International Journal of Advance and Innovative Research
Volume 9, Issue 4 October - December 2022
ISSN 2394 - 7780

A SURVEY ON DIFFERENT METHODS FOR MEDICINAL PLANTS IDENTIFICATION AND CLASSIFICATION SYSTEM

Maibam Maikel Singh and Thounaojam Rupachandra Singh
Department of Computer Science, Manipur University, Indo Myanmar Road, Canchipur, Manipur – 795003

ABSTRACT
Medicinal herbs are getting popular in the pharma industry because they have minimal side effects and become less expensive than contemporary pharmaceuticals. Several people have indicated a strong interest in the topic of automated medicinal plant identification as a result of these findings. There are numerous ways to make headway toward creating a robust classifier that can consistently identify therapeutic plants on an actual justification. This study discusses the efficacy and dependability of several machine learning methods that have been used in past few years to classify plants using images of their leaves. Additionally, analyses of their benefits and drawbacks are provided. The paper offers the image processing algorithms used to recognise leaves and recover important leaf properties for a few machine learning methods. These machine learning algorithms' efficiency in classifying leaf images based on common plant characteristics, such as form, vein, texture, and a combination of several other factors, is divided into three categories. The paper concludes with an overview of recent research and areas for future development in this area before looking at the publicly released leaf datasets for automated plant recognition.

Keywords: Medicinal herbs, Therapeutic plants, Image processing and Plant Images.

1. INTRODUCTION
Although healthcare technology has been created in large quantities for use in therapy, most developing nations now prefer traditional treatment due to the shortcomings of synthetic cannabinoids in the management of chronic disease. According to, medicinal plants make about a quarter of all internationally sold prescriptive drugs. Conventional medicines are frequently used in pharmaceutical firms. This is due to the benefits of medicinal herbs, which have fewer side effects and are significantly less expensive than produced drugs. Additionally, bioactive substances derived from herbal remedies such as phenolics, arytenoids, anthocyanins, and tocopherols function as antioxidants, anti-allergens, anti-inflammatory, antibacterial, and anti-hepatotoxic substances. [17,20]. Physically identifying herbal remedies, like other plant recognition, is challenging and time-consuming, but this is because there aren't enough expert opinions. Experts created several independent herb or foliage detection methods in order to overcome this situation, the bulk of which incorporated machine learning strategies. A branch of ai technologies called machine learning allows machines to recognise objects and draw conclusions without much to no human involvement. Machine learning has provided good detection, projection, and filtration results on a range of problems, including clinical diagnosis, financial analysis, predicting maintenance, and picture identification. Four categories of machine learning techniques are currently available: supervised, unsupervised, semi-supervised, and other. [11,14]. In supervised learning, the system makes decisions based on annotated input data, and the learning process is repeated until the model reaches the highest level of accuracy. Machine learning methods that can be trained utilising no labelled data are referred to as unsupervised learning. Some situations call for semi-supervised learning, where the algorithms are operated on both labelled and unlabelled data. This study discusses various promising and reliable machine learning methods that have lately been applied to classify plants or leaves. Additionally, research is done to weigh their advantages and disadvantages in order to decide how to improve future research[1]. This research project is divided into five sections: Section I examines the significance of systems for identifying medicinal plants; Section II reviews conventional methods for identifying leaves; Section III compares research strategies and discusses their advantages and disadvantages; Section IV examines conclusions from recent studies on plant identification; and Section V offers suggestions for plants identification's future advancement. The conclusion and work intended for the future are discussed in Section VII.

2. LITERATURE REVIEW
The numerous methods for classifying and recognising native herbs are described in this paper. In order to circumvent the issues that traditional classification algorithms have with identifying therapeutic herbs,

Kan et al [2017] suggested using photographs of plants leaves as the basis for an automated categorization strategy. The method first pre-processes photos of medicinal herb leaves, then generates 10 shape features (SF) and 5 texture features (TF), and last classifies medicinal herb leaves using a support vector machine (SVM). Twelve different medicinal foliar photos were classified using the classifier method, with a mean positive outcome of 93.3%. The results show that categorising medicinal plants automatically is doable by using multi-

ISSN 2394 - 7780

feature extraction from leaf images in conjunction with SVM. The article offers a helpful theoretical framework for the research and development of classification models for medicinal plants.

