www.revistageintec.net
ISSN: 2237-0722
revista GEINTEC
Gestão, Inovação e Tecnologias

A Survey on Different Methods for Medicinal Plants Identification and Classification System

Maibam Maikel Singh[1]; Thounaojam Rupachandra Singh[2]
[1]Department of Computer Science, Manipur University, India.
[1]maikelmaibam@manipuruniv.ac.in
[2]Department of Computer Science, Manipur University, India.
[2]rupachandrath@manipuruniv.ac.in

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's several avenues for progress in developing a strong classification that can reliably recognize medicinal herbs in real - time basis. The effectiveness and reliability of many machine learning techniques for plant categorization employing leaf pictures which have been employed in recent years are discussed in this paper. Also included are assessments of their advantages and disadvantages. For certain machine learning methods, the study provides the image processing techniques that are employed to identify leaf and retrieve significant leaf attributes. The effectiveness of these machine learning algorithms while identifying leaf images depending on typical plant properties, such as form, vein, texture, and a mixture of numerous aspects, is categorized. The leaf datasets that are publicly accessible for computerized plant identification are also examined, and the article closes with a summary of existing study and potential for improvement in this field.

Key-words: Medicinal Plants, Robust Classifier, Machine Learning, Multiple Features and Leaf Images.

1. Introduction

Medical technology is mass-produced for medical treatment, however due to the limitations of synthetic medications in treating and treating chronic illness, most first countries nowadays are turning to conventional medicine. Conventional medicines are widely employed in the pharmaceutical sector, with medicinal herbs accounting for a quarter of all internationally prescription pharmaceuticals, according to. This is owing to the advantages of medicinal herbs, which have less side effects and are far more cost efficient than manufactured medications. Moreover, bioactive chemicals isolated from medicinal herbs, like phenolics, arytenoids, anthocyanins, and tocopherols, work as antioxidants, anti-allergens, anti-inflammatory, antibacterial, and anti-hepatotoxic agents. [1,2].

However, manually recognizing medicinal herbs, like other plant identification, is difficult & time intensive, but this is owing to the lack of professional viewpoints. In response to these issues, experts developed a number of autonomous plant or leaf detection system, the majority of which used machine learning techniques. Machine learning is a subset of artificial intelligence that enables machines to recognize things and create judgments with little or no human input. On a variety of challenges, like clinical diagnosis, financial analysis, forecasting maintenance, image identification, machine learning has produced excellent recognition, forecasting, and filtration outcomes. There's several kinds of machine learning methods available today, which can be divided into four groups: supervised, unsupervised, and semi-supervised. [3,4].

The system provides judgments depending on labelled input data in supervised learning, and the learning process repeats till the predictor achieves the best accuracy. Unsupervised learning refers to machine learning techniques that could be trained using no labelled data. In certain circumstances, semi-supervised learning is required, in which the techniques are trained on labelled and unlabeled data. Multiple potential and trustworthy machine learning techniques that have recently been used for plant or leaf categorization are discussed in this paper. Furthermore, studies of their benefits and drawbacks are conducted in order to reach a decision for future study advancement [5].

In this research work, Section I analyses the importance of medicinal plants identification system, Section II provides the review on the traditional approaches that are used for leaf identification, Section III provides the comparison of the research strategies along with its benefits and drawbacks, Section IV discusses the inference from the recent research about plants identification, Section V produces the solution for future development in the plants identification area and section VI gives results of simulation. The conclusion and work intended for the future are discussed in Section VII.

2. Literature Review

This study describes various ways for identifying and classifying medicinal plants.

To overcome the problems of conventional classification algorithms in recognising medicinal plants, Kan et al [2017] suggested an automated classification approach depending on leaf images of medicinal herbs. The technique will initially preprocess medicinal plant leaf images, then calculate 10 shape features (SF) and 5 texture features (TF), and then identify medicinal plant leaves through a support vector machine (SVM) classification. The classification model was used to identify twelve various medicinal plant leaf images, with an average rate of success of 93.3 %. The findings demonstrate that utilising multi-feature extorion of leaf pictures in connection with SVM, it is possible to automatically categorize medicinal herbs. The report presents a useful conceptual framework for medicinal herbs categorization model research and development.

Alimboyong et al. [2018] suggested a computer vision method for recognizing ayurvedic medicinal plant species present in India's Western Ghats. A mixture of SURF and HOG attributes derived from leaf images, as well as a categorization utilizing a k-NN classifier, are used in the suggested approach. Studies have yielded findings that appear to be adequate for developing apps for real-world use.

