A Group Labelled Classification Model for Accurate Medical Plant Detection Used in Drug Preparation

Prabhat Kumar Thella*, Ulagamuthalvi Venugopal

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai 600119, India

Corresponding Author Email: thellaprabhatkumar@gmail.com

ABSTRACT

The use of medical plants in the preparation of medicines has been increased in recent years. Medical plants are an essential component in the production of medicinal products. Medicines are made from root powder or plant leaves. When the herbal medicine is reduced to powder, more experience is required to determine the medicinal product through pharmacognoses. Inaccurate medical plants can cause patients serious health problems. For standardization and quality control of medical drugs the correct identification of the powder shape of medical plants is important. Medical plants are currently classified using a chemical leaf-based assessment, physical assessment and biological assessment. In medicine industry it is extremely necessary to identify the right medicinal plants for the preparation of a medicine. Its leaves form, color and texture are the key features needed to recognize a medicinal plant. In hierarchical clustering technology, the coefficient of inconsistency is used to generate natural clusters. Intra class differences can be seen with the amount of clusters obtained for plant organisms. The aggregate of the corresponding vectors of each sample of a cluster is calculated for one cluster representation. In terms of its leaf samples, therefore, the multiple members of the valued interval type are used to represent the plants in an effective way. The proposed model performs classification of leaf features using group labelled clustering model and then perform locking of labelling. This paper considers a Group Labelled Classification (GLC) Model that examines feature on the front and back of a green leaf, along with morphological characteristics, to achieve a specific optimal combination of features that optimize the recognition rate. The proposed model efficiently extracts the relevant features only from the medical leaf for accurate medical leaf detection. The proposed model is compared with the traditional methods and the results show that the proposed model performance is better.

1. INTRODUCTION

Globally, millions of plant species play a major role in human life on Earth. Plants are very important for human health and provide significant knowledge for community growth. The food chain is based on plants. Nearly 3.3 billion people daily use medicinal plants as a medicine to improve their health. Medicinal plants form the basis and are useful in treating some chronic conditions in a medicine method called Ayurveda [1]. But botanists are scarce, and a common man cannot carry out this work in due course. Climate change has led to many changes in the geographical distribution of these species [2]. This has made the medical plant recognition an important task. It can also be extremely beneficial to identify plants with useful medicinal properties for certain people daily [3].

Plants have medicinal properties that can help combat certain important diseases [4]. Only if the plants are correctly described, this can be achieved. A facility should be available to identify medicinal plants quickly. Some advanced machine learning algorithms that use the leaf characteristics can be applied to automatically detect the leaf [5]. Furthermore, the process often involves people who are not experts in the field, with the use of a computer-based method to classify plants [6]. Therefore, massive potential seems to be available in this area, and a precise plant prediction can prove to be extremely necessary and important.

In India, the herbal medicines account for 80 percent of the population. Herbal medicines contain spices, herbal materials and herbal preparations [7]. The herbal plant contains the raw material of the plant, such as leaves and flowers, berries, grains, stems, wood, bark, roots, rhizome and other plant parts [8]. In addition to herbs, herbal products include fresh juices, gums, fixed oils, essential oils, resins and herbal dry powders. Medicines are made by extracting, dividing, purifying, concentrating or other physical or biological processes of medical plants [9]. Preparations are often made in alcoholic beverages and/or honey or in other materials by steeping or heating herbal material. Herbal preparations consisting of one or more herbs are the finished herbal preparations [10]. The word "herbal product mixing" can also be used if more than one herb is used.

The leaf is so easy to collect, differentiate and capture that it is the principal basis for the identification of medicinal plants [11]. This means that it is used for the identification. Medicinal plants are classified simply and feasibly using the process of automatic leaf-image recognition [12]. The creation of a medicinal plant identification system using that methodology has important consequences for the automatic classification and computerised instructions of medicinal

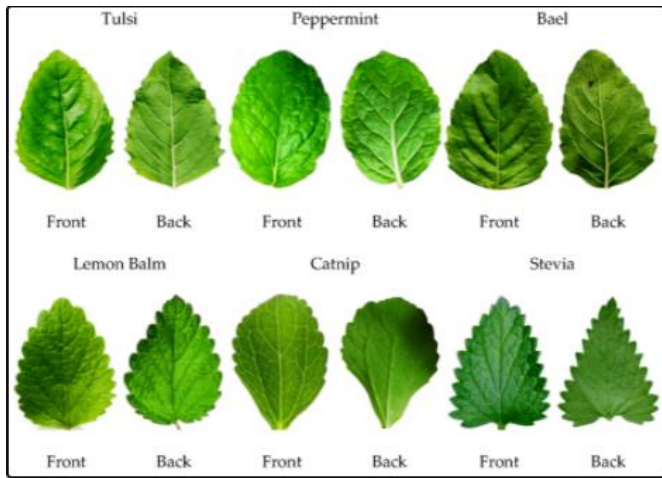plants. The leaf of a medical plant is indicated in Figure 1.

