

An Efficient Double Labelling Image Segmentation Model for Leaf Pixel Extraction for Medical Plant Detection

Prabhat Kumar Thella¹, Dr.V.Ulagamuthalvi²

¹Research Scholar, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India. E-mail: prabhat.thella@gmail.com

²Associate Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India.

ABSTRACT

Medicinal plants have less adverse effects and less expensive than modern medicine in the pharmaceutical industry. On this basis, many researchers have shown significant interest in the research into the recognition of automatic medicinal plants that benefits human health. In order to produce a robust classifier, there are different possibilities of improvement that enable medical plants to be classified correctly in real time. A large number of segmentation processes are used to handle various types of images and implementations. No recognition system can deal effectively with all sorts of images. Common parameters are seen as the basis for segmentation in various pixels. Common parameters here refer to value of intensity, colour and so on. The convex subjects, edges of various complex structures and closed contours are mainly used for these techniques. A pixel is separated independently from its grey values as pixel objects. As pixels of objects, isolated points or small areas are distinguished, because connectedness is not regarded as an essential feature. In this method an original image will be separated into pixels with small adjacent pixels. The pixel size is normally based on the operator's mask size. The object's characteristics are calculated on the object's edges, and the masks often consider the pixels from the background images. In this proposed work, an efficient Double Labelling Image Segmentation Model (DLISM) for leaf pixel extraction for accurate medical plant detection is introduced. The proposed model is contrasted with the traditional models and the results proved that the proposed model exhibits better performance levels.

KEYWORDS

Image Segmentation, Pixel Extraction, Leaf Edge Detection, Double Labelling, Pixel Labelling, Plant Identification.

Introduction

For the human race, plants are important. Herbs were in particular used by indigenous people since ancient times as medicines. Herbs are normally defined by professionals on the basis of years of sensory or locative experience [1]. Recent advancements in computational technology have substantially contributed to herbal data identification. This makes it easier for many people, especially those who lack herbal recognition experience. In addition to time consuming procedures, laboratory-based research includes skills in sample processing and data interpretation [2]. A simple and reliable medical plant recognition technique is therefore essential. Computing and mathematical analysis can be an effective method to recognise herbs. The non-destructive method is the method of choosing to identify plants quickly, especially for those who are unable to afford the use of costly analysis [3].

Segmenting process generates a collection of isolated pixels from the rest of the image. In general, this method is based on the grey image histogram [4]. Threshold is usually used for whole images, pixels that surpass the critical value given to a group of pixels or other images. The histogram would have a deep valley between two summits for photos of good quality [5]. The grey backdrop is concentrated around these two summits. Therefore, in the area of the valley, the optimal thresholds are to be located [6]. If the picture is bi-modal or is almost bi-modal, the histogram's threshold technique works well. In real time, however, most of the images are full of sound / irregularities leading to multimodal histograms [7].

To maximise the image contrast and intensity, the grey image conversion into geometrical data is implemented. Then, in order to convert the value of the image into its nearest threshold, a binary image of the grey scaled image is created with one or two potential values per pixel [8]. Therefore, erosion and dilation are a range of operations carried out to eliminate behind-the-scenes noise, which can affect digital pictures, such as grains, and hole. The pictures are regarded as homogeneous if they do not show significant contrasting variations. These photos display very narrow peaks when viewed in the histogram. The lack of uniform illumination on the picture is responsible for inhomogeneity [9]. In order to extend the narrow range to a more dynamic range, the picture is normalised. During

threshold conversion, the binary images from the process are reversed to make the background black. The Image edge detection model is indicated in Figure 1.

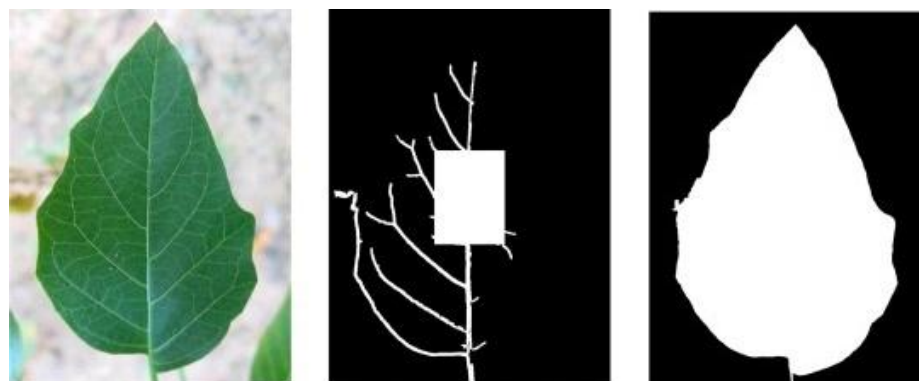


Fig. 1.Image Edge Detection Process

A clean or ready leaf picture for additional steps of the classification process cannot be obtained. A picture can have some noise, blur, lighting, or even shadowing effects. The presence of noise, blur and light in images complicate the successful characterisation of the picture. The shadows near leaf images influence their form and misinterpret the leaf during the stage of form analysis [10]. Photos should then be pre-processed in order to minimise noise, blur and light effects to improve the captured image for further functionality extraction [11].