Alimboyong et al. [2018] proposed a technique based on computer vision to identify the homeopathic herb species found in India's Western Ghats. The proposed method employs a combination of SURF and HOG features extracted from leaf pictures, and also a classification using a k-NN classifier. Researchers have also produced results that seem acceptable for creating apps for practical usage.

Prasad and Singh [2017] built an information sharing method employing deep qualities to explain the actual leaf images, from object identification to plant genetic studies. Studies have proven that such in-depth characteristics work better at identifying plant species than the standard methods. The study showed a brand-new, efficient approach for collecting leaves. The VGG-16 feature map is created by translating the picture into a lab colour space that is free of any specific technology. This feature map is re-projected to PCA subspace to improve species identification efficiency. For the purpose of proving durability, the study uses two different kinds of plant leaves.

To extract features from images, Turkoglu and Hanbay [2019] used image processing techniques like colour, vein properties, Fourier Descriptors (FD), and Gray-Level Co-occurrence Matrix (GLCM) methods. Research suggests using features gathered from leaves that have been cut into two or four sections rather than gathering attributes for the complete leaf. The effectiveness of each attribute extortion strategy is calculated separately and collectively using the Extreme Learning Machines (ELM) classifier. The suggested strategy was examined using the Flavia leaf database. The effectiveness of the suggested method was assessed using 10-fold cross-validation, which was then compared and tallied with methods from other works. The outcomes of the suggested strategy were assessed to be 99.10% on the Fluvial Leaf Database as well.

According to textural qualities, Naresh and Nagendraswamy [2016] suggested a symbolic system for identifying plant leaves. Modified local binary patterns (MLBP) have been proposed to recover textural data from plant leaves. Same types of plants can have varying leaf textures depending on the plant's maturity, age, and habitat. As a result, the clustering concept is used to select a large number of class representatives, and intra-cluster variations are tracked by interim valued kind metaphorical characteristics. The categorisation is made simpler by using a simple nearest neighbour classifier. The newly formed UoM Medicinal Plant Database, along with the publicly available Flavia, Foliage, and Swedish plant leaf databases, were used extensively for research. Values obtained using the proposed design are compared to results obtained using existing techniques. Even on this artificial database, studies using the Outex database have produced interesting results.

wei Tan, et al [2018] D-Leaf suggested a cutting-edge CNN-based strategy. To pre-process the leaf images and extract the properties, three different Convolutional Neural Network (CNN) algorithms—pre-trained AlexNet, fine-tuned AlexNet, and D-Leaf—were employed. The Support Vector Machine (SVM), Artificial Neural Network (ANN), k-Nearest-Neighbor (k-NN), Nave-Bayes (NB), and CNN machine learning methods were used to categorise these qualities. For the benchmark, morphological characteristics were calculated using a conventional morphometric method based on Sobel segmented veins. The D-Leaf model achieved a testing accuracy of 94.88% in compared to the AlexNet (93.26%) and fine-tuned AlexNet (95.54%) modelsAdditionally, CNN models beat conventional morphometric evaluations (66.55%). The properties obtained using CNN are discovered to follow the ANN classifier model well.

Grinblat et al. [2016] suggested using a deep convolutional neural network (CNN) In order to solve the problem of recognising plants from their leaf vein patterns. The three unique species of legumes—white bean, red bean, and soybean—are the writers' main areas of attention. Modern pipelines frequently use CNNs, which do away with the requirement for specially designed feature extractors. Furthermore, this deep learning model significantly boosts the effectiveness of the pipeline indicated before. shows that the design level achieves the claimed accuracy better. Finally, the created designs can be examined using a simple visualisation tool to identify intriguing texture features.

Ghasab et al. [2015] suggested a smart approach that employs ant colony optimization (ACO) as an attribute decision-making mechanism in order to identify different herb species from their leaf photos. For the purpose of identifying specific species, the ACO approach is utilised to delve deeper into the attribute search space and find the most discriminating features. The leaf pictures are used to extract a number of potential attributes, such as shape, morphology, texture, and colour, in order to build a features search space. The selected features are used by the support vector machine (SVM) to classify the species. To show the system's performance, about 2050 leaf photos were collected from the FCA and Flavia plant collections. The ACO-based method had an accuracy rate of 95.53%, according to the study's , demonstrating the viability of using the recommended approach for classifying various plant species.

Siravenha and Carvalho [2015] suggested a method for categorising and recognizing plants based on the shape of their leaves that looks at the contour-centroid distance's discriminating power in the Fourier frequency domain while ensuring some invariance (like rotation and scale). The effect of feature selection methods on categorization accuracy is also looked at. A feed forward neural network and a group of feature vectors in the principal components space were combined to achieve the accuracy of 97.45%.