Figure 1. Medical plant leaf

It is essential to research and identify plants properly in order to use and protect plant species. Unknown plants depend heavily on an experienced botanist's expertise. A manual method based on morphological characteristics is the most effective approach for accurate and easy identification of plants [13]. Many of the processes in which these plant species are classified are also 'based on human knowledge and skills. But this manual recognition process is also time consuming and laborious. Therefore, several researchers have conducted research to support automatic plant classification on the basis of their physical properties [14].

Systems that have been built so far use different steps to simplify the automated classification process, although the processes are very similar. These phases essentially include the preparation of the collected leaves, the preparation of some pre-processing to classify their basic features, leaf classification, database compilation, training for identification and the final results assessment. The leaves, while most widely used for plant identification, can be used for automatic processing of stem, flowers, petals, seeds and even the entire plant [15]. Non-botanical experts may use an automated plant identification system to classify plant species easily and without any difficulty [16].

By using a similarity function, the resemblance between objects is determined. It is especially useful for document organisation, recovery and browsing support [17]. Data Clustering is a system where sensibly identical and physically processed information is stored together. The number of disc accesses must be reduced in order to improve the performance of the database system. In the marketing, insurance, libraries, urban planning, biology, earthquakes and www document classification, clustering algorithms can be used.

The samples taken from various plants but belonging to the same genus can also have major shape and texture fluctuations. In a class of plant leaves there are therefore major changes within the class [18]. For successful classification, it is very important to identify certain intra-class variations [19]. So, selection of multiple members for a class using the principle of hierarchy clustering by grouping related leaf samples in terms of their shape and texture is suggested in the proposed model.

Verified and uncontrolled Machine Learning Approaches are generally used for accurate clustering and classification process. These methods classify new objects into a collection of discrete class labels thus reducing the role of empirical loss [20]. However, controlled procedures involve the application of a training set which provides a priori information on several class labels of objects. Unchecked methods do not need a training set, however, which includes information from class labels of objects for the purposes of entry [21]. Unmonitored methods can detect possible new cluster structures in a dataset. In addition, if class label data is not accessible they can be introduced. Therefore, unchecked computer education is to be applied in place of supervised methods, if the purpose of the trial is to discover the class labels best describing a collection of data.

2. LITERATURE SURVEY

Scale Invariant Feature Transform (SIFT) was used, as a form descriptor and colour moments, by Rahmani et al. [1]. Each plane is broken down into 9 grids. The image is broken down into an HSV. For each grid in each plane, colour moments are calculated and used as a vector function. For identification purposes the last Euclidean distance between test sets and training sets is used. The authors' databases are developed by acquiring in a natural light 40 leaf photographs of Malaysian herbs. 87.5% precision is achieved regardless of the image scaling and rotation. When used to extract key point characteristics, SIFT is computationally intensive.

Four geometric features such as convexity, solidity, excentricities and cyclical circularity were used by Masarczyk et al. [2] in combination with three colour features of the RDB in the form of roughness and granny indexes. The archive of 1000 leaf images obtained by the authors. The speed of the recognition method is increased by a three-stage comparison of function vectors. They have reached 85 percent identifying score. In the form of a design descriptor for identification, M.R. Dileep et al. [3] used four geometric features, HU invariable moments and Polar Fourier transforms coefficients. The RGB colour plane was used as colour descriptor for the mean, standard deviation, skewedness and curtosis. The machine was 100-trained and 50 scanned images were checked. The overall efficiency of classification is 92 percent. The process is computer-intensive and does not rotate fully invariant.

Kaya et al. [5] used a geometrical characteristic aspect by considering shape factor, rectangularity, perimeter diameter ratio, solidity, convexity and irregularity as shape descriptor. The tests have been carried out on the Flavia database with a precision of 89 percent. The findings were compared separately and tabled for the output of four different classificators based on geometric and Zernike instance features. In comparison with geometric descriptors, Zernike's moments are found to provide better accuracy in all classifiers. The experiment was conducted with the Naive Bayes classifier, K-NN, support vector machine and PNN classifier. From previous work, it is clear that the plant is identified with the geometric, colour and texture characteristics of the leaf.

