Identifying Medicinal Plant Diseases using Image Processing and Deep Learning Techniques

[1]Amruta K. Jadhav
Lecturer
Department of Computer Science & Engineering
MGM's Polytechnic College,
Aurangabad, MH, India

[2]Pravin Yannawar
Associate Professor
Department of Computer Science & IT
Dr. BAM University,
Aurangabad, MH, India

[3]Shivraj Marathe
Project Manager
Research & Development Department
Marg Ehealth Pvt. Ltd,
Delhi, India

Abstract: **Ayurveda unquestionably carries considerable income to India by unfamiliar trade through the fare of ayurvedic prescriptions. Plants' diseases cause significant creation and financial misfortunes in the horticultural and medicinal ventures around the world. Medicinal plants are acquiring consideration in the drug business due to having less destructive impacts responses and less expensive than present day medication. There are different freedoms for headway in delivering a strong classifier that can group medicinal plants precisely progressively. Checking of health and recognition of diseases in plants and trees is a basic issue. This paper presents a strategy for the identification of diseases in medicinal plants based on some significant features removed from its leaf images. The main piece of exploration on a plant disease to distinguish the disease based on CBIR (content-based image retrieval) that is principally worried about the precise identification of diseased medicinal plants. This paper presents a methodology where the plant is recognized based on its leaf highlights, for example, shading histogram and edge histogram. Vigilant edge recognition is likewise valuable to track down the solid edges of leaf of plants and that is utilized to draw the edge histogram which is one of the boundaries for testing. In this paper, different successful and dependable machine learning calculations for plant classifications utilizing leaf images that have been utilized lately are investigated. The survey incorporates the image handling strategies used to identify leaf and concentrate significant leaf features for some machine learning classifiers. These machine learning classifiers are arranged by their presentation when grouping leaf images based on commonplace plant highlights, in particular shape, surface, and a mix of various highlights. The leaf data sets that are freely accessible for programmed plants acknowledgment are investigated too and we finish up with a conversation of conspicuous continuous exploration and openings for improvement around here.**

Index Terms: **Medicinal Plants, Machine Learning, Leaf Identification, Classification**

I. INTRODUCTION

1.1 About Medicinal Plants and their usage in Pharmaceutical Industry

Present day medication is hugely delivered for clinical therapy however numerous first world nations are currently picking customary medication because of the constraint of manufactured medications in controlling and relieving persistent diseases (WHO 1999). Customary medications are utilized widely in the drug business, as guaranteed in (Karami et al. 2017), where a fourth of the around the world endorsed drug are extricated from medicinal plants. This is because of the advantages of medicinal plants that offer generously lower antagonistic responses and more practical when contrasted with engineered drugs (Lulekal et al. 2008). Moreover, bioactive mixtures like phenolics, carotenoids, anthocyanins and tocopherols that can be extricated from medicinal plants (Altemimi et al. 2017) fill in as cancer prevention agents, hostile to allergenic, calming, antibacterial and furthermore against hepatotoxic. In any case, the undertaking of distinguishing medicinal plants physically is confounded and tedious, like other plant acknowledgment and this is because of the accessibility of well-qualified conclusions (Sladojevic et al. 2016; Singh and Misra 2017; Wäldchen et al. 2018). Roused by these issues, analysts presented various programmed plants or leaf acknowledgment frameworks, where many of them used Machine learning draws near.

1.2 About Deep Learning

Machine learning is a part of man-made reasoning which permits machines to recognize examples and settle on choices with negligible human mediation. Machine learning has been utilized to get great acknowledgment, forecast and filtration results on numerous issues like clinical determination, monetary examination, prescient upkeep and image acknowledgment. Right now, there are different kinds of machine learning calculations, and these calculations can be arranged into three classes, directed, unaided and semi supervised. In administered learning, the calculation settles on choices based on the marked information, where the preparation cycle proceeds until the classifier ready to accomplish the most noteworthy exactness (El Mohadab et al. 2018).

