Plant leaf segmentation through connected pixel approach

Dr. Dayanand R B
Professor, Department of Computer Science,
KSIT, Bengaluru
Email-dayanandarb@ksit.edu.in

Daneshwari A Noola
Assistant Professor, Department of Information
Science, BLDEACET, Vijayapura
Email- daneshwari_a@rediffmail.com

*Abstract-***Agricultural plays one of the eminent role for human survival, it has become much more essential due to population increase and food demand, and hence the crop yield has to be produced according to the demand. However, one of the reason that quality and quantity of the crop gets compromised is the disease and in past various methodology has been proposed, however they lack on the various model metrics or the segmentation is achieved for the particular leaf, . In this paper, we have proposed a methodology named as SCPA (Segmentation through Connected Pixel Approach). The main objective of this paper is to achieve high accuracy segmentation. SCPA is the two step approach first we find the ROI(Region of Interest) of the particular leaf and in the second approach we find the instance based ROI i.e. for the whole plant, here both the step are performed simultaneously through incorporating one another. Moreover, SCPA is optimized iterative-based method and it is achieved through the approach of connected pixel approach. Connected pixels are the one where the edge of one pixel is connected to the other. When performed on the LSC dataset we achieve the accuracy of 95.10 %. This methodology is compared with the various state of art model and existing system by considering the model metric such as SBD, the results shows that SCPA model performs better than the other exiting method also the pictorial comparison of segmented leaf are shown and it shows our model identify it well when compared to others.***

Keywords: ROI (Region of Interest), segmentation, Plant Leaf segmentation

1 INTRODUCTION

In recent survey, it is found that economy of India highly depends on the agricultural productivity. Moreover, this is one the reason where detecting the disease has become one of the eminent task in recent days [1]. Negligence in this area causes the serious issue to the plant and these results in compromise in quantity, quality and productivity of crop [2]. Moreover, the traditional method for detecting the plant disease is through the eye observation by experts, to do that one need big team of experts that can monitor the plant all the time. Traditional method does not work in efficient manner, as it is costly, hence

The automatic detection of disease is required as it is efficient, accurate and cheaper.

In plants, there are several visible disease such as colored spots scorch, bacterial and fungal disease. This type of disease are visible and can be detected through the image processing technique.

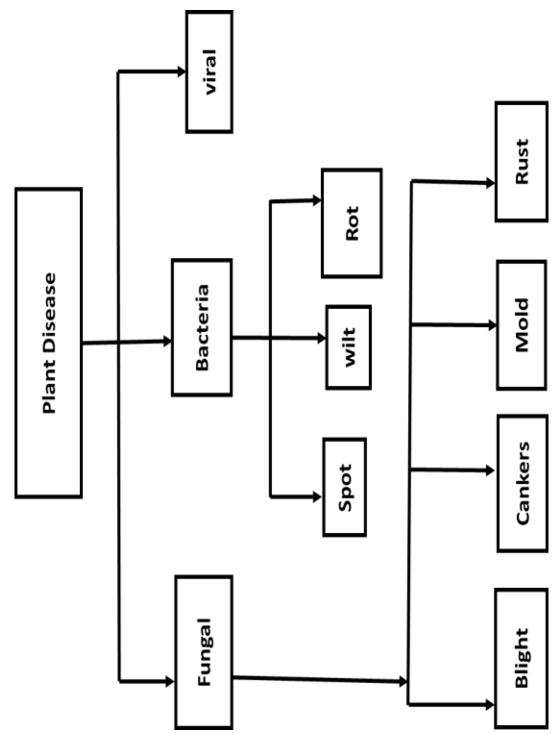


Figure 1 Types of plant Disease

The above diagram shows the different types of plant leaf disease, the figure shows the classification of various leaf disease. Moreover the plant disease caused due to the biotic agents.Plant disease are classified into three types based on the reason of the disease i.e. fungal, Bacteria and Viral. Moreover fungal is sub classified into various disease such as Blight, cankers, Mold and Rust whereas Bacteria is sub classified into disease such as spot wilt and root. All these disease causes the reduction in quality of the crops.