Tammina [9] have suggested using emerging self-organizing maps to visualise clusters in two-dimensional space of high-size biomedical data. However, it was still necessary to preselect the number of clusters and other parameters to achieve a better visualisation through their methods. Biclustering algorithms have been made prominent within the bioinformatics community in the ability to detect similar gene sets under various experimental conditions. Bhatt et al. [11] introduced one of the first works in this field. They

suggested that the algorithm be used to cluster gene expression results, an iterative greedy research need to be considered. While their approach did not involve a preliminary selection of the number of clusters, hyper parameters were needed.

Sabri et al. [13] proposes a scheme for the recognition of the leaves with a deeper convolutionary neural network. Pre-processing techniques shall be applied to remove the area from the image for the prediction of the leaves location. This procedure is performed prior to classification. The precise use of deep CNN with increase of data is defined and data increase is used to improve the model. A distinction between different methods of extracting functions is made in and also involves a combination of various features and classifiers used for the recognition process.

Within CNNs, the tested values are compared with the experimental and theoretical aspects for comparison. In the Leaf Identification using a Deep Convolutional Neural, Batvia et al. [14] proposed a similar approach which is also focused on CNN. Popular data collections of flavia and leaf datasets are used in this model. For identifying purposes, the author uses a CNN of nine layers and notes that in his experience the method presented is the best approach until now. The proposed CNN has been reported to be effective for leaf classification.

3. PROPOSED MODEL

Classification is the problem of deciding, based on a training set of data containing observations and which groups are considered to belong to a series of categories (subpopulations), of new observations. A broad variety of distance metrics can be used to test the similarity of medical leafs for accurate identification [22]. The correct distance metric is selected depending on the type of data under study (e.g., ratio, interval, ordinal, nominal or binary scale). For example, for ratio or interval data, the Euclidean distance is suitable, while for ordinal data of Manhattan scale is the distance calculated [23].

Subsequently a Group Labelled Classification algorithm is used in the system presented in this work. Hierarchical algorithms of clustering are commonly used by one connection method only, which may restrict the ability of some datasets to recognise cluster structures [24]. The leaf image numeric data is considered and its sequential representation of a pixel vector is taken to represent one-dimensional mathematical moments [25]. The pixel vector is depicted as the series v(n), defined as the Euclidean distance between the centre of mass and all points (n points) along the numeric pixels data array. The pixel vector representation is performed as:

$$A_p = \frac{1}{N} \sum_{i=1}^{N-1} \left[PV(I(i,N)) \right]^{N-i}$$
$$PV(I[i][j]) = 1 - 4\pi + \frac{2 * I(i, i+1)}{i + Th} + E^2 \tag{1}$$

The central mean of the numeric data is calculated for applying clustering of similar objects is performed as:

$$Mean(PV(I_n)) = \frac{1}{\theta} \frac{\overline{\sum_{i=1}^{} I_i + I_N}^{PV}}{Th + N} \tag{2}$$

where, P is the order of pixel vectors.

The proposed GLM model with a probability of creating n cluster sets uses group labelling of data fields. Initially the Labelling model clusters the data and then chooses appropriate features from the clusters, assigns labelling to every extracted feature in the initial level.

The labelling characteristics given are then connected to other features based on labelled vector, which are grouped to improve the degree of accuracy. If the number of items in a data set increases, the computer complexity of evaluating all possible numbers of clusters increases linearly. This can be a challenge even in parallel computing in datasets that contain numerous artefacts. The proposed model framework is depicted in Figure 2.
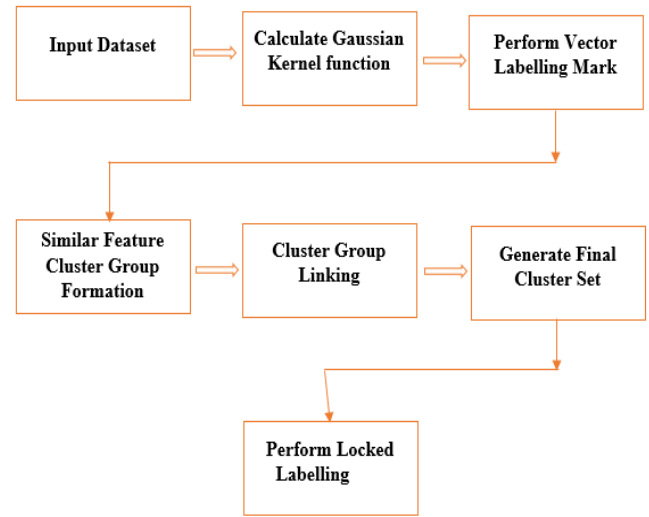

Figure 2. Proposed model architecture

The proposed model performs Group based Classification for merging of clusters with probabilistic adjacent values for reducing the numbers of total cluster set for medical plant identification by performing accurate pixel classification. The algorithm explains the proposed model operation.