There are additionally machine learning calculations that can be prepared without marked information and these calculations are arranged under solo learning (El Mohadab et al. 2018). Now and again, there is a requirement for semi-managed learning, where the calculations are prepared utilizing both named and unlabelled information (Zhu and Goldberg 2009). In this paper, different compelling and solid machine learning calculations like Multilayer Perceptron, Support Vector Machine (SVM), K-Nearest Neighbor (k-NN), Nearest Neighbor Generalized Exemplar, Fuzzy Lattice Reasoning, Naive Bayes Classifier, K-Star Instance Based Classifier, Hyper Pipes Classifier, Random Committee, Random Forest (RF), Bagging Classifier, Voting Feature Intervals, J48 Tree Classifier and Random Tree that are regularly used for plant or leaf classifications lately are audited (Dudani 1976; Güvenir et al. 1998; Gardner and Dorling 1998; Wilson and Martinez 2000; Hothorn and Lausen 2005; Noble 2006; Kaburlasos et al. 2007; Lira et al. 2007; Li and Zhang 2011; Kukreja et al. 2012; Mao and Wang 2013; Davies 2018; Panigrahi and Borah 2018).

II. Related Work

W Sandhika Biswa et al [1], utilizes FCM grouping and neural organization which distinguishes the late curse disease in potato. The images of potato leaves are caught under uncontrolled climate. It is anything but a very robust technique to recognize scourge disease in potato. In any case, here the division is discovered to be troublesome. Fluffy C-means (FCM) is an iterative calculation which discover the bunch communities. These middle's limit a uniqueness capacity and handle the covered information accuracy. It gives more precise outcome in situations where information is inadequate or questionable. In this technique calculation time is longer and it is affectability to commotion. Fluffy C-implies grouping Neural Network comprises of solo fluffy bunching and regulated fake neural organizations which help in accomplishing more ideal outcomes with generally barely any informational collections. The calculation comprises of essentially two stages: (a) Fuzzy c-mean bunching to isolate the disease influenced region alongside foundation (b)to separate influenced leaf region from foundation utilizing neural organization. This model has an excellent precision.

A.A. Joshi et al [2] utilizes KNN classifier and Minimum Distance Classifiers. Basic execution is the benefit of this technique. This technique learns complex models without any problem. However, it has high computational intricacy. K Nearest Neighbor is utilized for measurable assessment and example acknowledgment. It is anything but a simple, basic, and adaptable classification technique. KNN is hearty to uproarious preparing information, yet calculation cost is higher. The tone, zone savvy shape features and so forth are separated in include extraction measure. This is utilized as the contribution for additional classification. For every disease, a different data set has been utilized for preparing and testing. Proposed strategies have been tried for four referenced rice diseases utilizing two classifiers, k-NN and MDC and the exactness accomplished with two classifiers is 87.02%, 89.23% individually. While contrasting and the past methods, it is tracked down that the proposed strategy is 475 bosses as far as time intricacy, precision, number of diseases covered. Shading and shape features are separated in this work. Moreover, surface element can be incorporated and can be checked for the effect of this additional component on the presentation of a calculation. There are other rice diseases with the exception of four canvassed in this work. Future work can be to cover other rice diseases. Similar strategies can be applied to other crops with little adjustments.

John William Orillo et al [3], is utilizing Back Propagation and Artificial Neural Network strategies their paper. In this paper shading-based portrayal of leaf can be removed appropriately. Likewise, high calculations additionally should be possible by this procedure. High computational expense is a downside. Counterfeit Neural Network utilizes forward propagation which is the core of a neural organization. Probabilistic Neural Network is a feed forward calculation which is quicker and more accurate. In this investigation, the images of rice leave utilizing basic methods in image pr MATLAB. The shading attributes of appropriately separated and processed before its neural organization. The field tests w IRRI, Los Baños, Philippines regulated Buresh. The framework had the option to beat insight and gave an exactness extra remark was given to the framework examination of 3.5 identical rate in LCC, which be a test in manual visual investigation program carried out explicit improvement will help for future application.

Pooja Pawar et al [4] utilizes Artificial Neural Network procedure. It functions admirably for more than one yield of various sorts. Be that as it may, include choice is troublesome. Study includes gathering leaf tests of cucumber crop diseases. Work is completed to analyze cucumber crop disease and to give treatment to the identified yield disease. Wool buildup and fine mold are the two diseases in cucumber which are examined in this work. The principal request factual minutes and GLCM are utilized here to remove surface features from the dataset. Classification is completed by utilizing neural organization tool stash of MATLAB 7.10.1. Framework gives classification precision of 80.45%. The proposed work can be material for more than one yield of various sorts. In any case, if there should be an occurrence of other yield type, framework needs to separate features only that can arrange their harvest diseases precisely. In future, classification exactness can be expanded by utilizing extra surface highlights. In this model Gabor channel can likewise be utilized for the extraction of surface element.