In order to detect the plant leaf disease, image-processing technique is one of the popular technique, which detects the disease accurately. Generally any image process follow 5 steps process to classify any disease and they are[3][4] **Image acquisition :** Here the plant is selected which is affected by particular disease then plant leaf are selected and snapshot of leaf are loaded into the system.

Pre-processing: Here the image are resized and noise are removed if there is any.

Segmentation: This is one of the main process where one needs to take care as it finds the affected area where the leaf is damaged. This part of detection process requires lot of care.

Feature extraction: Here the particular feature of the leaf are extracted and then through the classification technique it is classified into the particular type of disease.

As we discussed earlier how important role segmentation play in leaf disease detection. In this research work, we focus on the segmentation. Segmentation is the method of grouping or separating the particular image into several parts, there are different method of achieving segmentation. Segmentation process is based on the various feature that exist in the image, these feature can include the information about the color, boundaries or anything.

In past several technique were proposed for segmenting the images [6-8], through this part we try to discuss some of the existing technique that has been utilized for the segmentation.

[9] used the integration of k-means clustering algorithm along with the neural network to cluster and classify the disease; they performed the methodology on the particular types of disease. [10] applied the statistical procedure for the detection of disease and later they classified into the fungal disease, the evolution of the statistical features are performed like GLRM and GLCM, for classification, they used the KNN algorithm. Moreover, [11] proposed a technique of three step to detect the disease. First is pre-processing then segmentation and later they used the statistical analysis. For filtering, they used the median filter and affected part of the leaf are segmented through the SPR (statistical Pattern Recognition). [12] Proposed a methodology of the integration of Image processing technique and GSM, it used the Otsu's method along with the K-means clustering, for classifying the disease they used SVM as the classifier. [13] proposed a technique for disease identification and classification through the machine learning technique , here for the feature set extraction the color co-occurrence is applied, it was one of the important approach considering the root disease and stem disease however failed in case of leaf disease . [14] used four step methodology at first RGB image is taken as the input image here the color transformation is formed then the through the particular threshold value the mask of green pixels are removed, for further segmentation texture statistics are computed. [15] used the Gabor filer to extract the feature and ANN is used for the classification, this method helps in improvising the recognition Rate. Few other researcher such as in paper [16] researcher tries to achieve segmentation through the contagious function and [17] follows the data visualization technique to segment the plant leaf, taking a step ahead [18] uses the CNN based method is used for leaf segmentation. This method tries to achieve the region of interest with MLP(Multi Layered perceptron) which dynamically transform the feature description. All these above method does try to achieve the segmentation they do succeed in achieving the accuracy, however the instance based has been considered only by the few of the researcher, hence by considering these related work and their flaws we have proposed connected pixel approach which achieve the instance based ROI.

2 Motivation and contribution of the proposed work.

Agriculture is one of the important aspects from the economic point of view. Moreover, plant disease does affect the quantity as well as the quality of the crop. Traditional approach of disease detection is very much expensive and accuracy is less. Hence In this paper we propose a methodology named as SCPA (Segmentation through connected pixel Approach). Here at first we find the region of interest for the particular leaf, then we find the Instance based segmentation i.e. segmentation of the whole plant. The contribution of the paper is as follows:

1. SCPA achieves the segmentation with the complex background.
2. Segmentation is done in the two-step; in first step, segmentation is achieved for the particular leaf later in the second step segmentation is achieved for the whole plant.
3. Generate the accurate segmentation
4. Extensive simulation is performed by considering the dataset

5. Accuracy is above 95% in case of two different sub dataset of LSC.
6. SCPA not only achieves the segmentation for single leaf but for the whole plant.
7. Performs better than the other existing technique for segmentation.

This research work is organized in a way where first section presents the introduction where the importance of crops and several types of plant disease has been discussed, followed by that we discuss some of the existing methodology that exist for leaf segmentation. Second section presents the motivation and contribution of the research work. Third section presents the proposed methodology with the optimization technique. Last but not the least section presents the performance Evaluation of SCPA.