Prasad and Singh [2017] developed an information exchange from object recognition to plant genetic analysis using deep attributes to describe the original plant leaf image. Such deep attributes have been shown to outperform the current in plant species identification in experiments. The study demonstrated a novel and effective leaf collection method. The image is then translated into a device-independent lab colour space, which is then utilised to construct the VGG-16 feature map. To boost the effectiveness of species identification, this feature map is re-projected to PCA subspace. The study utilizes two kinds of plant leaf collections to demonstrate the sturdiness.

Turkoglu and Hanbay [2019] employed image processing methods like colour, vein attributes, Fourier Descriptors (FD), and Gray-Level Co-occurrence Matrix (GLCM) approaches to retrieve features from images. Rather than obtaining attributes for the entire leaf, research advises using features collected from leaves separated into 2 or four pieces. Extreme Learning Machines (ELM) classifier calculates the separate and aggregate performances of every attribute extortion approach. The Flavia leaf database was used to test the proposed method. The suggested technique's efficiency was evaluated using 10-fold cross-validation, that was then contrasted and tabulated with approaches from other studies. Also on Fluvial leaf database, the suggested method's results were evaluated to be 99.10 %.

Naresh and Nagendraswamy[2016]proposed a symbolic method for classifying plant leaves depending on textural characteristics To retrieve textural information from plant leaves, modified local binary patterns (MLBP) have been suggested. Plant leaves of the similar species can have different textures according on their maturity, acquisition, and environmental factors. As a result, numerous class representatives are chosen by using clustering principle, and intra-cluster differences are recorded by interval valued kind symbolic attributes. A basic nearest neighbor classifier is used to simplify the categorization. Considerable research was conducted using the recently constructed UoM Medicinal Plant Database, and also the publicly accessible Flavia, Foliage, and Swedish plant leaf databases. The suggested methodology's values are verified to those achieved by current approaches. Investigations using the Outex database are also being explored, and the findings are encouraging even on this synthetic database.

wei Tan, et al [2018] proposed D-Leaf a novel CNN-based approach. Three distinct Convolutional Neural Network (CNN) algorithms, pre-trained AlexNet, fine-tuned AlexNet, and D-Leaf, were used to pre-process the leaf images and extort the attributes. Five machine learning algorithms were used to classify these attributes: Support Vector Machine (SVM), Artificial Neural Network (ANN), k-Nearest-Neighbor (k-NN), Nave-Bayes (NB), and CNN. For benchmark, a traditional morphometric approach depending on Sobel segmented veins was used to calculate morphological parameters. In comparison to AlexNet (93.26 %) as well as fine-tuned AlexNet (95.54 %) models, the D-Leaf model obtained a testing accuracy of 94.88 %. Furthermore, CNN models outperformed standard morphometric assessments (66.55 %). The ANN classification model is found to follow well with the characteristics recovered from CNN.

Grinblat et al [2016] proposed for the challenge of plant recognition via leaf vein patterns, a deep convolutional neural network (CNN) was used. authors focus on the three distinct legume species in specific: white bean, red bean, and soybean. The usage of a CNN, which is typical in modern pipelines, eliminates the need for customised feature extractors. Moreover, the efficiency of the mentioned pipeline is greatly improved by this deep learning model. Indicated that better the model level achieves the stated accuracy. Ultimately, interesting vein patterns can be discovered by examining the generated models using a basic visualization tool.

Lee, et al [2017] learned Convolutional Neural Networks (CNN) are used to extract valuable leaf attributes directly from unprocessed forms of input data, and a Deconvolutional Network (DN) technique is used to obtain understanding about the selected characteristics. inform a few surprising consequences: (1) varied orders of venation are superior depicted attributes contrasted with outline shape, as well as (2) examine multi-level depiction in leaf data, illustrating the hierarchical conversion of attributes from lower to upper level of abstraction, relating to species categories. Demonstrate that the findings correspond to the hierarchical botanical descriptions of leaf attributes.

The results present insight into the development of novel hybrid attribute extortion algorithms that can boost the discriminative capability of plant classification methods even more.