Harshal Waghmare. et al [5] utilizes Multi Vector Support Machine in the paper discovery and Classification of diseases of Grape Plant Using Opposite Color Local Binary Pattern Feature and Machine Learning for Automated Decision Support System. It performs exact classification. In any case, exactness improves just when the testing and preparing proportion increments. The disease in grape plant is grouped utilizing Multiclass SVM in this paper. In this paper a solitary leaf is given as contribution for performing division and examinations it. Investigation is done through high pass channel. By this the diseases some portion of leaf is distinguished. The portioned leaf surface is recovered utilizing fractal-based surface element which are locally invariant in nature and in this manner gives a decent surface module. The removed surface example is then grouped to prepare Multiclass SVM classifiers into healthy or diseased classes separately. In this paper significant disease usually saw in Grape's plant, for example, Downy Mildew, Powdery Mildew, Black decay, and so on are contemplated to do the analysis. Trial results shows that coordination of image preparing methods with DSS utilizing multiclass SVM surrenders exactness to 96.66% for grape plant disease classification. The precision of the framework can be additionally improved by improving the preparation proportion. The motivation behind this work is to give a mechanized Decision Support System (DSS) to perform classification among healthy and diseased leaf effectively too effectively accessible for Farmers.

Balasubramanian Vijaya Lakshmi.et al [6] utilizes FRVM-Accuracy, FRVM-Sensitivity, FRVM-Specificity strategies in Kernel based PSO and FRVM: A programmed plant leaf type location utilizing surface, shape, and shading highlights. It produces better precision, not touchy to commotion. It lessens time intricacy and there is no constraint in speed and size. Be that as it may, the division is a difficult errand. Leaves with like shape and size are hard to group. Trouble to arrange leaves with confounded backgrounds. The fundamental point of the proposed classification procedure is to order the kind of leaf. There are a few issues during this examination work, which is recorded as beneath: (i) The leaves with the conventional design and comparative shapes are difficult to classify. (ii) If the leaves are influenced by shadow or any disease, which change the shade of the leaf is hard to group the kind of leaf. The leaf with confounded foundations is difficult to distinguish. The proposed framework can stretch out to characterize the disease of the plant leaf. It very well may be finished by deciding the impacts of adding numerous leaf features and furthermore stretch out the proposed work to order the medicinal leaf images, which is utilized in the Ayurvedic prescriptions for relieving the human diseases.

Kaur et al [7] fostered a system for the identification and classification of plant leaf and stem diseases. Neural network is utilized here. This technique is exact and extremely viable in leaf disease acknowledgment, likewise it diminishes the computational intricacy. Be that as it may, it kills the force surface highlights. For leaf and steam recognition an image-preparing based methodology is done here. Here five diseases influenced on plants are tried. They are: Early burn, Cottony form, gray shape, late sear, small whiteness. The proposed approach is image handling based. In this model at first, the images in dataset are fragmented utilizing the K-Means strategy, in the second step the portioned images are gone through a pretrained neural organization. The arrangement of leaf images taken from AlGhor region in Jordan is utilized in this model. This model fundamentally supports exact and programmed recognition of leaf diseases.

Dheeb Al Bashish. et al [8] worked in the Detection and Classification of Plant Leaf Diseases by utilizing Deep Learning Algorithm. Assurance of harvest in natural mixtures is an unpredictable matter. In this framework concentrated profound learning models were created, based on explicit convolutional neural organizations structures, for the identification of plant diseases through leaves images of healthy or diseased plants. Quicker R-CNN is utilized here for object acknowledgment and its Region Proposal Network. The Object location framework called Faster R-CNN has two modules in it. In this model the principal module is a profound completely associated convolutional neural organization. This CNN proposes districts. For preparing reason, the framework considers secures containing an article or not, based on the Intersection-over-Union (IoU) between the item proposition and the ground-truth. Then, at that point the subsequent module is the Fast R-CNN finder that utilizes the proposed locales. Box proposition technique is utilized in this model to trim the highlights. Likewise, same moderate component map which are hence taken care of to the rest of the element extractor. This is done to foresee a class and class-explicit box refinement for every proposition. Utilizing a solitary brought together organization, the whole interaction is occurring in this model. This model permits the framework to share full-image convolutional features with the identification organization, subsequently empowering almost without cost area proposition.

