JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

An Adaptive Machine Learning Approach for Image Based Plant Leaf Disease Identification with Performance Improvement

[1]Pooja Gedam, [2]Prof. Jeetendra Singh Yadav

[1]MTech Scholar, [2]Assistant Professor
Department of Computer Science and Engineering
Bhabha Engineering Research Institute, Bhopal, India

Abstract : Artificial Intelligence offers vast opportunities for application in agriculture; there still exists a lack of familiarity with high tech machine learning solutions in farms across most parts of the world. AI systems also need a lot of data to train machines and to make precise predictions. Tomatoes (Solanum lycopersicum) can be grown on almost any moderately well-drained soil type. This research presents an adaptive machine learning approach for image based plant leaf disease identification with performance improvement. Simulation is performed using Python sypder 3.7 version. The overall accuracy is achieved 98% in different plant leaf disease identification.

***IndexTerms* - Sypder, Python, Accuracy, AL, Plant, Disease, Machine Learning.**

I. Introduction

The agri-e-number cruncher as a shrewd application assist the brilliant farmer with picking the most appropriate crop and reasonableness dependent on a few reliance factors. The farmer can utilize the savvy number cruncher and simply pick the ideal crop to be developed over his favored inclusion space of homestead. Then, at that point any remaining required data sources dependent on different reliance factors are naturally distinguished and taken by the e-adding machine and gives the assessment results. This yield result gives valuable information on assessment of manures cost/amount, water, seeds, development hardware cost and Work Day endeavors/cost with Work Day exertion circulation on schedule graph of crop life cycle, crop yield alongside extrapolated market cost at the gather time and its productivity. All the required sources of info which are both direct and non-straight in nature are taken by farmer's information base, outer data sources referenced before. The sources of info get prepared by machine learning strategies and produce the assessment with achievability concentrate so the farmer can pick the ideal crop for development.

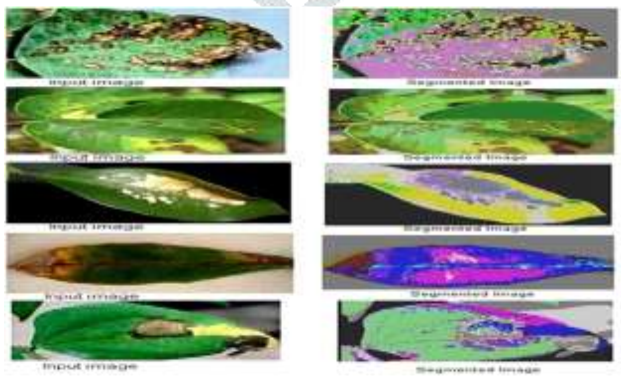

Figure 1: Diseased image with segmented image

The crop care administration direction traverses directly from the planting of seeds as start point till the hour of gathering as endpoint. The complex organized information examined from IoT sensors from the fields are investigated alongside the information gathered from wellsprings of data locales alongside domain master inputs any place required through Artificial Intelligence methods. After the investigation of complete information, the general restorative thing to do is inferred out of PID (Corresponding Indispensable and Differential) regulator instrument. In like manner, the restorative measures are made aware of the farmer on their PDA to focus on the activity dependent on seriousness and desperation.

Artificial Intelligence offers vast opportunities for application in agriculture; there still exists a lack of familiarity with high tech machine learning solutions in farms across most parts of the world. Exposure of farming to external factors like weather conditions, soil conditions and presence of pests is quite a lot. AI systems also need a lot of data to train machines and to make precise predictions. In case of vast agricultural land, though spatial data can be gathered easily, temporal data is hard to get. For example, most of the crop-specific data can be obtained only once in a year when the crops are growing. Since the data infrastructure takes time to mature, it requires a significant amount of time to build a robust machine learning model. This is one reason why AI sees a lot of use in agronomic products such as seeds, fertilizer, pesticides and so on rather than in-field precision solutions.