3 PROPOSED METHODOLOGY

Throughout the above discussion, it is observed that in order to detect the disease segmentation plays one important role. In here, we propose a methodology named as SCPA (Segmentation through Connected pixel Approach) which is optimization problem; figure 2 shows the process flow of SCPA model. Here, at first we take input as an image, then this image are decomposed into GCP (Group of connected pixel). Connected pixels are the one where the each pixel are touches to one of the neighbor pixel. Connected pixels gives us the advantage of better feature learning.

Then through applying the method, we learn the feature of connected pixel, later in order to find the region we construct a particular graph, which helps in finding the region of particular leaf. Moreover then we achieve the instance based ROI.

Let's consider a particular set of C images dataset of $D = \{D_1, D_2, \ldots \ldots . D_i\}$, Moreover we achieve the special feature learning for each image i.e. G. Then each image of D_n gets decomposed into various connected group of pixel, from here now onwards we call this as GCP, these GCP acts as the domain for the two-phase methodology i.e. achieving RoI and Instance based RoI.

$$C = \sum_n C_n,\ n\epsilon\{1,2,\ldots,D\}. \quad (1)$$

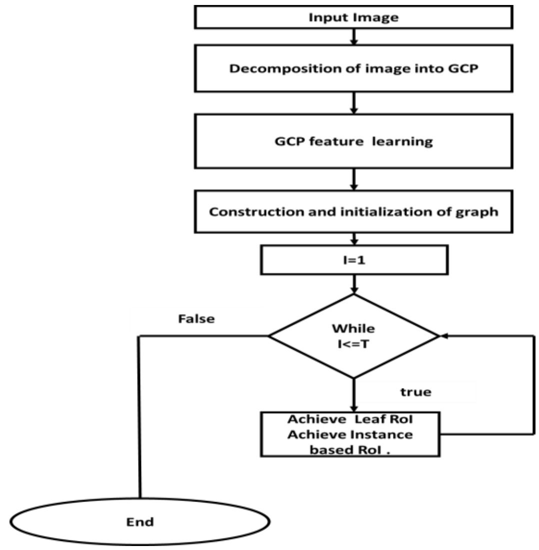


Figure 2: process flow of dual approach leaf disease detection method

The equation 1 shows the connected pixel, moreover to learn the particular feature we need the weights of the pixel, later this is integrated in order to achieve the instance based segmentation

$$w_m = [w_{m,1}, w_{m,2}, \ldots \ldots, w_{m,G}]^L \quad (2)$$

3.1 Connected pixel and feature learning

Each image is decomposed into various GCP and the similarity among the connected pixel are presented through the below equation. k_m and $k_{m'}$ are the feature that has been learnt, later both are combined for the particular connected pixel .

$$P(m\,, m') = exp\left(-\frac{Z^2(k_m, k_{m'})}{\delta}\right) \quad (3)$$

δ is kept as constant for the average distance between the connected pixel.

3.2 Construction of Two-Phase Segmentation Model.

Let us consider any graph $\mathbb{H}$ where $\mathbb{H} = (E = E_n\ and\ \tau = \tau_n)$,which represents the relationship among the connected pixel. E_n is the all connected pixel in D_n and τ set present the neighboring relationship. Moreover Edge weight are computed through the below equation.

$$\mathbb{E}(m\,, m') = \mathbb{a}(m\,, m') * P(m\,, m') \quad (3)$$

$\mathbb{a}(m\,, m')$is the count of the neighbor pixel, here the weight of the pixel plays crucial role and $\mathbb{E}$ is the matrix which is associated with the generated graph.

3.3 Problem Definition

In order to achieve the ROI, We need weights extract the weights which is given as $\mathbb{W} = [w_1, w_2, \dots \dots, w_C]^T$ and for the instance we need $\mathbb{T} = [t_1, t_2 \dots \dots \dots \dots, t_C]$. Moreover, the Leaf segmentation and Instance based ROI, these two are formed simultaneously and this is achieved through the optimization of the following equation.