Robert G. de Luna.et al [9] done identification of Philippine natural medication plant leaf utilizing fake neural organization. It has better exactness and straightforward execution. Yet, the computational time is too enormous. It primarily center around the medicinal home-grown plant in Philippine as it were. Diseases in natural leaves like akapulko, amplaya and so on

Canister Liu et al [10] done identification of apple leaf diseases based on profound convolutional neural organizations. Exactness, strength are the upsides of this paper alongside the counteraction of over fitting and high element extraction capacity. The trouble in recognizing structure of the model is a major issue. This paper has proposed a novel profound convolutional neural organization model to precisely distinguish apple leaf diseases, which can consequently find the discriminative features of leaf diseases and empower an end-to-end learning pipeline with high exactness. A sum of 13,689 images were created by image handling advances. The outcomes are acceptable, and this proposed model can acquire an acknowledgment precision of 97.62%. This rate is higher than the acknowledgment capacities of different models. The proposed model diminishes the quantity of boundaries significantly, has a quicker assembly rate when contrasted with the standard AlexNet model. Exact identification of the four normal sorts of apple leaf diseases can be distinguished utilizing this CNN model. This model has a high precision and gives an achievable answer for identification and acknowledgment of apple leaf diseases. Likewise, because of the limitation of natural development laws and the flow season in which the apple leaves have fallen, different diseases of apple leaves are difficult to gather. In future work, for identifying apple leaf diseases continuously, other profound neural organization models, like Faster RCNN (Regions with Convolutional Neural Network), YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector), are intended to be applied. Moreover, more sorts of apple leaf diseases and a huge number of high-quality regular images of apple leaf diseases should be assembled in the ranch to recognize more diseases in an ideal and exact way.

Santanu Phadikar et al [11] recognized rice diseases utilizing design acknowledgment strategies. Self-Organizing Map (SOM) method is utilized which has basic and computational productivity. However, the image change in recurrence space does not offer better classification. It for the most part centers around rice disease identification utilizing design acknowledgment. In the paper, train images are acquired by separating features of the contaminated pieces of the leave. Here four distinct kinds of images are applied for testing purposes. Utilizing straightforward computationally effective procedures this model concentrates highlights.

III. Proposed Methodology

3.1 Dataset Used

The images in Basil/Tulsi Plant Leaf, information bases were caught under controlled climate from the region of Ahmednagar and Amravati, yet every data set is varied as far as the quantity of images and the sort of species utilized. Dataset has been downloaded from [12]

3.2 Classification Results on the Dataset

The classification performed on this dataset is extremely difficult because of the around comparable leaves structures for each plant species in the dataset. Notwithstanding the similitudes inside the images, the analysts figured out how to acquire 82.1% classification precision by applying various features to various leaf types in the classification cycle. The features addressing lobed leaves were extricated utilizing Local Edge Orientation Histogram (LEOH), Hough Histogram and Directional section histogram (DFH), while for non-lobed leave, Edge Orientation Histogram (EOH), Hough and DFH were utilized to separate highlights. In the meantime, the creators utilized Hough Histogram and DFH to extricate features addressing compound leaves. LEOH is a descriptor where an image is depicted utilizing directional difference in shading, though Hough histogram performs gathering of edge focuses through casting a ballot over a bunch of defined image objects to beat flawed pieces of an item inside an image.