Digital image processing transforms an image and performs operations such as creating, processing and quality improvement. Picture acquisition, improvement and restoration, Morphological processing, segmentation, presentation and description and object recognition are key steps in digital image processing. Image segmentation is the essential part of the image processing, which divides an image into meaningful areas, depending on features like grey, spectrum, texture and colours [12]. The characteristics for medicinal plants are very necessary. Input and output are used as images in some methodologies. The inputs and outputs are attributes derived from these images in segmentation methods [13].

Images are divided into various areas or parts, which do not overlap, separating subjects from background areas [14]. The key stages of segmentation are unnecessary data exclusion, subject recognition and the identification of characteristics. Picture segmentations are characterised by regions, borders, area, perimeter, statistics and entropy [15]. In terms of colour, intensity or texture, the region pixels are divided. Segmenting methods are divided into two major categories: layer-based segmentation methods and blocks [16]. The proposed model considers only block-based segmentation type for getting accurate results. The following two properties, similarity and discontinuity, are utilised by this segmentation process [17]. The approach to similarity partition an image into areas where a partitioned image is divided into edges based on sudden intensity shifts [18]. The proposed model segments the leaf images into multiple partitions and extract pixels for better analysis for accurate medical plant species recognition.

Literature Survey

The Leaf Shape Plant Recognition Proposed by KinjalMunot et al. [1] using morphologic leaves of plants as their formal features, i.e., appearance ratio, eccentricity, rectangularity, area, circularity and invariant moments. These features are then used for recognition through hypersphere classification of moving median centres (MMCs). Rafael et al. [2] created a plant species identification hand-held device based on plant leaf form using isolated leaf pictures in a simple context. The user captures plant leaf pictures and sends these pictures to a server that automatically identifies the plant leaf or provides top matches.

Bouligand– Minkowski's fractal-dimensional method of identifying the leaf leaves of the Brazilian Atlantic forest and Brazilian Cerrado scrubland suggested a way to evaluate leaf complexity model proposed by Rahul Basak et al. [4]. The tests were carried out with ten species of Brazilian Atlantic forest, as well as four more species belonging to a total of 14 species of Brazilian Cerrado. A mobile image recovery method based on the adaptive grid-based

matching algorithms and a nearest neighbouring search system, proposed by Song Yuheng et al., [5]. They also suggested a recovery system with features from the leaf vein.

In order to classify plant-leaves surface and extract textures for the purposes of plant-leaf identification, Dharmendra Kumar et al. [7] proposed a new texture signature model, based on the BouligandMinkowsky fractal-scale process. They developed a dataset of 20 species of plant leaves and carried out methodological experiments. The method for extracting surface texture features by means of Gabor filters of different scales and orientation for classifying Brazilian flora has been proposed by C Ananthi et al. [8]. Twenty separate species have been experimented in this model.

Cerutti et al. (2013) suggested a plant identification model based approach. They suggested a two-stage method for polygonally approximated contour segmentation and used high geometric descriptors to semiconduct plants based their leaf images. Du et al. (2013) suggested a method for extracting the form of the leaf and leaf vein, using fractal dimensions. For plant classification, the ring projection fractal property of the leaf form is extracted.

I KirubaRaji et al. [9] suggested a method of plant classification based on plant leaves shape characteristics. For classification using linear discriminatory analysis, shape features, such as invariant moments, multiscale distance matrix and some morphological features were used. PatilPriyanka Vijay et al. [10] have suggested an automated classification scheme to classify the plants' veins as three distinct legume species, namely, soya, red and white beans. The method of classification is based on the photos of plant leaves scanned. They introduced a new segmentation system, implemented with a straightforward transformation and adaptive threshold. Furthermore, classification-related functions, such as vector support, penalised discriminant testing and random forest have been considered based on vein measurements.

Mohd. Aquib Ansari et al.[12] extracted plant leaves characteristics using a different degree of resolution of Pyramid Oriented Gradient Histogram. For leaf image modelling, a clustering approach was applied. For recognition purposes, they calculated distance from multiples of picture sets. Tree image segmenting from complex backgrounds has been developed by R. Adamset al. [14]. The trees from various backgrounds are broken into a single series of tree image pixels, based on the visual variations of the tree and the surrounding objects. The picture of the tree is divided according to gray-level coincident matrix colour and texture features. Only when no other green plants are close to the picture of the tree is the correct segmentation stated ideal. For different plant types in different lighting and soil types, Rozario, Liton Jude et al.[17] have suggested segmentation into field areas of green vegetation. For enhancing segmentation precision, the mean-shift segmentation algorithm is used. The picture was later separated by a mean-shift technique into vegetation and non-vegetation. RGB and HSI have been calculated for a mean vector value.