Algorithm Group Labelled Clustering (GLC) Model
{
Input: Image Dataset (IS), Weight (Wg), Threshold (Th), Pixel Vector(Pv), Labelling Mark (LM)
Output: Cluster Set (CS)

In any case, the clustering occurs by comparing the values identical to the cluster group.
For each Img in range(IS(i)),

$$sim(I(i, i+1)) = \frac{\sum_{i=1}^{N} |i_v - j_v|^2}{\sum_{j=i}^{N} |j_v + i_v|^2} \tag{3}$$

where, 'sim' is the comparison difference of instances in the dataset. The distance induced by Gaussian Kernel function for pixel extraction is:

$$\|\varphi(i) - \varphi(j)\|^2 = (\varphi(i), \varphi(j), \varphi(j) - \varphi(i)) - \min(sim(I(i, j))) \tag{4}$$

Based on the similarity cases, the relevant values need to be grouped. To initialize the grouping process, labelling mark need to be performed for the pixel vectors considered. The labelling is carried out by specifying a labelling mark vector LM as:

$$LM(G(i)) = \sum_{i \in IS(N), j=i} IS(j) + th_j^l (sim(I(i) + I(j)_i)) + LM(i_j, j_i) + \mu \quad (5)$$

After the features are marked, all the features of the cluster are linked to a cluster set. The categories are calculated as:

$$L\left(G(I(i,j))\right) = \sum_{i=1}^{N_i} \sum_{i=1}^{N} |i_v - j_v|^2 \, LM(G(i) \in IS(i) * \mu_{i,j} + \sum_{i=1}^{N} \lambda_n \left(\sum_{k=1}^{c} LM(IS(i))^{-1}\right) \quad (6)$$

where, $\mu_{i,j}$ is a pixel similarity degree for the surrounding area. The pixels and the adjacent pixels are compared for the relevancy and the irrelevant features are identified and relevant features are linked as:

$$Lk(IS(n)) = \sum_{i=1}^{N} \sum_{j=i}^{N_j} L(Cg(i) + \theta T_t(Mean - A) + \frac{|IS(i+1)^j|}{Min(sim(IS(i,j))) + Th} \quad (7)$$

Here θ reflects the way the cluster groups link. After linking all the cluster groups, the labelling process is updated with the Locked Labelling (LL) for creating the cluster set that is performed as:

$$LL_{i,j}(IS(i)) = \sum_{i=0}^{N} Lk(IS(i,i+1)) + \frac{(Max(Sim(IS(i))) - Min(Sim(IS(i+1))))}{\theta * \left(\frac{N-2x+\lambda i}{2}\right)! \left(\frac{N+2x-\lambda j}{2}\right)!} - Th \quad (8)$$

The optimization technique is used for reducing the number of features considered based on the labelled mark and the leaf exact shape. The optimization is performed as:

$$Iopt_N(X,Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} LL_{i,j}{}_{i,i+1}^{N} \|\lambda(x) - \lambda(y)\|^2 + \sum_{j=i}^{M-N} Min(Lk(IS(i))) \quad (9)$$

$$Fopt(x, y, \lambda) = 2 \sum_{i=1}^{M} \sum_{j=1}^{N} Iopt_N^M (1 - G(IS(i,j))) - \left(\sum_{j=i+1}^{M+N} LL_{i,j} - min(Iopt)\right) \quad (10)$$

Finally, the Cluster set after performing multi labelling is performed as:

$$ClusterSet(DS(i)) = \sum_i \sum_j ((MLF(i) + Cg(i)) + (MLF(j) + (Cg(j) - F(j)) \quad (11)$$

$$ClusterSet(DS(i)) = \left(\frac{\lambda_1}{2}\right)^{\frac{1}{N-1}} \left[\left(\frac{1}{(1-G(i_1,j_1))}\right)^{1/M-1} + \left(\frac{1}{(1-G(j_1,i_1))}\right)^{1/N-1}\right] + Th \quad (12)$$

$$ClusterSet(DS(i)) = \left(\frac{\lambda_1}{2m}\right)^{\frac{1}{N-1}} \left[\sum_{j=1}^{N} \left(\frac{1}{(1-Fopt(i_1,j_1))}\right)^{1/M-1}\right] \quad (13)$$

$$FinalClusterSet(FCS(i)) \Rightarrow \frac{4\lambda}{\mu^2} \sum_{l=1}^{N} Fopt_n^m \left(-\exp\left(\frac{-\|i_j - j_i\|^2}{\lambda^2}\right)\right) + \left[\sum_{j=1}^{N} \left(\frac{1}{(1-Fopt(i_1,j_1))}\right)^{1/M-1}\right] \quad (14)$$