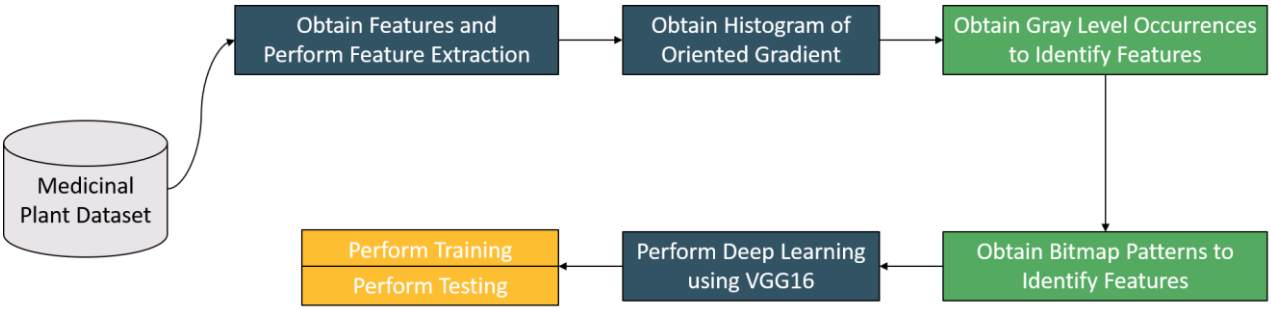

Figure 1 Proposed Methodology

3.3 Leaf Identification

A definitive objective of performing leaf discovery is to decide the presence and area of leaves in images. The location cycle includes image pre-handling, image division, image improvement and limitation. The execution of pre-preparing is pivotal to considerably lessen calculation time and commotion, which along these lines prompts higher exactness. This is demonstrated as the process exhibited 99% identification rate when identifying the upgraded Ficus Deltoidea leaf images utilizing Principal Component Analysis (PCA). The creators utilized 345 images from 5 varieties of Ficus Deltoidea. In another model, Gao and Lin (2018) accomplished 99% recognition rate also however the creators applied OTSU technique, which is an ideal division approach that actions contiguous checked edge point of leaf precisely in pre-handling stage. The OTSU strategy receives most extreme between-class difference as basis and chooses ideal division limit. We utilized two datasets where the first dataset comprises of 440 images containing 88 unique species and the second dataset has 232 images including 70 distinct species.

Figure 2 Some Examples of Leaf Images used in Dataset

3.4 Feature Extraction

In machine learning, highlight extraction capacities to lessen the dimensionality of information and to work with the classification interaction (Kumar and Bhatia 2014). For the examination on plant identification, plant features like shape, surface and shade of the leaves, natural products, blossoms and barks are regularly used to remove new dimensionality decreased features that best address plant. The plant features permit analysts to separate between various stages or classes of plants to even precisely recognizing the types of plant. Among all the plant highlights, leaves, particularly the state of the leaves are frequently utilized in plant acknowledgment research in correlation with blossoms, leafy foods. This is because of the accessibility of leaves consistently, a lot simpler to gather and protect contrasted with other plant organs. We in this manner proposed a visual consistency-based element extraction technique to separate significant highlights, in particular viewpoint proportion, vertical erraticism, rectangularity, convexity, even balance and shape intricacy to address leaf lamina.

We additionally performed leaf pivot, where the leaf must be turned to a specific direction utilizing Inertia Axis technique. This is to emulate human propensity while noticing an article. The discoveries incorporate features from an equivalent plant show a high consistency with irrelevant uniqueness in laminas, and among every one of the highlights, viewpoint proportion, vertical whimsy and shape intricacy show huge divergence in laminas from various plants.

In the interim, we endeavored to extricate mathematical features like the longest breadth, width, length, angle proportion and structure factor of the leaf in images. Furthermore, the creators removed morphological features from leaf like smooth factor, rectangularity, thin factor, edge proportion of breadth, edge proportion of physiological length, physiological width, and vein highlights. We have figured out how to accomplish 88.33% exactness when arranging plant utilizing extricated features.

3.5 Classification based on Texture

Leaf identification based on leaf surface has acquired extensive interest and this is because of the difficulties and intricacy experienced when recognizing leaf utilizing surface data. One of the difficulties is the variety of leaf surface that could be smooth, furrowed, warty, creased, notched, or even covered with trichomes (Cope and Muenscher 2001). The exploration that used surface as the element for classification lately have shown extraordinary precision. We figured out how to accomplish 98.7% precision when arranging leaf based on leaf surface utilizing Learning Vector Quantization (LVQ) and Radial Basis Function (RBF). LVQ calculation permits client to pick the quantity of preparing occasions and it likewise ready to characterize successfully from a little dataset. In the meantime, RBF reestablishes missing information inside an image by assessing multivariable capacities through a straight mix of univariate capacities.

3.6 Classification based on Multiple Features

Usage of numerous features for leaf or plant identification has been seen as a superior option when contrasted with utilizing a solitary element to distinguish plant. Numerous features are needed to have the option to separate different classifications of plants. We utilized Probabilistic Neural Network (PNN) to arrange 30 types of medicinal plants based on morphology, shape, surface and a mix of these highlights. The creators guaranteed that the classification utilizing numerous features has extraordinary potential in accomplishing high exactness even though the most noteworthy precision achieved in their analysis is just 74.67%.