Proposed Model

The identification and classification of leaf is important for agriculture, forestry, rural medicine and other businesses. The image processing techniques play a critical role in medical plant detection. The thresholding technology uses histogram to transform a grey image into a binary image. Histogram is a graphical display of a digital gray-scale distribution and shows how many times per gray-level occurs. The background and the foreground are distributed over the histogram of the image intensity. The degree of strength of the first object differs from the context. In this research work, histogram data containers is separated into equal intervals and fit the data with a suitable polynomial and evaluate the polynomial coefficients. Initially the medical plant image is considered as input and its polynomial $p(x)$ is represented as:

$$p(x) = a_0 + a_1x + a_2x + + a_nx^n \quad (1)$$

$a_0, a_1, a_2...$ are the coefficients. The equation describes, on the basis of values n , a straight line ($n = 1$), quadratic parabolic ($n = 2$), a curve ($n = 3$), etc., unknown to ($n+1$), x an autonomous variable. For image analysis tasks, segmentation is an important method. Many of the current image description and recognition techniques in particular rely heavily on the detaching performance. The 2D image processing has a wide variety of useful applications such

as visualising, measuring objects and length, identifying anomalies, quantifying and classifying tissues and more. The aim is to simplify images into numeric data that is easier to analyse and make the data more meaningful. The medical leaf considered undergo segmentation for accurate analysis on the smaller leaf portions for accurate shape and type prediction. The process of medical leaf segmentation is performed as

$$\lambda = \frac{\sum_{i=1}^P I_i + T_h}{(p+1)} + \theta + \int_i^{p+1} I(i) + npix(i+1) \quad (2)$$

$$Seg(I(x, y)) = \sqrt{\frac{\sum_{i=1}^P (I_i - \lambda)^2 + ((p+i)_c - Th)^2}{(P+1) + \lambda}} \quad (3)$$

Here p is the polynomial representation of image, I is the image considered, T_h is the threshold value, λ is the initial segmented modules, $p+i$ is the neighbour pixel value extracted. Let r_i be the greyness of the neighbouring pixels in any structure with the neighbouring n and m numbers around the middle pixel r_c . The mean and standard deviation of all p neighbouring numbers are calculated accordingly:

$$\omega = \frac{\sum_{i=1}^P r_i(I(x, y))}{p} \quad (4)$$

$$\delta = \sqrt{\frac{\sum_{i=1}^P (r_i - \omega)^2 + r_i - Seg(I(i))}{p(I(i+1))}} + Th \quad (5)$$

Here ω is the mean level and δ is the standard deviation of the image pixels calculated. The pixel intensity levels are calculated for considering the image content as relevant or irrelevant. The image intensity values are calculated as,

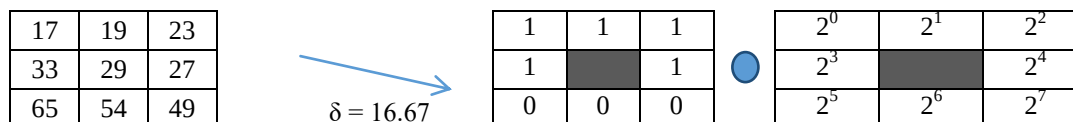
$$PixInt(I(i, j)) = \sum_{i=1}^P \delta * 2^{i-1} + \omega - Th \quad (6)$$

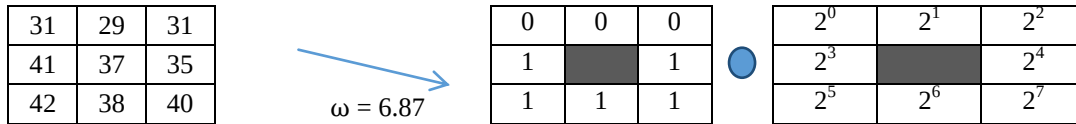
$$where \quad PixInt(I) = \begin{cases} 1 & \text{if } (r_c - \omega) < r_i < (r_c + \delta) \\ 0 & \text{otherwise} \end{cases}$$

After performing image segmentation, the image segment is considered and the initial labelling process is done as

$$IL(\epsilon) = \frac{\omega + r_{i_{avg}} + \max(\sum_{p(i)} (r_c - \omega) + (r_c + \delta)))}{\max(\sum_{i-}) * Th} \quad (7)$$

If the grey value of the neighbour pixel r_i lies between the interval $[(r_c - \omega), (r_c + \delta)]$ then the vector V is 1; otherwise 0. The pixels are represented in the array that is indicated in matrix format.