II. METHODOLOGY

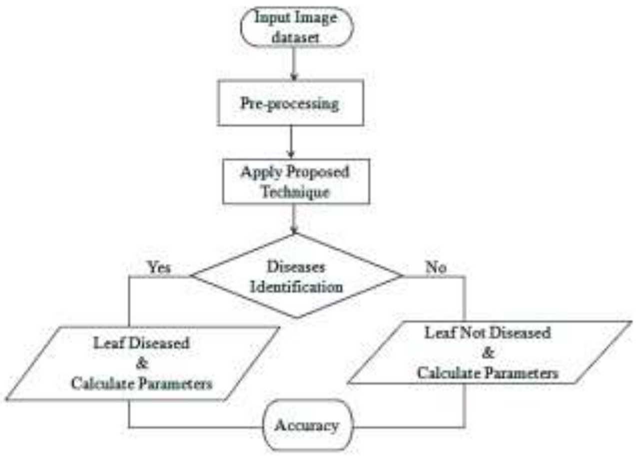

Figure 2: Flow Chart

The proposed methodology is based on the following sub modules-

- Data Selection and Loading
- Data Pre-processing
- Feature Extraction & Feature Optimization
- Splitting Dataset into Train and Test Data
- Classification
- Prediction
- Result Generation

Data Selection and Loading

- The data selection is the process of selecting the data in form of image for detecting the plant spices.
- In this research, the random dataset is used for detecting the plant disease.

Data Pre-processing

- Data pre-processing is the process of removing the unwanted data from the dataset.
- Missing data removal
- Encoding Categorical data
- Missing data removal: In this process, the null values such as missing values and Nan values are replaced by 0.
- Missing and duplicate values were removed and data was cleaned of any abnormalities.
- Encoding Categorical data: That categorical data is defined as variables with a finite set of label values.
- That most machine learning algorithms require numerical input and output variables.

Feature Extraction & Feature Optimization

Spider Monkey Optimization (SMO) is a global optimization algorithm inspired by Fission-Fusion social (FFS) structure of spider monkeys during their foraging behavior. SMO has gained popularity in recent years as swarm intelligence based algorithm and is being applied to many engineering optimization problems. Similar to the other population-based algorithms, SMO is a trial and error based collaborative iterative process. The SMO process consists of six phases: Local Leader phase, Global Leader phase, Local Leader Learning phase, Global Leader Learning phase, Local Leader Decision phase and Global Leader Decision phase.

Classification: Support Vector Machine

A support vector machine takes these data points and outputs the hyperplane (which in two dimensions it's simply a line) that best separates the tags. This line is the decision boundary: anything that falls to one side of it we will classify as blue, and anything that falls to the other as red.

Support Vector Machine (SVM) is a supervised calculation that can classify cases by isolating an informational index into at least two classes using a separator. SVM works by: Mapping information to a high-dimensional component space so that information points can be sorted (kernelling), in any event, when the information are not otherwise linearly separable.

A separator between the categories is found; at that point the information is transformed in such a manner that the separator could be drawn as a hyperplane.

III. SIMULATION AND RESULT DISCUSSION

The simulation is performed using python spyder 3.7 software. The python is widely used for the machine learning implementation.

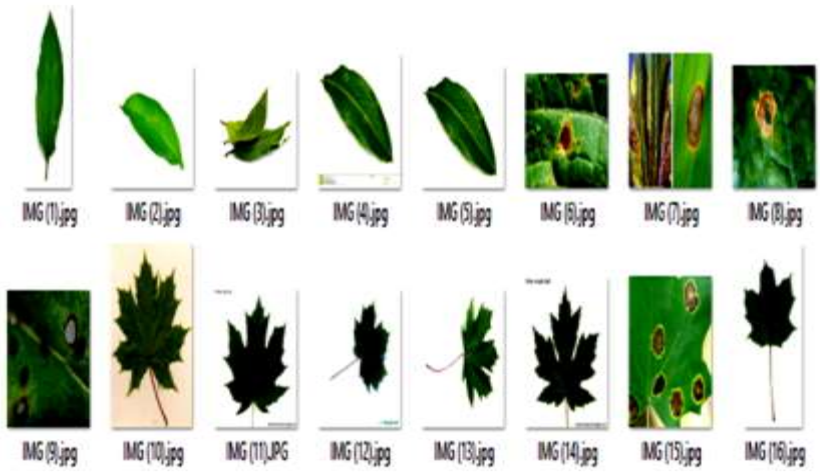

Figure 3: Sample of dataset

Figure 3 is showing the plant leaf image input data. Total 32 images taken with 7 different disease, which includes tomato, banana, ginger, mango, norway maple, onion and paper mulberry.