4. RESULTS

The proposed model is implemented in ANACONDA SPYDER and the dataset is considered from the link https://www.herbalgram.org/resources/expanded-commission-e/ and https://www.kaggle.com/dalipkamboj/mango-leaf-species-indian/version/1. This dataset includes thirty species of safe herbs such as the sandalwood album, Muntingia calabura, Plectranthus amboinicus / Coleus amboinicus (Indian Mint) and Brassica juncea (Eastern Mustard), as well as many others. The dataset is composed of 1600 pictures. The 100 to 200 high-quality images are represented in each species. The directories are numbered according to the name of the species. The plucked leaf is from various plants in local gardens of the same genus. After taking a picture, it is keen to avoid plugging many leaves to create the dataset. The dataset selects healthy leaves for extraction of features.

The Group Labelled Classification (GLC) Model is compared with the Hybrid Support Vector Machine and Deep Learning Neural Network (SVM-DLNN) model and the results are illustrated clearly. The proposed model is compared with the traditional method by considering parameters like Classification Accuracy Levels, Classification Time Levels, Labelling Mark Time Intervals, Similarity Identification Accuracy levels, Optimization Time Levels. This analysis used a total of 700 leaves. Twenty-Five separate medical varieties have collected a total of 32 leaves. Two datasets were used for the 25 plant types. The first data set was used for training with 70% of medical plant samples; the other dataset used for study with 30% of samples. The clusters obtained from different parameters are indicated in Table 1.

Table 1. Clusters obtained for different values of δ

Cluster Index	Sample indices forming the cluster		
	d=0.1	d=0.5	d=1
10	{1,6}	{1,7}	{4,8,12}
20	{5,12}	{2,4,6}	{3,6,9}
30	{3,6,9,11}	{3,5,7,9}	{1,2,4,6,8}
40	{1,2,6,9}	{2,5,9}	{1,3}

The different leaves datasets are considered and the overall classification performance values are represented in Table 2.

Table 2. Overall classification performance of the proposed shape description scheme for different plant leaf datasets

Leaf Dataset	Number of representatives	60:40:00 F-Measure Average for overall species	50:50:00 F-Measure Average for overall species
UoM Medicinal Plants	1126	0.7201 ± 0.0149	0.7172 ± 0.0142
	986	0.8040 ± 0.0129	0.7479 ± 0.0108
	1056	0.8722 ± 0.0063	0.8305 ± 0.0137
Flavia	967	0.7789 ± 0.0080	0.7351 ± 0.0108
	1567	0.8613 ± 0.0125	0.8189 ± 0.0119
	1278	0.8917 ± 0.0068	0.8695 ± 0.0065
Foliage	1244	0.6952 ± 0.0155	0.6907 ± 0.0114
	1835	0.7605 ± 0.0122	0.7467 ± 0.0155
	2123	0.8250 ± 0.0131	0.7977 ± 0.0103
Swedish	1398	0.8257 ± 0.0108	0.8140 ± 0.0119
	1576	0.8744 ± 0.0159	0.8612 ± 0.0117
	1498	0.9116 ± 0.0138	0.8862 ± 0.0127

Models of classification forecast category names, and models of prediction predict ongoing valued functions. Classification is the most popular data mining technique used to create a model that can identify the population of documents over a number of preclassified attributes. The method of classification includes classification and learning. In the study of training results, classification algorithms are analysed. In classification test results, the exactness of the classification rules is estimated. The rules may be extended to the new data sets if the precision is sufficient. This preclassified attribute is used in the classification-training-algorithm to establish the parameters needed for adequate discrimination. The proposed model considers relevant features and the clustering is perforemd for accurate medical plant detection. The cluster set formation is clearly indicated in Figure 3.
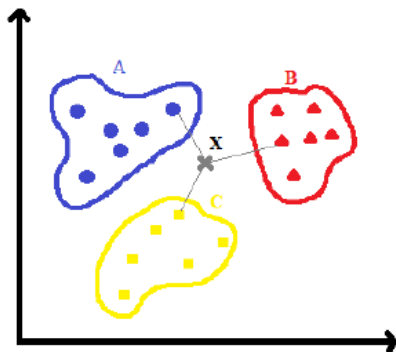