Moreover, we could likewise sum up that shape is a transcendent element and essential for leaf identification. We endeavored to order 6 kinds of Indian medicinal plants from 63 images utilizing Artificial Neural Network (ANN) classifier. The features utilized for classification are the blend of smallness, whimsy, viewpoint proportion, Hu minutes, RGB and surface data. The different features considered have added to extensively high classification precision, which is 94.4%. The classification performed on 54 images utilizing ANN and VGG16 showed that both the classifiers have equivalent execution.

IV. CONCLUSION

The identification of medicinal plants physically requires a lot of work hours, and the cycle is helpless to human blunder. For these issues, programmed plant identification could be an answer however the programmed identification framework advancement requires a high number of assets which incorporates a huge information base, profound information on morphology of the plants and PC programming abilities. Right now, a large portion of the examination on programmed plant identification frameworks are tried with set up datasets that were created under controlled climate. Henceforth, more exploration on images different brightening conditions and complex foundation ought to be performed. Aside from that, the size of the dataset ought to be significantly huge to consider better preparing. This would permit the created identification framework to be more exact. Based on the survey, it is additionally discovered that identification utilizing different highlights like shape, shading and surface would likewise altogether influence the precision of the classifier. Accomplishing higher precision could affect the advancement in the use of medicinal plants in the clinical area and improving the programmed plant identification framework would quite influence the protection and safeguarding of our current circumstance.

REFERENCES

[1] Sandika Biswas, Bhushan Jagyasi, Bir Pal Singhy and Mehi Lal, “Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment,” IEEE Canada International Humanitarian Technology Conference - (IHTC), 2014.

[2] A. A. Joshi and B. D. Jadhav, “Monitoring and Controlling Rice Diseases Using Image Processing Techniques,” 2016 International Conference on Computing, Analytics and Security Trends (CAST) College of Engineering Pune, Dec 2016.

[3] John William Orillo, Gideon Joseph Emperador, Mark Geocel Gasgonia, Marifel Parpan, and Jessica Yang, “Rice plant nitrogen level assessment through image processing using artificial neural network,” IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), pp. 1-6. 2014.

[4] Pooja Pawar, Varsha Turkar, and Pravin Patil, “Cucumber disease detection using artificial neural network,” IEEE International Conference on Inventive Computation Technologies (ICICT), vol. 3, pp. 1-5, 2016.

[5] Harshal Waghmare, Radha Kokare and Yogesh Dandawate, “Detection and Classification of Diseases of Grape Plant Using Opposite Colour Local Binary Pattern Feature and Machine Learning for Automated Decision Support System,” 3rd International Conference on Signal Processing and Integrated Networks (SPIN), 2016.

[6] Balasubramanian Vijayalakshmi and Vasudev Mohan, “Kernel based PSO and FRVM: An automatic plant leaf type detection using texture, shape and colour features,” Computer and Electronics in Agriculture, vol. 125, pp. 99- 112, 2016.

[7] Kaur, Iqbaldeep, Gifty Aggarwal, and Amit Verma, “Detection and Classification of Disease Affected Region of Plant Leaves using Image Processing Technique,” Indian Journal of Science and Technology, vol. 9, no. 48, 2016.

[8] Dheeb Al Bashish, Malik Braik, and Sulieman Bani Ahmad, “A Framework for Detection and Classification of Plant Leaf and Stem Diseases, “International Conference on Signal and Image Processing”, 2010.

[9] Robert G. de Luna, Renann G. Baldovino, E. A. Cotoco, Anton Louise P. de Ocampo, C. Ira, Valenzuela, A. B. Culaba, and E. P. Dadios Gokongwei, “Identification of philippine herbal medicine plant leaf using artificial neural network,” 9th IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), pp. 1-8, 2017.

[10] Bin Liu, Yun Zhang, Dong Jian He and Yuxiang Li, “Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks,” Symmetry, vol. 10, no. 11, 2017.

[11] Santanu Phadikar and Jaya Sil, “Rice disease identification using pattern recognition techniques,” 11th IEEE International Conference on Computer and Information Technology, pp. 420-423, 2008.

[12] Mrs. Disha Sushant Wankhede, Dr. Selvarani Rangasamy, “Leaves: India’s Most Famous Basil Plant Leaves Quality Dataset”, IEEE Dataport, 2020, DOI: https://dx.doi.org/10.21227/a4f6-4413.