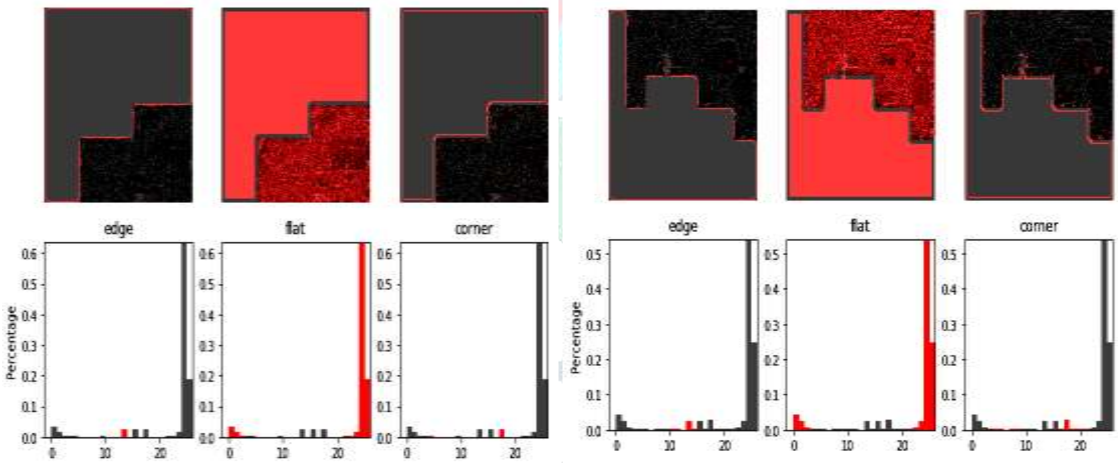

Figure 4: Training of image data

This image data trained by using the image features in terms of edge, flat corner, due to training its learn about various edges, flats and corners at different texture.

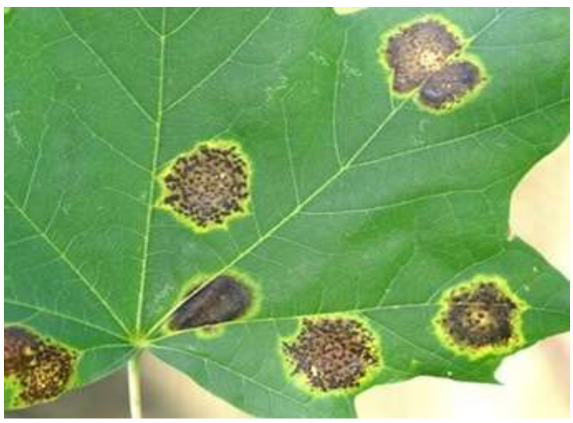

Figure 5: Tomato Leaf input original image

Run: 1
SMO is optimizing "F1"
['At iteration 1 the best fitness is 47160.628221655264']
['At iteration 2 the best fitness is 31652.20550507322']
['At iteration 3 the best fitness is 31652.20550507322']
['At iteration 4 the best fitness is 27471.69488449077']
['At iteration 5 the best fitness is 17202.35797549416']
Results of 1 run are saved in 'csv' file.
Run: 2
SMO is optimizing "F1"
['At iteration 1 the best fitness is 61587.3247115257']
['At iteration 2 the best fitness is 31211.69562119171']
['At iteration 3 the best fitness is 25140.548865874254']
['At iteration 4 the best fitness is 22267.175150865212']
['At iteration 5 the best fitness is 22267.175150865212']

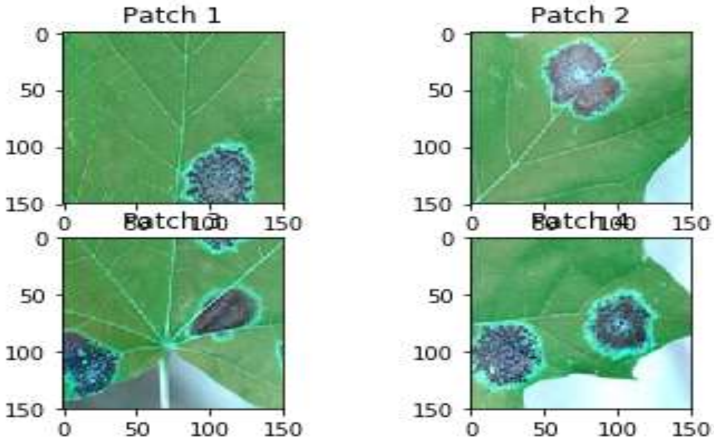

Figure 6: Patch Segmentation Sub plot

Table 1: Simulation Result

Sr. No.	Parameters	Values (%)
1	Accuracy	98
2	Classification error	2
3	Precision	100
4	Recall	96
5	F-measure	97
6	Sensitivity	96
7	Specificity	100

Table 1 is presenting simulation parameters value, which is calculated by the following standard formula-
Precision = True Positive/(True Positive + False Positive)
Recall = True Positive/(True Positive + False Negative)
F1-Score = 2x (precision x recall)/(precision + recall)
Accuracy = (TP + TN)/(TP + TN + FP + FN)
Error Rate = 100 – Accuracy