Figure 3. Clusters set

Cost-sensitive classification systems are meta-students that make their base classification cost-sensitive. In addition, several classifiers are examined in the process based on an error, which is designed to improve the classifier's accuracy. In many applications of classification, the expense of misclassification is also a significant parameter. The classification failures are equally common for all error-based classification methods, that is not the case for all applications in the real world. The classification accuracy levels of the proposed and traditional methods are indicated in Figure 4. The proposed model classification accuracy is more when compared to traditional method.
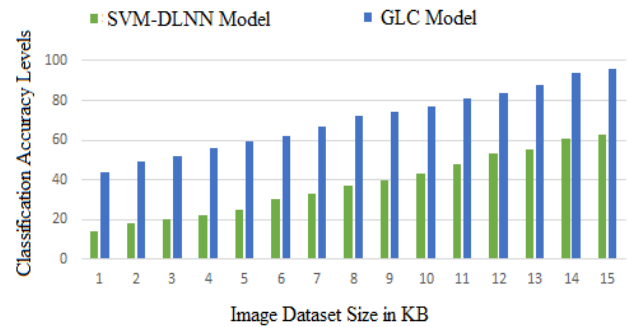

Figure 4. Classification accuracy levels

Classification is the prediction of a specific class, given that its characteristics are related to data. Each class has its own characteristics and weaknesses. One thing that is absolutely important to remember here is that ML calculations help us differentiate groups from knowledge generalisation. The classification time levels of the proposed and existing methods are clearly represented in Figure 5. The classification time levels of the proposed model is less than the existing method.
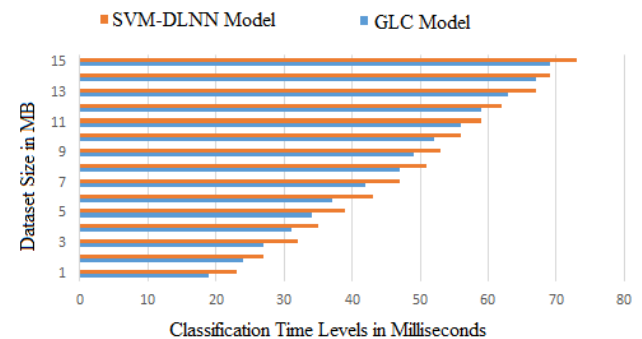

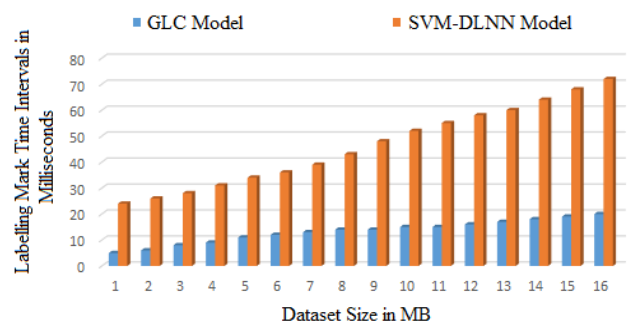**Figure 5.** Classification time levels

Figure 6. Labelling mark time intervals

The Labelling mark time intervals of the proposed and existing models are included in Figure 6. The Labelling mark time levels of the proposed model is less that improves the

system performance levels.

The identification of similarity values of the features is performed and the similarity identification accuracy values are indicated in Figure 7. The proposed model accuracy levels are high than the existing model.
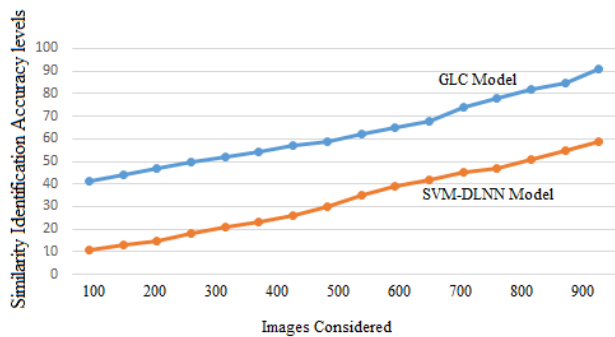

Figure 7. Similarity identification accuracy levels

The feature optimization is performed in the proposed model that reduces the cost and improves the performance levels. The optimization time levels of the proposed and existing models are indicated in Figure 8.
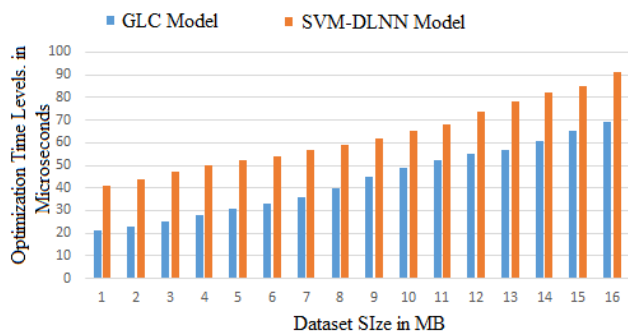