Based on the geometry of their leaves, Zhao et al. [2015] suggested a novel method for classifying plants. Contrary to other studies that concentrated on simple leaves, the recommended approach can correctly detect both simple and compound leaves. Introduce a brand-new feature that keeps track of both global and local shape information separately so that they can be analysed independently while being categorised. Additionally, they assert that when comparing two leaf people, it is preferable to "count" the number of various shape patterns rather than precisely replicate the extorted shape features. The proposed counting-based shape identifier in addition to being efficient in terms of computation and storage, but it also provides discrimination for categorization. Studies on five separate datasets of leaf images demonstrate that the system outperforms cutting-edge approaches in terms of accuracy rate, performance, and storage needs.

Sabu et al. [2017] suggested a computer vision approach for identifying ayurvedic medicinal plant species found in India's Western Ghats. The proposed method employs a set of SURF and HOG features extracted from leaf pictures, as well as a classification using a k-NN classifier. Investigations have produced results that seem to be sufficient for creating apps with practical applications.

Priyankara and Withanage [2015] developed a leaf image-based plant classification approach by combining SIFT features with a Bag of Word (BOW) framework and a Support Vector Machine (SVM) classifier. The algorithm's accuracy rate was 96.48% after training on 20 different species. According to the results, an Android application was developed that communicates with the server and enables users to search for plant species via images of plant leaves that were taken using a smartphone.

Ekerolu and Nan [2016] developed a smart identification model that recognises and categorises 27 different types of leaves using a back propagation neural network. With a classification result of 97.2%, the findings show that the designed system performs better than fresh studies.

Elhariri, et al. [2014] developed a categorization strategy relying on Random Forests (RF) and Linear Discriminate Analysis (LDA) techniques for detecting distinct types of plants. The proposed approach consists of three stages: pre-processing, feature extraction, and categorization. Since many different plant species have distinctive leaves, the categorization method used in this research is focused on the herbs leaves. Leaves can be distinguished from each other by characteristics such shape, colour, texture, and border. The Uci Repository provided the database for these testing, which includes a variety of plant species with up to 340 picture data per species. This was applied to both the training and testing datasets using 10-fold cross-validation. Form, first order texture, the Gray Level Co-occurrence Matrix (GLCM), HSV colour moments, and vein features were combined to provide LDA a classification performance of 92.65% compared to RF's 88.82%.

Lee, et al. (2017) discovered From raw forms of input data, Convolutional Neural Networks (CNN) are employed to extract useful leaf qualities, and a Deconvolutional Network (DN) technique is utilised to learn more about the chosen features. Give a few unexpected results to consider: (2) Examine multi-level representation in leaf data, illuminating the hierarchical conversion of features from lower to higher level of abstraction, corresponding to species categories. (1) Different orders of venation are superior displayed features contrasting with outline shape. Show that the results are consistent with the hierarchical botanical descriptions of leaf characteristics. The findings yield insights into the creation of unique hybrid attribute extraction algorithms that have the potential to further enhance the classification accuracy of plant classification techniques.

Table 1: Conclusions drawn from the Current Research

Authors	Methods	Merits	Demerits
Elhariri, et al [2014]	Random Forests	Accurate results	Ineffective for real-time predictions
Priyankara and Withanage [2015]	Support Vector Machine(SVM)	Produces improved recognition performance	Does not suitable for large data

 ISSN 2394 - 7780

Siravenha and Carvalho [2015]	Feed forward neural network	Achieved higher accuracy	Approximation of a solution
Zhao et al [2015]	Counting-based shape descriptor	Computationally fast and storage cheap.	Does not used leaf texture feature
Ghasab et al [2015]	SVM	Achieved Higher accuracy	Does not suitable for more number of data
Şekeroğlu and İnan [2016]	Back propagation neural network	Reduces the error rate	Time consuming
Grinblat et al [2016]	Convolutional neural network	obtained and understood the relevant vein patterns	Higher computational complexity
Naresh and Nagendraswamy [2016]	SNN	Produces better recognition performance	Couldn't integrate attributes of additional techniques such as flower, fruit/bark
Sabu et al [2017]	K-NN classifier	Better performance	Need to use other classifier with various feature extraction models
Lee et al [2017]	Convolutional neural network	Improved recognition	Does not tested with real time data
Prasad and Singh [2017]	PCA	Robustness	Very expensive
Kan et al [2017]	Support vector machine	classifier offers accurate results	Leaf vein features and edge features are not considered
Alimboyong et al [2018]	K-NN	classifier adequate for constructing apps for real world usage	Time consuming
wei Tan, et al [2018]	CNN	Flexible	Need to include more tropical plant species
Turkoglu and Hanbay [2019]	ELM	Higher accuracy	Need to improve performance ratios with other feature extraction models