The proposed Double Labelling Image Segmentation Model framework is represented in Figure 2 that clearly illustrates the process of image segmentation and cluster formation.

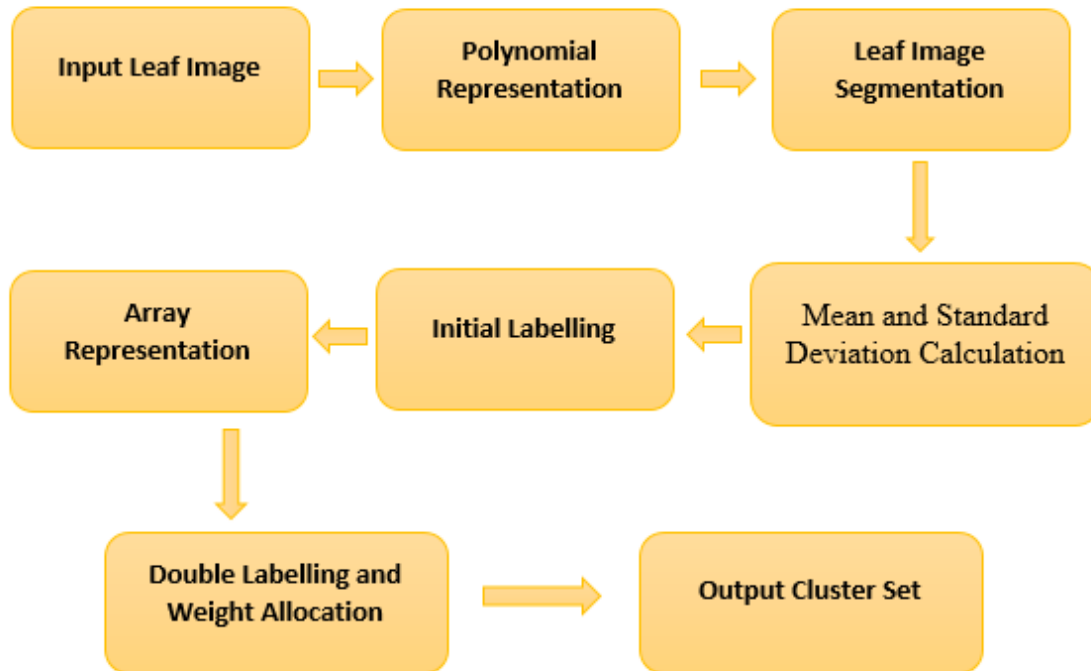


Fig. 2. Proposed Model Framework

The Contrast, dissimilarity and entropy of the considered image is calculated as

$$contrast = \sum_{i=1}^{p-1} P_{i,j} + I(i-j)^2 \quad (8)$$

$$Dissimilarity = \sum_{i=1}^p ri + IL(\varepsilon) - |i-j| * P(i,j) \quad (9)$$

$$Entropy = \sum_{i=1}^p -\ln(P_{ij})P_{ij} + Th \quad (10)$$

The pixels extracted from the images are stored in an array and then they are analysed for the similarity levels and double labelling is performed as

$$DL(Sim_{score}) = \frac{1}{\lambda} \sum_{i=1}^p \left\{ \max \left[\frac{1}{1 + abs \left(IL(\varepsilon)_n^{(ij)+} - PixInt_n^{rc} \right)}, \frac{1}{1 + \min \left(Seg(I(i,j))_{ri}^{(ij)-} - Seg(I(j,j+1))_{rj}^{ij} \right)} \right] \right\} \quad (11)$$

The weights are allotted based on the double labelling done and the clusters are established which can be further analysed to detect the medical plant type. The weights are allotted to each pixel as

$$W_{i,j} = \exp\left(\frac{-(I_{i,j}(DL(Simscore)) - \omega(i+1))^2}{1L(\varepsilon) - |i-j| + Th}\right) + \sum_{k=1}^c \left\{ \frac{\|I_j - \delta_i\|}{\|I_j - \omega_j\|} \right\}^{\frac{2}{p}-1} \quad (12)$$

The text values are stored in the matrix and is represented as

$$T^l(X, Y) = I^{(i)}(X, Y) + \frac{1}{p} \sum_{i=1}^{i=N} W_{i,j}^N \left((X_i - Y_j^{(M)}) + X(W) + Y(W) \right) + Th \quad (13)$$

Results

The proposed image segmentation model is implemented in ANACONDA SPYDER and the leaf image dataset is considered from the link, <https://data.mendeley.com/datasets/nnytj2v3n5/1>. In order to improve consistency, the output of diverse image segmentation analysis was evaluated in all processing factors. In the processing stages, the selection of an image with corresponding problem is calculated to recognise disadvantages on each step, as well as to recommend, simulate and evaluate solutions for optimal and effective solutions. The proposed Double Labelling Image Segmentation Model (DLISM) is compared with the traditional Adaptive Dropout Depth Calculation (ADDC) method and the results show that the proposed model performance is better in terms of accuracy and time levels. The medical leaf relevant pixel extraction model is represented in Figure 3.