Ghasab et al [2015] proposed an intelligent method that provides ant colony optimization (ACO) as a attribute decision-making method to recognise distinct plant species from their leaf images. The ACO method is used to look deeper into the attribute search space to obtain greatest discriminate traits for individual species detection. A collection of possible features like shape, morphology, texture, and colour are retrieved from the leaf photos to construct a feature search space. Support vector machine (SVM) uses the chosen attributes to categorize the species. Approximately 2050 leaf images were acquired from two separate plant datasets, FCA and Flavia, to demonstrate the system's effectiveness. The report's analysis indicate that the ACO-based technique had an average accuracy of 95.53 %, indicating the feasibility of applying the suggested algorithm categorization of different plant species.

Zhao et al [2015] proposed a novel way for identifying plants based on the geometry of their leaves. Unlike previous research that has focused on simple leaves, the suggested technique can properly identify both simple and compound leaves. Suggest a new attribute that records global and local shape data separately, allowing them to be analysed separately while categorization. Moreover, they claim that while contrasting two leaf individuals, it is preferable to "count" the amount of different shape patterns instead of matching the extorted shape attributes point-by-point. Not only is the suggested counting-based shape descriptor discriminative for categorization, and it is also computationally and storage-efficient. Investigations on five different leaf picture datasets show that the system beats state-of-the-art methods with regard to recognition accuracy, performance, as well as storage requirements.

Sabu et al [2017] proposed a computer vision method for identifying ayurvedic medicinal plant species present in India's Western Ghats A collection of SURF and HOG attributes derived from leaf images, as well as a categorization utilizing a k-NN classifier, are used in the suggested approach. Investigations have yielded findings that appear to be adequate for developing apps for real-world use.

Siravenha and Carvalho [2015] proposed a technique for identifying and classifying plants depending on their leaf forms that investigates the discriminatory ability of the contour-centroid distance in the Fourier frequency domain while ensuring certain invariance (like rotation and scale). Additionally, the impact of feature selection approaches on categorization accuracy is examined. The accuracy of 97.45 % was reached by merging a collection of features vectors in the principle components space and a feed forward neural network.

Priyankara and Withanage [2015] described SIFT characteristics coupled with a Bag of Word (BOW) framework and a Support Vector Machine (SVM) classifier to create a leaf image-based plant classification method. The algorithm was trained to categorize 20 species and achieved an accuracy rate of 96.48 %. Depending on findings, an Android application was created that interacts with the server and allows people to access plant species via photos captured of plant leaves utilizing smartphone.

Şekeroğlu and İnan [2016] introduced a smart identification model that uses a back propagation neural network to recognize and identify 27 various sorts of leaves. The findings indicate that the designed system outperforms new studies with a classification results of 97.2 %.

Elhariri, et al [2014] presented for identifying various kinds of plants, a categorization strategy depending on Random Forests (RF) and Linear Discriminate Analysis (LDA) techniques was used. The suggested method is divided into three stages: pre-processing, feature extortion, as well as categorization. Because many varieties of plants have distinct leaves, the categorization method used in this study is based on the leaves of the plants. Features like shape, colour, texture, and border distinguish leaves from one another. The database for these tests was obtained from the Uci Repository comprises of a collection of diverse plant species with a maximum of 340 image data. Using 10-fold cross-validation, this was employed both for training and testing databases. With a mixture of form, first order texture, Gray Level Co-occurrence Matrix (GLCM), HSV colour moments, and vein characteristics, LDA obtained recognition rate of 92.65% versus RF's 88.82 %.

Table 1 - Implications from the Present Work

Authors	**Techniques**	**Advantages**	**Disadvantages**
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
Prasad and Singh [2017]	PCA	Robustness	Very expensive
Turkoglu and Hanbay [2019]	ELM	Higher accuracy	Need to improve performance ratios with other feature extraction models
Naresh and Nagendraswamy [2016]	SNN	Produces better recognition performance	Couldn't integrate attributes of additional techniques such as flower, fruit/bark
wei Tan, et al [2018]	CNN	Flexible	Need to include more tropical plant species
Grinblat et al [2016]	Convolutional neural network	obtained and understood the relevant vein patterns	Higher computational complexity
Lee,et al [2017]	Convolutional neural network	Improved recognition	Does not tested with real time data
Ghasab et al [2015]	SVM	Achieved Higher accuracy	Does not suitable for more number of data
Zhao et al [2015]	counting-based shape descriptor	Computationally fast and storage cheap.	Does not used leaf texture feature
Sabu et al [2017]	K-NN classifier	Better performance	Need to use other classifier with various feature extraction models
Siravenha and Carvalho [2015]	Feed forward neural network	Achived higher accuracy	Approximation of a solution
Priyankara and Withanage [2015]	SVM	Produces improved recognition performance	Does not suitable for large data
Şekeroğlu and İnan [2016]	Back propagation neural network	Reduces the error rate	Time consuming
Elhariri, et al [2014]	Random Forests	Accurate results	Ineffective for real-time predictions