$$opt(\mathbb{W}, \mathbb{X}) = \|\mathbb{W}\|_2^2 + \gamma_1 \sum_{m:e_m \in \mathcal{V}} B(w_m) + \gamma_2 \sum_{1 \le n < n' \le n} R(t_n, t_{n'}) \quad (4)$$

$$+\gamma_3 \sum_{m:b_m \in \mathbb{E}} K(w_m, w_{m'}) + \gamma_4 \sum_{b_{mm'} \in \tau} U(t_m, t_{m'}) + \gamma_5 \sum_{b_{mm'} \in \tau} V(w_m, w_{m'})$$

Such that $\|w_m\|_1 = 1, w_m \ge \bar{0}$, $t_m \in \{0,1\}$, for $1 \le m \le C$. Here $\bar{0}$ indicates the entire zero vector and there are five constants that are γ_1, γ_2, γ_3, γ_4 and γ_5 are the constants and w_m is the image configuration of $D_n. w_{m'}$ is the defined similarity. $\mathbb{W}, \mathbb{X}$ Are the combined optimization such that both can be useful for one another. $\|\mathbb{W}\|_2^2$ is the regularization, $K(w_m, w_{m'})$ is to maintain between one another i.e. leaf and the instance based. Moreover, our methodology is twostep process for finding the region of interest. First, we find the important feature then we detect the region of interest.

3.4 Optimization

Through optimizing the equation 3 we achieve the ROI and Instance based ROI. Our model is designed in such a way that at each iteration one variable is optimized at the meantime the other variables are fixed, later the role of the other variables are switched accordingly. Moreover, the iteration are performed and variables are optimized until the absolute ROI is achieved.

Such that $\|w_m\|_1 = 1, w_m \ge \bar{0}$ for $1 \le m \le Cx$

3.5 Achieving Instance Based ROI

In order to achieve the instance based ROI, we optimize the equation 3 though fixing $\mathbb{W}$,

$$opt(\mathbb{X}) = \gamma_3 \sum_{m:b_m \in E} K(w_m, t_m) + \gamma_5 \sum_{b_{mm'} \in} U(w_m, t_{m'}) + \gamma_2 \sum_{1 \le m < \widehat{m'} < c} R(t_n, t_{n'}) \quad (6)$$

s.t $t_m \in \{0,1\}$ for $1 \le m \le C$

Once the iteration are done and both the equation i.e. 5 and 6 are minimized we achieve the leaf ROI and Instance based ROI. Thus, we find the ROI of particular plant through segmentation. Moreover, to evaluate the algorithm we perform the segmentation on the given LSC dataset, performance evaluation is discussed in the next section of our work.

4 PERFORMANCE EVALUATION

4.1 Dataset

In order to evaluate the performance of proposed methodology we have use LSC dataset[19], this particular dataset is publicly available and originally it was released in conjunction with the ECCV workshop on CV problems . This particular dataset constitutes of three sub-dataset i.e. A1, A2, A3, each consists of 128, 31 and 27 images. A1 and A2 has Arabidopsis images and A3 has Tobacco plant.

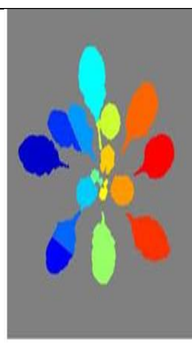
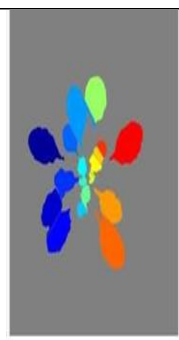
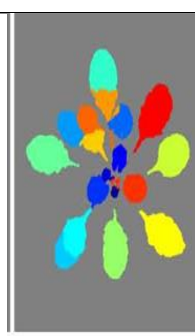
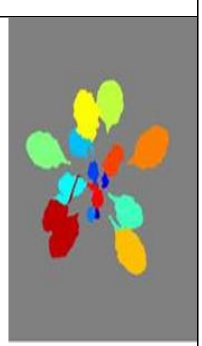
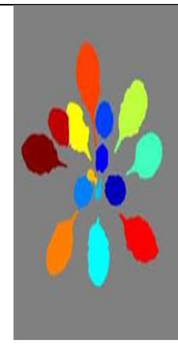
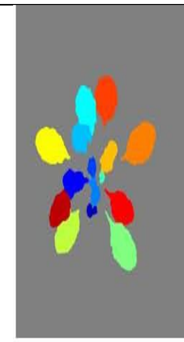
4.2 Pictorial comparison