Fig. 3. Pixel Extraction Layout

Segmenting images is the method of splitting a digital image into many segments (sets of pixels, also known as image objects). The objective of segmentation is to make representing the image more meaningful and easier to analyse. The input image provided need to undergo segmentation and pixel extraction that is indicated in the Figure 4.

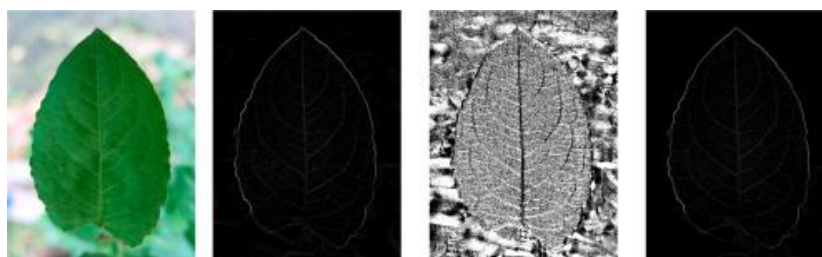


Fig. 4. Medical Leaf Relevant Feature Extraction

Segmentation is done in the input image by defining the area frontiers. Regional image segmentations are initially used to look for seed points in the input image, and proper regions are used to enter object boundaries. The medical leaf segmentation model is indicated in Figure 5. The leaf edges and useful information is extracted from the leaf image and the pixels are analysed.

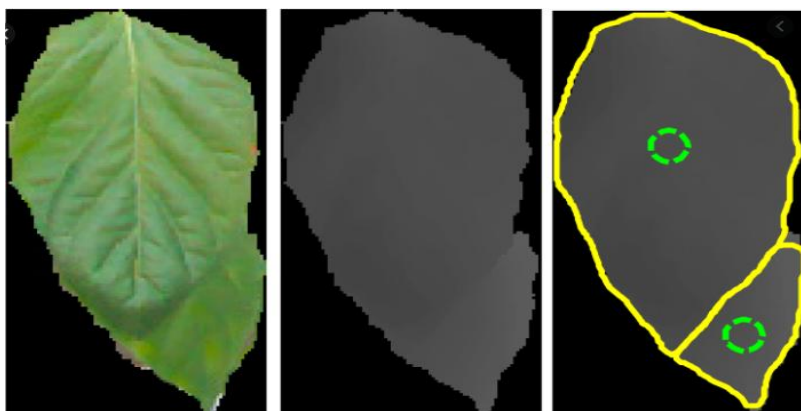


Fig. 5.Medical Leaf Segmentation Model

The main goal of dividing an image into several parts for further study is to obtain the only details possible or fragment. The image partitioning is based on certain image characteristics such as colour, texture, pixel strength, etc. There are many image segmentation techniques including thresholding, region-based method, edge-based method, clustering method, etc. The Figure 6 represents the Image Segmentation Time Levels. The proposed model takes less time in performing image segmentation when compared to the traditional model.

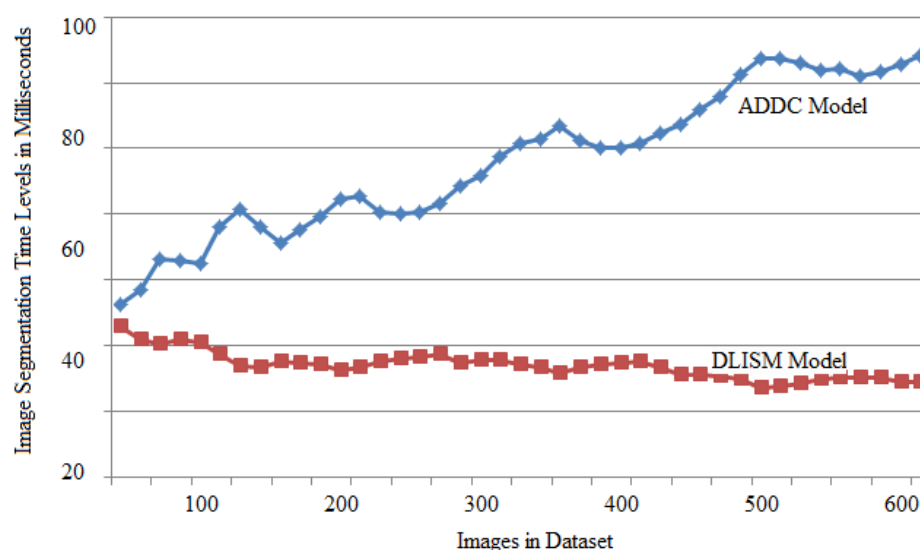


Fig. 6.Image Segmentation Time levels

The segmentation of images is a major part of image processing, almost everywhere it is used to process the images, so that our model can recognise pixels inside the image. The division divides the image into many segments or items. The problem depends on how much the fractionation of the picture is borne. When an image object is separated, segmentation must stop at that moment. The proposed model performs labelling of pixels for relevant features consideration only. The Image Labelling Accuracy Levels of the proposed and existing methods are illustrated in Figure 7.

