Plant Leaves International Journal of Pure and Applied Mathematics Special Issue 882 Disease detection using Image Processing Techniques," IOSR J. of Compu. Eng

[10]. Y. Q. Xia, Y. Li, and C. Li, "Intelligent Diagnose System of Wheat Diseases Based on Android Phone," J. of Infor. & Compu. Sci., vol. 12, pp. 6845-6852, Dec. 2015.