Figure 8. Optimization time levels

The overall medical plant detection accuracy levels are indicated in Figure 9. The proposed model detection rate is high than the existing model. The medical plant detection accuracy of the model is better when compared to existing methods which indicate the performance levels are also high.
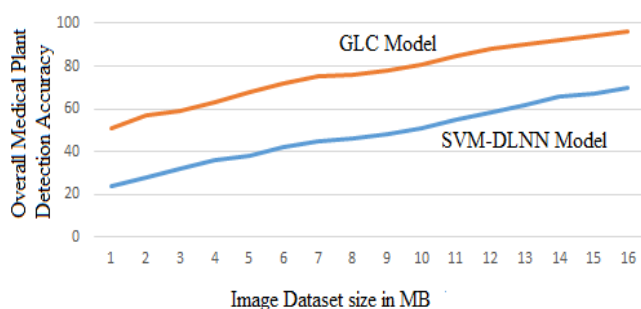

Figure 9. Overall medical plant detection accuracy

5. CONCLUSION

Machine learning methods allow information to be found in broad sets which may help respond to various questions: for example, distinguish a pure product using a combination, a real product/fake product, raw material of high quality/low quality, etc. For several more years a progressive transformation will begin. Machine Learning is an area of mathematics/programming science which can solve various issues without having to write simple instructions and which includes creating and applying algorithms. Uncontrolled learning is part of machine learning for analysing the information on the internal structure for which there is no input information. A new approach has been proposed to identify medicinal plants from front and back pictures of the leaf. Unique combinations of ethical, colour, and texture features that optimise the identification rate of green leaves have been established. The number of parameters used as inputs of the models and their importance when separating different species, make the models more solid than the previous studies and more medicinal plant species can be identified. The group labelled classification model performs labelling to the features considered for improving the accuracy and reducing the time complexity in medical plant recognition process. The proposed model when compared to the traditional methods exhibits better performance in accuracy in the process of identification of medical plants.

REFERENCES

[1] Rahmani, M.E., Amine, A., Hamou, R.M. (2016). Supervised machine learning for plants identification based on images of their leaves. International Journal of Agricultural and Environmental Information Systems (IJAEIS), 7(4): 17-31. https://doi.org/10.4018/IJAEIS.2016100102

[2] Masarczyk, W., Głomb, P., Grabowski, B., Ostaszewski, M. (2020). Effective training of deep convolutional neural networks for hyperspectral image classification through artificial labelling. Remote Sens., 12(16): 2653. https://doi.org/10.3390/rs12162653

[3] Dileep, M.R., Pournami, P.N. (2019). AyurLeaf: A deep learning approach for classification of medicinal plants. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), pp. 321-325. https://doi.org/10.1109/TENCON.2019.8929394

[4] Azlah, M.A.F., Chua, L.S., Rahmad, F.R., Abdullah, F.I., Alwi, S.R.W. (2019). Review on techniques for plant leaf classification and recognition. Computers, 8(4): 77. https://doi.org/10.3390/computers8040077

[5] Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H., Tekinerdogan, B. (2019). Analysis of transfer learning for deep neural network based plant classification models. Computers and Electronics in Agriculture, 158: 20-29. https://doi.org/10.1016/j.compag.2019.01.041

[6] Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., Stefanovic, D. (2019). Solving current limitations of deep learning based approaches for plant disease detection. Symmetry, 11(7): 939. https://doi.org/10.3390/sym11070939

[7] Mukti, I.Z., Biswas, D. (2019). Transfer learning based plant diseases detection using ResNet50. 2019 4th International Conference on Electrical Information and Communication Technology (EICT), pp. 1-6. https://doi.org/10.1109/EICT48899.2019.9068805

[8] Chaki, J., Dey, N., Moraru, L., Shi, F. (2019). Fragmented plant leaf recognition: Bag-of-features fuzzy-color and edge-texture histogram descriptors with multi-layer perceptron. Optik, 181: 639-650.

https://doi.org/10.1016/j.ijleo.2018.12.107

[9] Tammina, S. (2019). Transfer learning using VGG-16 with deep convolutional neural network for classifying images. International Journal of Scientific and Research Publications, 9(10): 9420. https://doi.org/10.29322/IJSRP.9.10.2019.p9420