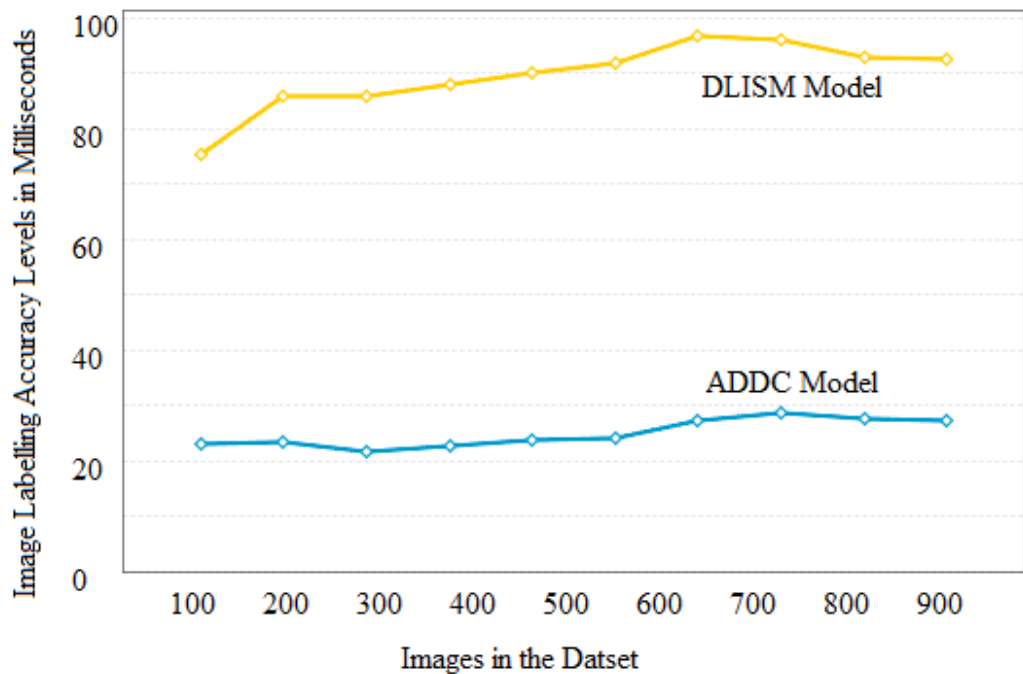


Fig. 7. Image Labelling Accuracy Levels

The Image Labelling time levels of the proposed and the existing models are clearly indicated in Figure 8. In image processing, the task of pixel labelling is very important in the proposed model. After completion of the operational file, the division of the image in many chunks makes it simple for further processing. Segmented object recognition increases the precision of an image and reduces the loss. The segmented image pixels and extracted and labelling is performed for considering those pixels for medical plant detection.

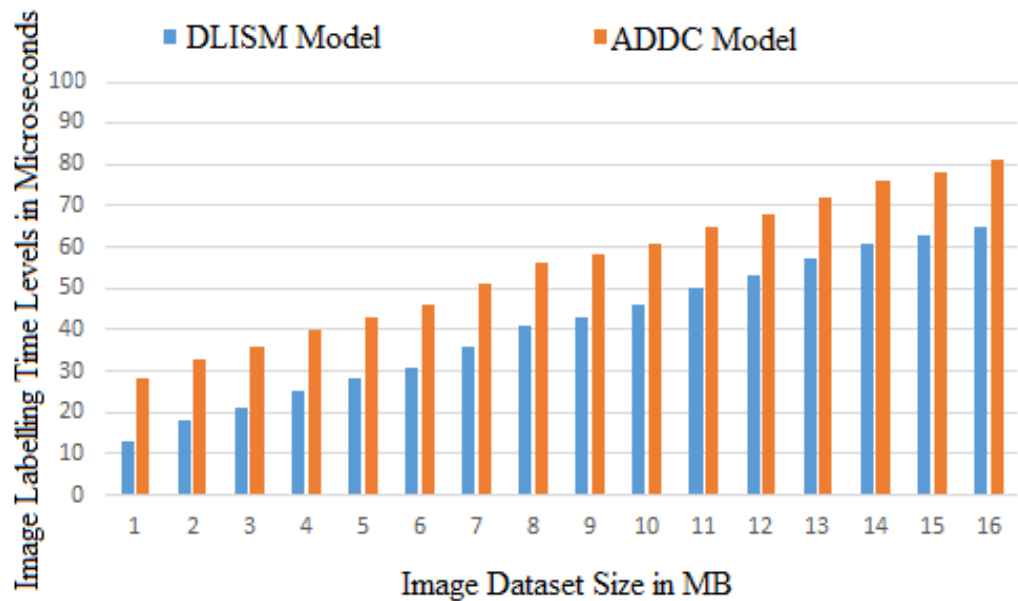


Fig. 8. Image Labelling Time Levels

Pixel extraction is part of the reduction of dimensionality process by dividing and reducing the initial collection of raw data into groups which can be more regulated. The main feature of these massive data sets is that they have several variables. In order to process these variables, several computer resources are needed, so extraction of the