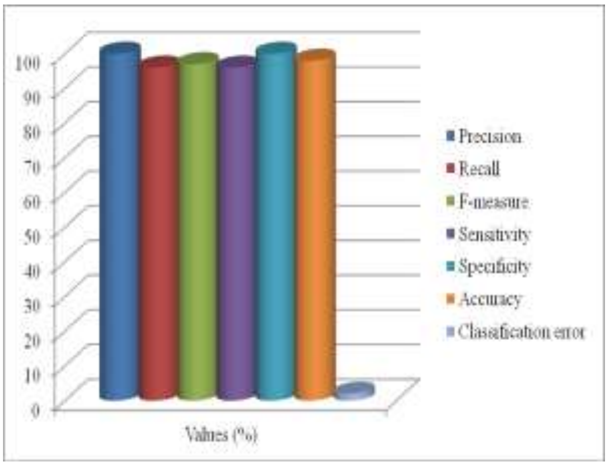

Figure 7: Comparison graph

Table 2: Comparison of proposed work with previous work

Sr No.	Parameters	Previous Work [1]	Proposed Work
1	Method	Restructured residual dense network model	SVM and SMO
2	Accuracy (%)	95	98
3	Error Rate (%)	5	2

Table 2 is showing the results parameters comparison of the previous work and the proposed work. The accuracy achieved by the proposed approach is 98% while previous it is 95%.

IV. CONCLUSION

This research proposed an adaptive machine learning approach for image based plant leaf disease identification with performance improvement. The spider monkey optimization and support vector machine is used to optimize and identified the plant decease prediction. The accuracy achieved by the proposed approach is 98% while previous it is 95%. Error rate is 2% by the proposed work while 5% by the previous work. Therefore proposed methodology achieved better result than the existing results.

REFERENCES

1. C. Zhou, S. Zhou, J. Xing and J. Song, "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network," in IEEE Access, vol. 9, pp. 28822-28831, 2021, doi: 10.1109/ACCESS.2021.3058947.
2. P. K. V, E. G. Rao, G. Anitha and G. K. Kumar, "Plant Disease Detection using Convolutional Neural Networks," 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), 2021, pp. 1473-1476, doi: 10.1109/ICOEI51242.2021.9453045.
3. D. Varshney, B. Babukhanwala, J. Khan, D. Saxena and A. k. Singh, "Machine Learning Techniques for Plant Disease Detection," 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), 2021, pp. 1574-1581, doi: 10.1109/ICOEI51242.2021.9453053.
4. C. S, S. Ghana, S. Singh and P. Poddar, "Deep Learning Model for Image-Based Plant Diseases Detection on Edge Devices," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-5, doi: 10.1109/I2CT51068.2021.9418124.
5. P. Kanade and J. P. Prasad, "Machine Learning Techniques in Plant Conditions Classification and Observation," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 729-734, doi: 10.1109/ICCMC51019.2021.9418386.
6. H. D. Gadade and D. K. Kirange, "Machine Learning Based Identification of Tomato Leaf Diseases at Various Stages of Development," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 814-819, doi: 10.1109/ICCMC51019.2021.9418263.
7. K. Kirti, N. Rajpal and J. Yadav, "Black Measles Disease Identification in Grape Plant (Vitis vinifera) Using Deep Learning," 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2021, pp. 97-101, doi: 10.1109/ICCCIS51004.2021.9397205.
8. Deepa, R. N and C. Shetty, "A Machine Learning Technique for Identification of Plant Diseases in Leaves," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 481-484, doi: 10.1109/ICICT50816.2021.9358797.
9. X. Li and L. Rai, "Apple Leaf Disease Identification and Classification using ResNet Models," 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), 2020, pp. 738-742, doi: 10.1109/ICEICT51264.2020.9334214.
10. S. Kumar, K. Prasad, A. Srilekha, T. Suman, B. P. Rao and J. N. Vamshi Krishna, "Leaf Disease Detection and Classification based on Machine Learning," 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 2020, pp. 361-365, doi: 10.1109/ICSTCEE49637.2020.9277379.
11. K. Trang, L. TonThat and N. G. Minh Thao, "Plant Leaf Disease Identification by Deep Convolutional Autoencoder as a Feature Extraction Approach," 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2020, pp. 522-526, doi: 10.1109/ECTI-CON49241.2020.9158218.
12. H. Yu and C. Son, "Leaf Spot Attention Network for Apple Leaf Disease Identification," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 229-237, doi: 10.1109/CVPRW50498.2020.00034.