[10] Durairajah, V., Gobee, S., Muneer, A. (2018). Automatic vision based classification system using DNN and SVM classifiers. 2018 3rd International Conference on Control, Robotics and Cybernetics (CRC), pp. 6-14. https://doi.org/10.1109/CRC.2018.00011

[11] Bhatt, P., Sarangi, S., Shivhare, A., Singh, D., Pappula, S. (2019). Identification of diseases in corn leaves using convolutional neural networks and boosting. Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019), pp. 894-899. https://doi.org/10.5220/0007687608940899

[12] Tan, J.W., Chang, S.W., Abdul-Kareem, S., Yap, H.J., Yong, K.T. (2018). Deep learning for plant species classification using leaf vein morphometric. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 17(1): 82-90. https://doi.org/10.1109/TCBB.2018.2848653

[13] Sabri, N., Aziz, Z.A., Ibrahim, Z., Rosni, M.A.R.B.M., Ghapul, A.H.B.A. (2018). Comparing convolution neural network models for leaf recognition. International Journal of Engineering Technology, 7(3): 141-144.

[14] Batvia, V., Patel, D., Vasant, A.R. (2017). A survey on ayurvedic medicine classification using tensor flow. International Journal of Computer Trends and Technology, 53(2): 68-70.

[15] Ghasab, M.A.J., Khamis, S., Mohammad, F., Fariman, H.J. (2015). Feature decision-making ant colony optimization system for an automated recognition of plant species. Expert Systems with Applications, 42(5): 2361-2370. https://doi.org/10.1016/j.eswa.2014.11.011

[16] Salve, P., Sardesai, M., Manza, R., Yannawar, P. (2016) Identification of the plants based on leaf shape descriptors. In: Satapathy S., Raju K., Mandal J., Bhateja V. (eds) Proceedings of the Second International Conference on Computer and Communication Technologies. Advances in Intelligent Systems and Computing, vol 379. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2517-1_10

[17] Chi, X., Zhang, Z., Xu, X., Zhang, X., Zhao, Z., Liu, Y., Wang, Q., Wang, H., Li, Y., Yang, G., Guo, L., Tang, Z. Huang, L. (2017). Threatened medicinal plants in China: Distributions and conservation priorities. Biological Conservation, 210: 89-95. https://doi.org/10.1016/j.biocon.2017.04.015

[18] Fuentes, S., Hernández-Montes, E., Escalona, J., Bota, J., Viejo, C.G., Poblete-Echeverría, C., Tongson, E., Medrano, H. (2018). Automated grapevine cultivar classification based on machine learning using leaf morpho-colorimetry, fractal dimension and near-infrared spectroscopy parameters. Computers and Electronics in Agriculture, 151: 311-318. https://doi.org/10.1016/j.compag.2018.06.035

[19] Gonzalez Viejo, C., Fuentes, S., Torrico, D., Howell, K., Dunshea, F.R. (2018). Assessment of beer quality based on foamability and chemical composition using computer vision algorithms, near infrared spectroscopy and machine learning algorithms. Journal of Science of Food and Agriculture, 98(2): 618-627. https://doi.org/10.1002/jsfa.8506

[20] Wei, C.S., Lei, F.H., Ai, W.X., Feng, J., Zheng, H., Ma, D., Shi, X.H. (2017). Rapid identification of 6 kinds of traditional Chinese medicines containing resins and other components based on near infrared refectance spectroscopy and PCA-SVM algorithm. Chinese Journal of Experimental Traditional Medical Formulae, 23(9): 25-31.

[21] Prasad, S., Kumar, P.S., Ghosh, D. (2017). An efficient low vision plant leaf shape identification system for smart phones. Multimedia Tools Appl., 76(5): 6915-6939. https://doi.org/10.1007/s11042-016-3309-2

[22] Munisami, T., Ramsurn, M., Kishnah, S., Pudaruth, P. (2015). Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers. Procedia Computer Science, 58: 740-747. https://doi.org/10.1016/j.procs.2015.08.095

[23] Babatunde, O.H., Armstrong, L., Leng, J., Diepeveen, D. (2015). A survey of computer-based vision systems for automatic identification of plant species. Journal of Agricultural Informatics, 6(1): 61-71. http://dx.doi.org/10.17700/jai.2015.6.1.152

[24] Patil, A., Bhagat, K.S. (2016). Plant identification by leaf shape recognition—a review. International Journal of Engineering Trends and Technology, 35(8): 359-361.

[25] Naresh, Y.G., Nagendraswamy, H.S. (2016). Classification of medicinal plants: An approach using modified LBP with symbolic representation. Neurocomputing, 173: 1789-1797. https://doi.org/10.1016/j.neucom.2015.08.090