pixels helps obtain the best feature in these large data sets by selecting and combining variables into functions, thus reducing data volumes effectively. The Pixel extraction time levels of the proposed and traditional models are included in the Figure 9. The proposed model pixel extraction time levels are less when compared to existing ones.

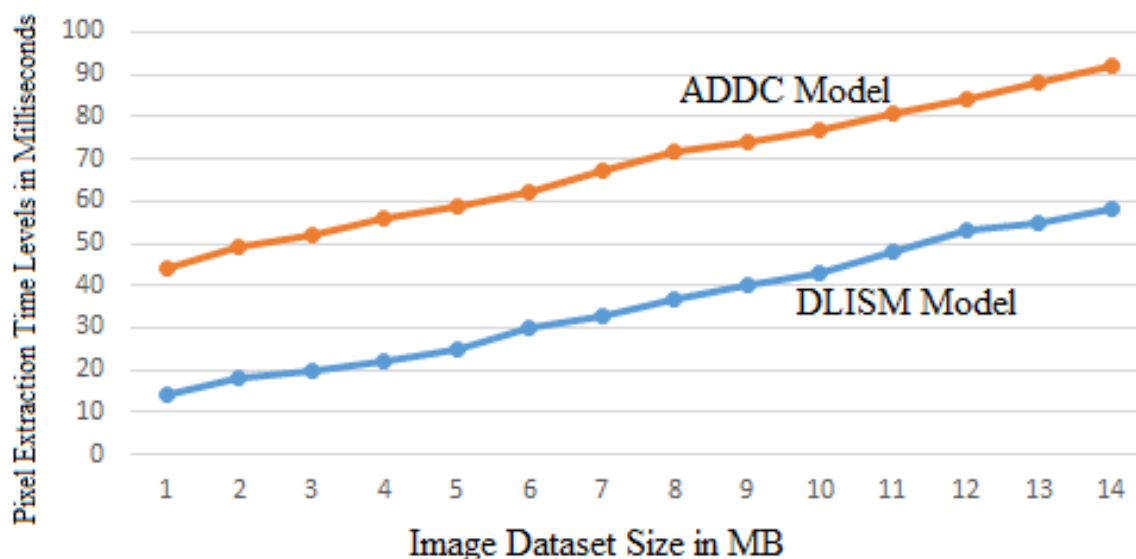


Fig. 9. Pixel Extraction Time Levels

Picture segmentation is a dynamic and demanding process, and many factors, such as noise, low contrast, lighting, and object boundaries irregularity, affect it. The proposed model segmentation accuracy levels are contrasted with the existing model and the result is indicated in Figure 10.

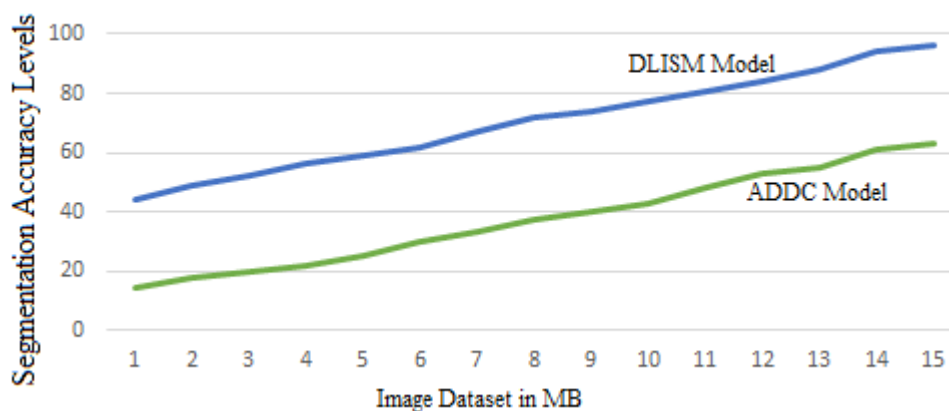


Fig. 10. Segmentation Accuracy Levels

The proposed model segmentation accuracy levels are contrasted with the existing model and the accuracy levels of the segmentation process is represented in Table 1.