Table 1 shows the comparison of segmented leaf with various technique along with the ground truth. Here first raw shows the input image i.e. a) and b) second raw shows the ground truth i.e. image c and d. The comparison is done with the some of the disease identification technique and as we can observe our model excels throughout the segmentation and it is depicted in image j) and k).

$$opt(\mathbb{W}) = \gamma_3 \sum_{m:e_m \in E} C(w_m, t_m) + \sum_{b_{mm'} \in} U(w_m, t_{m'}) \quad (5)$$

$$+\gamma_1 \sum_{m:b_m \in M} B(w_m) + \|\mathbb{W}\|_2^2$$

Table 1 comparison of leaf segmentation with the various technique

Input_Image	 *a)*	 *b)*
Ground_ truth	 *c)*	 *d)*
IPK[20]	 *e)*	 *f)*
MSU[21]	 *g)*	 *h)*
ERIS[22]	 *i)*	 *j)*
Proposed	 *k)*	 *l)*

4.3 Model Metric

The below table shows the performance of proposed model based on the metric described below:

Sensitivity is the performance metric, which is measure of capability of model to identify with the disease whereas specificity is the measure of capability of the model to identify the one without disease. Moreover, higher percentage of sensitivity and specificity shows the better model. In here for the different sub-dataset i.e. A1, A2 and A3 our model gets 70.59, 97.90 and 94.89. Accuracy is a metric that shows the efficiency of model and how well it can predict; here our model achieves the accuracy of 86.55, 99.23 and 99.53 for A1, A2 and A3 respectively. Similarly, F1- measure is the measure of balance between precision and recall. Here for A1, A2 and A3 our model achieves the 0.7516, 0.9412 and 0.7256.

Model Metric	**Dataset**		
	A1	A2	A3

Sensitivity	70.59%	97.90%	94.89%
specificity	97.26%	99.47%	99.53%
Accuracy	86.55%	99.23%	99.53%
F1-measure	0.7516	0.9412	0.7265

4.3.1 SBD Comparison table

Symmetric Best dice is one of the performance metric which is considered for the quantitative evaluation, the below table shows the comparison of proposed technique with various existing technique. The higher SBD indicates the more optimized model, in the table we observe that our model achieves the marginally improved SBD of 75.16%

Methodology	LSC-A1(SBD)
ERIS	70.3 %
MSU	68.1 %
IPK	64.6 %
PS	75.16%

5. CONCLUSION

In this paper, we have proposed a twostep optimization methodology through considering the connected pixel. In first step, we achieve the leaf ROI and in second step, we achieve instance-based leaf ROI. In order to evaluate the we have used standard dataset of LSC dataset. Moreover, the model is compared with the various existing technique and our model outperforms the other model by considering the segmented image and SBD as the performance metric. For further evaluation, we have used various model metric such as sensitivity, specificity, accuracy and F1-measure, our model achieves the average value of these four metrics as 87.79 %, 98.75%, 95.10, and 0.8061 respectively on all the three sub-datasets. Moreover our segmentation model excel considering all the important metrics, in the future work we would be performing classification of the disease which identifies the disease and classify into the same.

REFERENCE

[1]. Lin xuefeng and Wu Shengliang, "Study on integrated development of modern agriculture evaluation," *2011 International Conference on Business Management and Electronic Information*, Guangzhou, 2011, pp. 308-312.

[2]. Pramod Aggarwal, Shalika Vyas, Philip Thornton, Bruce M. Campbell, Martin Kropff, Importance of considering technology growth in impact assessments of climate change on agriculture,Global Food Security, Volume 23, 2019, Pages 41-48.

[3]. S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing," *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 768-771.

[4]. Arnal Barbedo, J.G. SpringerPlus (2013) 2: 660. https://doi.org/10.1186/2193-1801-2-660

[5]. ijai Singh, A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques, Information Processing in Agriculture, Volume 4, Issue 1, 2017, Pages 41-49, ISSN 2214-3173.