3. Inferences from Recent Works

Thousands of species of plants exist on the planet, most of which has therapeutic value, others would be endangered, and yet others are dangerous to humans. Plants are not just a vital resource for humans, and they also serve as the foundation for certain food chains. It is critical for proper research and categorize plants in order to use and conserve them. Recognizing unfamiliar plants depends heavily on an experienced botanist's inherent knowledge. As a result, a large number of research has been done to assist the early detection of plants depending on physical properties. Although the methods are very identical, the algorithms produced so far use a different number of actions to automate the procedure of automatic categorization. The procedures entail preparing the gathered leaves, identifying their special properties, classifying the leaves, building the dataset, training for identification, and lastly analyzing the outcome. Though leaves are the most popular way to identify plants, stems, flowers, petals, seeds, or even the entire plant could be employed in an automated procedure. To improve quality of the data, data may contain undesired content that needs be eliminated from source data. And data might have more number of features so it will take time for identification. Finally can use improved machine learning models for classification.

4. Solution

Plants provide important contributions to human lives and play a key influence in the world population's well-being. These are a primary source of food, raw materials, pharmaceuticals, and other necessities. Ever since, certain plants have become famous for being utilized as a specialised treatment for a particular sickness or illness. The majority of people are conscious of their value and are interested in learning more about how to utilize specific plants to treat specific conditions. Till now, various plants, particularly herbal medicine plants, have had a major impact on human health all over the globe. For classifying those plants first some preprocessing models can be introduced using noise removal, edge detection and enhancement models. Then optimization based feature selection will be used for significant feature selection to reduce the time consumption. And then improved machine learning based models will be used for classification.

5. Results and Discussion

This part explains experimental outcomes of proposed model. Proposed medicinal plant identification and classification model is implemented using MATLAB 2013b. here, provides experimental findings for identifying the Swedish Leaf testing database through proposed technique. A contrast with numerous competitive and modern methodologies K-NN classifier, BNN was undertaken to illustrate the effectiveness of the LBP-SVM methodology. Then splitted the database to two parts at arbitrarily. A first half is utilised for training, while the next half is employed for testing. Selected 53 instances per species for a training and the rest 22 photographs for a testing to compare.

Experimental Metrics

Recall is described as the portion of appropriate cases, which have been acquired over the overall number of relevant instances. Both precision and recall are hence dependent on an understanding and the measure of relevance. greater recall indicate that method retrieved many appropriate outcomes. Mathematically, recall is defined to be as below:

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

Precision is defined as fraction of appropriate instances between acquired cases. Large precision proves that method retrieved considerably large appropriate outcomes compared to inappropriate ones. Precision is defined as below:

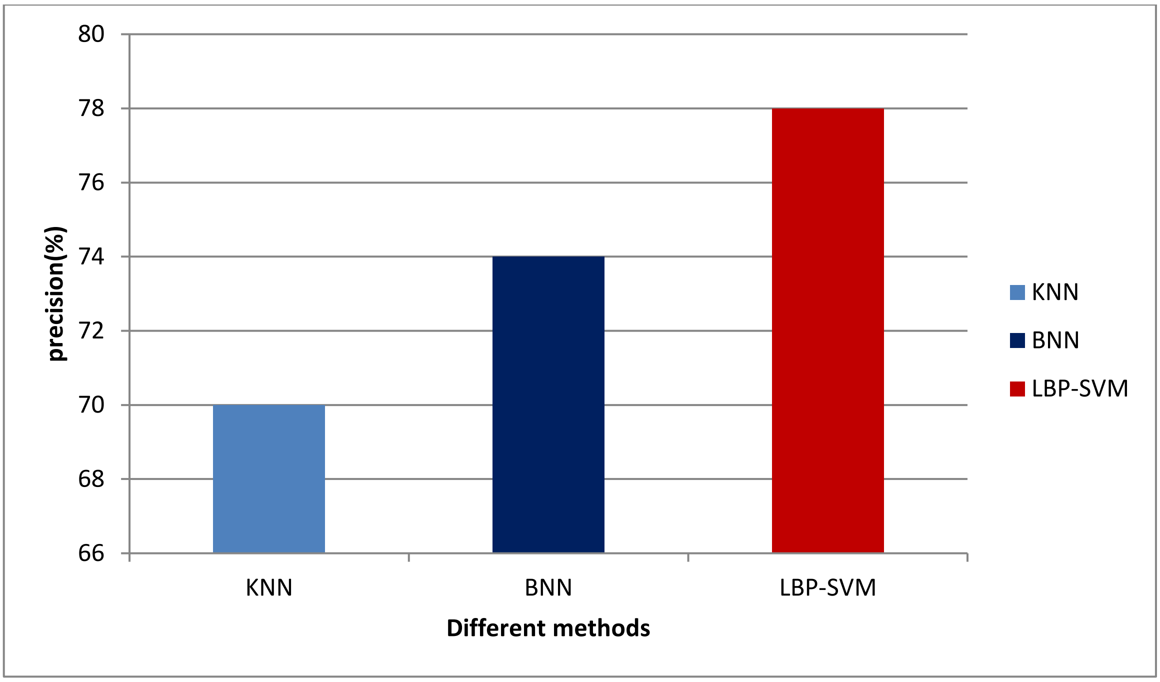
$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (2)$$