Table 1. Segmentation Accuracy Levels

Method Name	Accuracy Level	Average Time for Segmentation(milliseconds)
Proposed DLISM Model	96%	36
Existing ADDC Model	64%	94

Conclusion

The simplest techniques, and minimal uses are the segmentation techniques based on grey level techniques, such as thresholding and region-based techniques. They can increase the efficiency through the integration of artificial intelligence techniques. Techniques based on atlas or table-search textural features have excellent results on the segmentation of medical pictures. A fully automated segmentation method is proposed in this paper for medicinal leaf images in a complex setting. First, from the colour picture the gradient is derived directly. On the basis of a feature that is found in the leaf area and by effective enhancement and extraction, a more precise image of the foreground marker is produced. Finally, the pixels dominated by the pixel marker is used to segment images. The proposed approach is better than several of the major completely automated image segmentation techniques like FCN deep learning through several experimental tests based on a self-built database and another commonly used database. The speed is also high than the segmentation of the OTSU threshold model. In essence, the proposed approach has solved the problem of completely automatic leaf image segmentation in complex contexts and has encouraged the recognition of the leaf image in complex backgrounds.

References

- [1] KinjalMunot, Nishi Mehta, Sakshi Mishra, Ratnesh N, Chaturvedi, "A Review on Image Segmentation Techniques with an Application Perspective," Volume 8, No. 9, November-December 2017. IJARCS.
- [2] Rafael, C., Gonzalez, Richard, E., Woods. Digital Image Processing. *University of Tennessee, MedData Interactive*. © 2008 by Pearson Education, Inc.
- [3] Naveen Tokas, ShrutiKarkra, Manoj Kumar Pandey, "Comparison of Digital Image Segmentation Techniques- A Research Review," *International Journal of Computer Science and Mobile Computing*, Vol.5 Issue.5, May- 2016, pg. 215- 220.
- [4] Rahul Basak, Surya Chakraborty, Aditya Kumar Mondal, SatarupaBagchiBiswas, "Image Segmentation Techniques: A Survey," Volume 05 Issue. 04 |Apr-2018, *International Research Journal of Engineering and Technology (IRJET)*.
- [5] Song Yuheng, Yan Hao, "Image Segmentation Algorithms Overview".
- [6] C Mariyammal, S Sasireka, "Survey on Image Segmentation Methods," *International Advanced Research Journal in Science, Engineering and Technology*. Vol. 5, Issue 11, November 2018.
- [7] Dharmendra Kumar, Kanak Kumar, "Review on Different Techniques of Image Segmentation Using MATLAB," *International Journal of Science, Engineering and Technology*. 2017, Volume 5 Issue 2.
- [8] C Ananthi, Azha, Periasamy, S Muruganand, "Pattern Recognition of Medicinal Leaves Using Image Processing Techniques," *Journal of NanoScience and NanoTechnology*, Volume 2, Issue 2, Spring Edition, ISSN 2279 – 0381.
- [9] I KirubaRaji, K KThyagarajan, "An Analysis of Segmentation Techniques to Identify Herbal Leaves from Complex Background," *ICCC 2014*, 1877-0509 © 2015, Published by Elsevier B.V.
- [10] PatilPriyanka Vijay, Prof N. C. Patil, "Gray Scale Image Segmentation using OTSU Thresholding Optimal Approach," *Journal for Research*, Volume 02, Issue 05, July 2016.
- [11] PinakiPratimAcharjya, Ritaban Das, DibyenduGhoshal, "A Study on Image Edge Detection Using the Gradients," *International Journal of Scientific and Research Publications*. Volume 2, Issue 12, December 2012.
- [12] Mohd. Aquib Ansari, DikshaKurchaniya, Manish Dixit, "A Comprehensive Analysis of Image Edge Detection Techniques," *International Journal of Multimedia and Ubiquitous Engineering*, Vol.12, No.11 (2017), pp.1-12.
- [13] Venkatesh Jatla, Marios S. Pattichis, Charles and Nick Arge "Image Processing Methods for Coronal Hole Segmentation, Matching, and Map Classification" *In IEEE Transactions on Image Processing*, Vol.29, pp:

1641 - 1653, 2019.

- [14] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.16 issue.6, pp. 641-647, 2014.
- [15] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in plant science*, vol. 7, 2016.
- [16] Mathworks, "Texture Analysis Using the Color-Level CoOccurrence Matrix (GLCM)," 2017.
- [17] Rozario, Liton Jude, Tanzila Rahman, and Mohammad Shorif Uddin. "Segmentation of the Region of Defects in Fruits and Vegetables." *International Journal of Computer Science and Information Security*, Vol.14, no. 5, pp:121-1252016.
- [18] J.Barbedo. "A new automatic method for disease symptom segmentation in digital photographs of plant leaves" *In European Journal of Plant Pathology*, Vol.147, Issue:2, pp:349-364, 2017.