[6]. Scharr, Hanno & Minervini, Massimo & French, Andrew & Klukas, Christian & Kramer, David & Liu, Xiaoming & Luengo, Imanol & Pape, Jean-Michel & Polder, Gerrit & Vukadinovic, Danijela & Yin, Xi & Tsaftaris, Sotirios. (2015). Leaf segmentation in plant phenotyping: a collation study. Machine Vision and Applications. 27. 10.1007/s00138-015-0737-3.

[7]. P. Arbelaez, M. Maire, C. Fowlkes and J. Malik, "Contour Detection and Hierarchical Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898-916, May 2011.

[8]. Jifeng Ning, Lei Zhang, David Zhang, Chengke Wu, Interactive image segmentation by maximal similarity based region merging, Pattern Recognition, Volume 43, Issue 2, 2010, Pages 445-456

[9]. H.Al-Hiary, S. Bani-Ahmad, M.Reyalat, M.Braik & Z.AlRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, Vol. 17, No.1, pp. 31-38.March 2011.

[10]. J. D.Pujari, R. Yakkundimath & A.S.Byadgi, "Statistical Methods for Quantitatively Detecting Fungal Disease from Fruits" Images", International Journal of Intelligent Systems & Applications in Engineering Advanced Technology & Science (IJISAE), volume 1 no.4, pp. 60–67, ISSN:2147-6799, 12th Dec2013.

[11]. Youwen Tian, Lide Wang and Qiuying Zhou, "Grading Method Of Crop Disease Based On Image Processing", CCTA 2011, part II, IFIP AICT 369, pp. 427-433, 2012.

[12]. Pallavi S. Marathe, G.H. Raisoni, Savitri Phule, "Plant Disease Detection Using Digital Image Processing And GSM", IJESC, volume 7, issue 4, April 2017.

[13]. Mrunalini R Badnakhe, Deshmukh Prashant R. An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases. Int Conf Adv Inf Technol 2011;20. 2011 IPCSIT.

[14]. Arivazhagan S, Newlin Shebiah R, Ananthi S, Vishnu Varthini S. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agric Eng Int CIGR 2013;15(1):211–7.

[15]. Kulkarni Anand H, Ashwin Patil RK. Applying image processing technique to detect plant diseases. Int J Mod Eng Res 2012;2(5):3661–4.

[16]. Chen, Yuhao, Sriram Baireddy, Enyu Cai, Changye Yang, and Edward J. Delp. "Leaf Segmentation by Functional Modeling." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 0-0. 2019.

[17]. Aggarwal, Shilpi, Madhulika Bhatia, and Hari Mohan Pandey. "A New Method to Classify Leaves Using Data Visualization: Spreading Awareness About Global Warming." In Proceedings of ICETIT 2019, pp. 1129-1139. Springer, Cham, 2020.

[18]. 3. Prashar, Kapil, Rajneesh Talwar, and Chander Kant. "CNN based on Overlapping Pooling Method and Multi-layered Learning with SVM & KNN for American Cotton Leaf Disease Recognition." In 2019 International Conference on Automation, Computational and Technology Management (ICACTM), pp. 330-333. IEEE, 2019.

[19]. H. Scharr, M. Minervini, A. P. French, C. Klukas, D. M. Kramer, X. Liu, I. Luengo, J. M. Pape, G. Polder, D. Vukadinovic et al., “Leaf segmentation in plant phenotyping: a collation study,” Machine Vision Applications, vol. 27, no. 4, pp. 585–606, 2016.

[20]. J. Pape and C. Klukas, “3-d histogram-based segmentation and leaf detection for rosette plants,” in Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP), 2014, pp. 61–74.

[21]. X. Yin, X. Liu, J. Chen, and D. M. Kramer, “Joint multi-leaf segmentation, alignment, and tracking from fluorescence plant videos,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 6, pp. 1411–1423, 2018.

[22]. J. Yu *et al*., "Exemplar-Based Recursive Instance Segmentation With Application to Plant Image Analysis," in *IEEE Transactions on Image Processing*.