Accuracy is one metric used for the evaluation of classification models. In layman terms, accuracy is defined as the division of forecast method that is got correct. For two-class categorization, accuracy is computed with regard to positives and negatives as below:

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

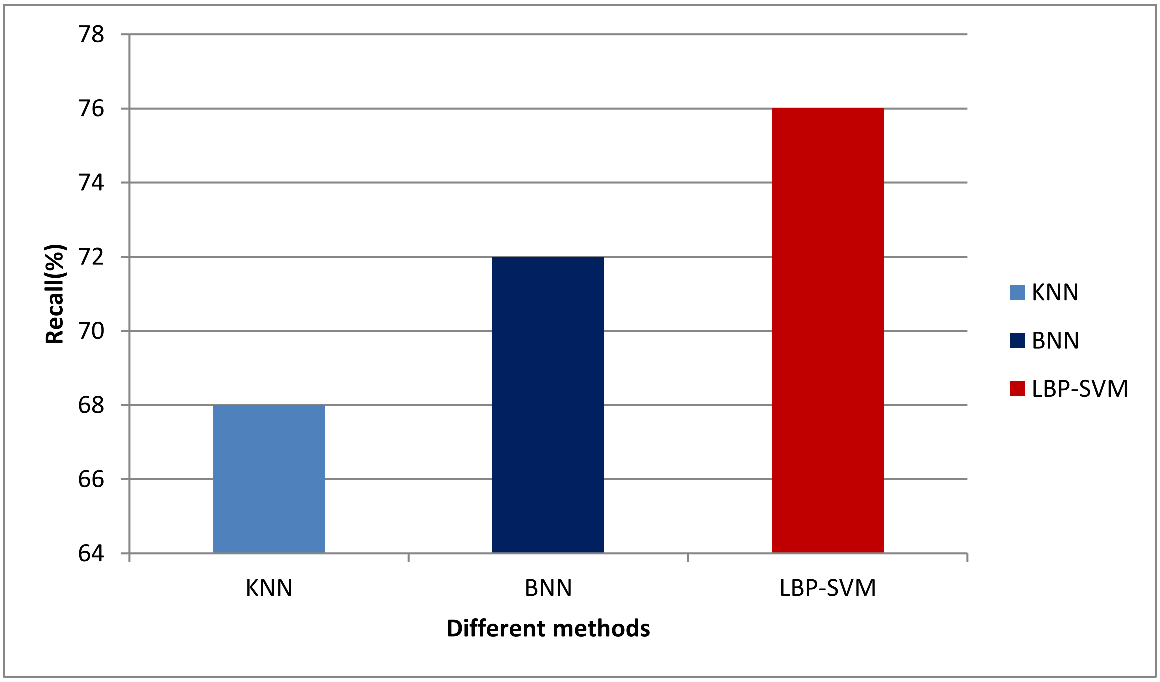
Here TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Figure 1 - Precision Results vs. Classification Methods



Accuracy performance measure contrasted among suggested LBP-SVM and present KNN, BNN are shown in figure.1. by outcomes, it is concluded that suggested LBP-SVM model provides the higher precision outcome of 78% while the existing KNN, BNN models produces only 70% and 74% accordingly.