3. INFERENCES FROM RECENT WORKS

There are thousands of tree species on the earth, the majority of which possess medicinal value. However, some are hazardous to humans, some are at risk of extinction. In addition to being an essential resource for people, plants also form the base of several food chains. In attempt to utilize and preserve plants, accurate study and classification are essential. The ability of a skilled botanist to identify unknown plants significantly relies on their natural knowledge. As a matter of fact, numerous studies have been conducted to aid in the timely identification of herbs using their physical characteristics. The algorithms developed up to this point use a higher number of operations to automate the process of automatic classification, despite the fact that the approaches are nearly similar. The steps include cleaning the collected leaves, determining their unique characteristics, categorizing the leaves, creating the dataset, training for classification, and finally evaluating the results. Although leaves are the most common technique to identify plants, an automated procedure might also use stems, flowers, petals, seeds, or the entire plant. Unwanted content may be present in the data and must be removed from the source data in order to improve the quality of the data. Additionally, the data may contain a greater number of attributes, making identification more difficult. Furthermore, categorization can benefit from advanced machine learning models.

4. SOLUTION

Plants have a significant impact on human existence and the welfare of the global populace. Food, raw materials, medicines, and other requirements are mostly obtained from them. Since then, specific plants have gained notoriety for being used as a specialised remedy for a specific disease or illness. Most people are aware of their usefulness and are eager to learn more about how to use particular plants to cure particular ailments. Until now, a variety of plants, especially those used in herbal therapy, have had a significant influence on human health all over the world. Some pre-processing models can be added employing noise reduction, edge detection, the advanced models to first identify those plants. Then, in order to save time, important feature selection will be done using optimization-based feature selection. After that, classification will be performed using better machine learning-based models.

5. RESULTS AND DISCUSSION

This section explains the experimental results of the suggested model. Here, experimental results for Swedish Leaf testing database identification utilising developed model are shown using MATLAB 2013b to implement the proposed medical plant recognition and classification model. To demonstrate the efficiency of the LBP-SVM methodology, a comparison with a number of contemporary and competitive K-NN classifiers was made. then randomly divided the database into two pieces. The first half is used for testing, and the second half is used for training. 53 examples of each species were chosen for training, and the remaining 22 photos were used for comparison testingMetrics that are Tested Recall is defined as the ratio of acquired appropriate cases to the total number of pertinent instances. Therefore, knowledge and the relevance metric are necessary for both precision and recall.

Greater recall suggests that the approach obtained numerous suitable results. Mathematically, recall is defined to be as below:

$$\text{Recall}=\frac{TP}{TP+FN} \qquad (1)$$

The proportion of acceptable instances between obtained situations is how precision is defined. Large precision demonstrates that the procedure got far more appropriate results than improper ones. Precision is defined as below:

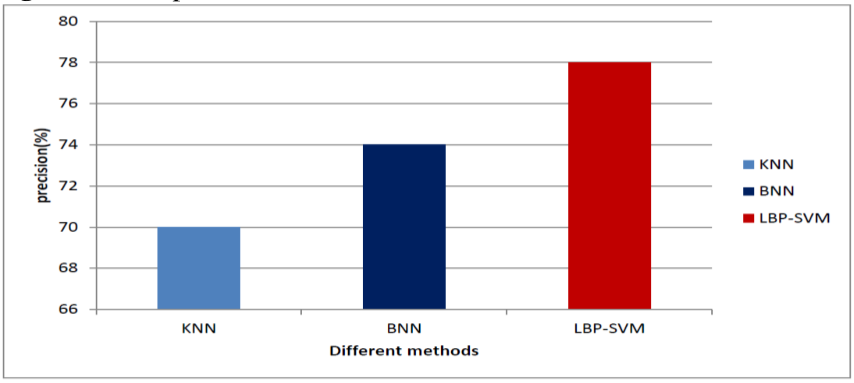
$$\text{Precision}=\frac{TP}{TP+FP} \qquad (2)$$

One parameter for assessing classification techniques is accuracy. In simpler terms, accuracy is the percentage of the forecasting method that is accurate. For two-class categorization, accuracy is computed with regard to positives and negatives as below:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \qquad (3)$$

Here TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Figure 1: Comparison of Classification methods and Results of Precision



The chart compares the suggested LBP-SVM with the current KNN and BNN in terms of accuracy performance. 1. Based on results, it is determined that the recommended LBP-SVM model offers a higher precision result of 78%, compared to the existing KNN and BNN models, which only yield 70% and 74%, respectively.