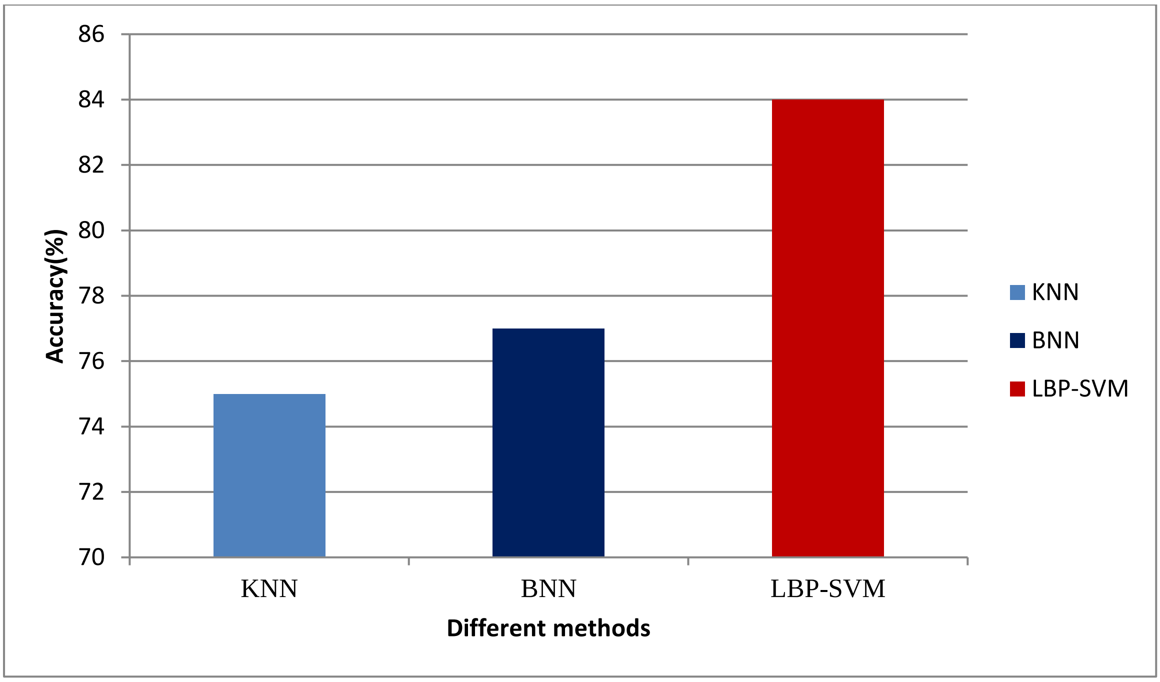
Figure 2 - Recall Results vs. Classification Methods



ISSN: 2237-0722
Vol. 11 No. 4 (2021)
Received: 11.06.2021 – Accepted: 14.07.2021

Figure:2. Shows performance measure contrast among suggested LBP-SVM as well as current KNN, BNN in terms of recall. By outcome it is concluded that suggested LBP-SVM method produces higher recall results 76% while the existing KNN, BNN models produces only 68% and 72% accordingly.

Figure 3 - Accuracy Results vs. Classification Methods



Performance measure contrasted among suggested LBP-SVM as well as present KNN, BNN are shown in figure: 3. interrns of accuracy. By outcomes it is concluded that suggested LBP-SVM model provides higher accuracy outcomes of 84% while the existing KNN, BNN models produces only 75% and 77% consequently.

6. Conclusion and Future Work

Plants are necessary for human survival. Herbs, particularly, are employed by indigenous populations as folk medicines from old period. Herbs are typically recognized by clinicians based on decades of intimate sensory or olfactory experience. Recent improvements in analytical technology have made it much easier to identify herbs depending on scientific evidence. This helps a lot of individuals, particularly those are not used to recognising herbs. additionally for time-consuming methods, laboratory-based analysis necessitates expertise in sample healing and data explanation. As a result, a simple and reliable method for identifying herbs is required. Herbal identification is anticipated to benefit from the combination of computation and statistical examination. This nondestructive technique will be the preferred approach for quickly identifying herbs, especially for individuals who can not able to use expensive analytical equipment. This work reviews about different methods for plants recognition and also reviews their advantages and disadvantages. In future research in the area of plants identification, improved machine learning classifier with some preprocessing and feature selection models will be used to solve the accuracy related issues and enhance the performance.

References

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.

Yalcin, H., 2017, August. Plant phenology recognition using deep learning: Deep-Pheno. *In 2017 6th International Conference on Agro-Geoinformatics* (pp. 1-5). IEEE.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

Şekeroğlu, B. and İnan, Y., 2016. Leaves recognition system using a neural network. *Procedia Computer Science*, 102, pp.578-582.

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.