Figure 2: Recall Results vs. Classification Methods

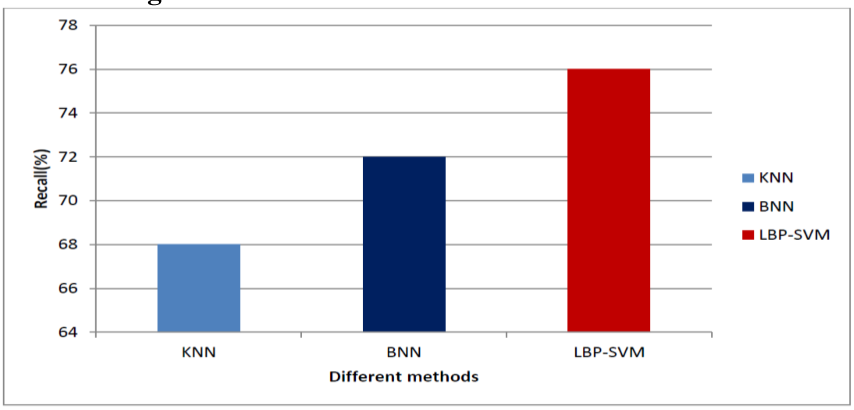
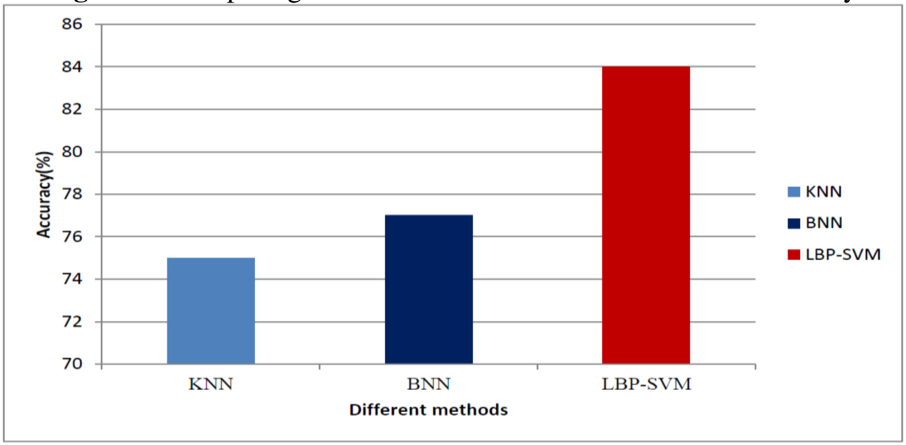


Figure 2 compares the suggested LBP-SVM to the current KNN and BNN in terms of recall performance. By conclusion, it can be said that the suggested LBP-SVM method yields greater recall outcomes (76%), whereas the conventional KNN and BNN models yield, respectively, just 68% and 72%.

Figure 3: Comparing Classification models and Results of Accuracy



Accuracy vs. Classification (Figure 3) The following figure compares suggested LBP-SVM performance to current KNN and BNN performance: 3. accuracy concerns. The suggested LBP-SVM model offers greater accuracy results of 84%, while the existing KNN and BNN models provide, respectively, only 75% and 77%.

6. CONCLUSION AND FUTURE WORK

Human survival depends on plants. Indigenous peoples have used plants in particular as traditional medicinal for a long time. Clinicians frequently identify herbs based on years of personal sensory or olfactory experience. Herb identification based on scientific data has become much simpler thanks to latest advances in testing technologies. Many people benefit from this, especially those who are not accustomed to identifying herbs. Laboratory-based research also requires knowledge in sample repair and data interpretation for time-consuming procedures. Therefore, a quick and accurate technique of detecting plants is needed. Herbal identification is anticipated to benefit from the combination of computation and statistical examination. When it comes to recognising plants, this non invasive method will be preferable, especially for people who are unable to use costly analytical technology. This study examines various techniques for identifying herbs as well as their benefits and drawbacks. Improved machine learning classifiers with some preprocessing and feature selection models will be employed in next research in the field of species identification to address accuracy-related problems and increase performance.

REFERENCES

1. Anubha Pearline, S., Sathiesh Kumar, V. and Harini, S., 2019. A study on plant recognition using conventional image processing and deep learning approaches. Journal of Intelligent & Fuzzy Systems, 36(3), pp.1997-2004.

2. Kan, H.X., Jin, L. and Zhou, F.L., 2017. Classification of medicinal plant leaf image based on multi-feature extraction. Pattern Recognition and Image Analysis, 27(3), pp.581-587.

3. Alimboyong, C.R., Hernandez, A.A. and Medina, R.P., 2018, October. Classification of plant seedling images using deep learning. In TENCON 2018-2018 IEEE Region 10 Conference, 1839-1844.

4. Prasad, S. and Singh, P.P., 2017, November. Medicinal plant leaf information extraction using deep features. In TENCON 2017-2017 IEEE Region 10 Conference (pp. 2722-2726). IEEE.

5. Turkoglu, M. and Hanbay, D., 2019. Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts. Applied Mathematics and Computation, 352, pp.1-14.

6. Naresh, Y.G. and Nagendraswamy, H.S., 2016. Classification of medicinal plants: an approach using modified LBP with symbolic representation. Neurocomputing, 173, pp.1789-1797.

7. Wei Tan, J., Chang, S.W., Abdul-Kareem, S., Yap, H.J. and Yong, K.T., 2018. Deep learning for plant species classification using leaf vein morph metric. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 17(1), pp.82-90

ISSN 2394 - 7780

8. Grinblat, G.L., Uzal, L.C., Larese, M.G. and Granitto, P.M., 2016. Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture, 127, pp.418-424.
9. Lee, S.H., Chan, C.S., Mayo, S.J. and Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 71, pp.1-13.
10. Ghasab, M.A.J., Khamis, S., Mohammad, F. and Fariman, H.J., 2015. Feature decision-making ant colony optimization system for an automated recognition of plant species. Expert Systems with Applications, 42(5), pp.2361-2370.
11. Sabu, A. and Sreekumar, K., 2017, March. Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach. In 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 145-149). IEEE.
12. Zhao, C., Chan, S.S., Cham, W.K. and Chu, L.M., 2015. Plant identification using leaf shapes—A pattern counting approach. Pattern Recognition, 48(10), pp.3203-3215.
13. Sabu, A., Sreekumar, K. and Nair, R.R., 2017, December. Recognition of ayurvedic medicinal plants from leaves: A computer vision approach. In 2017 Fourth International Conference on Image Information Processing (ICIIP) (pp. 1-5). IEEE.
14. Yigit, E., Sabanci, K., Toktas, A. and Kayabasi, A., 2019. A study on visual features of leaves in plant identification using artificial intelligence techniques. Computers and electronics in agriculture, 156, pp.369-377.
15. Siravenha, A.C.Q. and Carvalho, S.R., 2015, August. Exploring the use of leaf shape frequencies for plant classification. In 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images, 297-304.
16. Priyankara, H.C. and Withanage, D.K., 2015, April. Computer assisted plant identification system for android. In 2015 Moratuwa engineering research conference (MERCon) (pp. 148-153). IEEE.
17. Yalcin, H., 2017, August. Plant phenology recognition using deep learning: Deep-Pheno. In 2017 6th International Conference on Agro-Geoinformatics (pp. 1-5). IEEE.
18. Şekeroğlu, B. and İnan, Y., 2016. Leaves recognition system using a neural network. Procedia Computer Science, 102, pp.578-582.
19. Elhariri, E., El-Bendary, N. and Hassanien, A.E., 2014, December. Plant classification system based on leaf features. In 2014 9th International Conference on Computer Engineering & Systems (ICCES) (pp. 271-276). IEEE.
20. Bambil, D., Pistori, H., Bao, F., Weber, V., Alves, F.M., Gonçalves, E.G., de Alencar Figueiredo, L.F., Abreu, U.G., Arruda, R. and Bortolotto, I.M., 2020. Plant species identification using color learning resources, shape, texture, through machine learning and artificial neural networks. Environment Systems and Decisions, 40(4), pp.480-484.