Journal of Theoretical and Applied Information Technology
15th October 2024. Vol.102. No. 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INTEGRATION OF AUTOMATED GRABCUT ALGORITHM WITH DEEPLABV3+ TO ENHANCE IMAGE SEGMENTATION FOR ACCURATE LEAF DISEASE DETECTION AND CLASSIFICATION

SREYA JOHN[1] , Dr. P. J. ARUL LEENA ROSE[*]

[1]Research Scholar, Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and technology, Kattankulathur Campus, Chennai- 603203, India
[*]Professor, Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and technology, Kattankulathur Campus, Chennai- 603203, India
Email: [1]sj2457@srmist.edu.in, [*]leena.rose527@gmail.com

ID 55439 Submission	Editorial Screening	Conditional Acceptance	Final Revision Acceptance
26-08-2024	29-08-2024	06-09-2024	08-10-2024

ABSTRACT

Detection and classification of plant diseases play a significant role in various fields such as plant pathology, agriculture and environmental studies. To produce effective segmentation of leaf images, this study provides a hybrid strategy that blends deep learning approaches with the enhanced GrabCut algorithm. The proposed method automates the original GrabCut algorithm in order to build initial masks, which are then refined using the powerful DeepLabv3+ model. A detailed correlative analysis is also performed to demonstrate that the suggested model with an efficacy of 95.99% outperforms existing deep learning models such as Unet, SegNet, and other commonly used segmentation methods. The results are obtained using evaluation metrics such as pixel accuracy, intersection of union, precision, recall, and F1 score, demonstrating that incorporating deep learning into the GrabCut algorithm significantly enhanced the leaf image segmentation process.

KEYWORDS: *Segmentation, Leaf Disease Detection, Grabcut Algorithm, Deeplabv3+ Network*

1. INTRODUCTION

Recently many technological advances have taken place in the field of Agriculture. The developments in this area began centuries ago inorder to meet the demands of a broadening population. The increased consumption and high usage of resources resulted in famine and shortage of natural resources in many parts of the world. This excess utilization of natural resources disrupted the biological cycle of life which culminated in climatic changes. Increased temperature and less amount of rainfall became a threat to the farmers as water is an important element for irrigation and other purposes. They also faced many other problems such as financial instability, resource limitations and various types of diseases due to overexposure to the high temperature. All these factors affected the agriculture sector tremendously.

Subsequently, life on earth became difficult and forced us to chase alternate ways to achieve high quantity yield in an eco friendly manner. This led to the introduction of various types of technology-driven devices in the agricultural field. The simplicity and ease in handling these devices increased their popularity. This type of farming was later known as precision agriculture or smart farming which enhanced the productivity, efficiency and sustainability in the agricultural practices [1]. The farmers were able to make informed decisions, adjust to the changing conditions, and fulfill the rising need for food in a more effective and environmentally responsible way by the incorporation of technologies and procedures into agricultural practices. However, there were still significant challenges that slowed the sector's expansion. The prevalence of numerous variants of plant diseases is just one of many such issues. For a very long time farmers and other agriculturist were clueless on handling this issue. The traditional way of tackling the plant diseases was the naked eye observation technique which was a tedious, tiring and time consuming process.

[*]Corresponding author E-mail: leena.rose527@gmail.com

Additionally, it was only able to offer treatment for a small number of locally prevalent ailments, even though experiments on treatment methods were time-consuming. Studies and research have been conducted on the application of technology in this field to provide a reliable method for disease detection. This resulted in the successful development of computerized plant disease detection systems which was a sudden relief to the farmers and agriculturists.

Automated crop disease identification is a significant technological advancement in the field of agriculture. Using this automated technology we can easily identify the diseases in large scale farming within a short span of time just as it is done in the small scale farming. It uses various technologies to identify and detect diseases in crops [2]. The main aim is to provide the farmers with a quick and accurate method that enables timely intervention and treatment of the disease. Recent trends in this area are the usage of machine learning and deep learning algorithms to implement and classify the diseases. Nowadays, these models are widely used as they are very effective in the early detection of various diseases. Also many technological modifications are brought to such systems as to provide aid to the famers regarding cultivation and many other things [3]. A strong, perfectly built framework is the necessary step in providing such an accurately performing model [4]. Therefore, it is very important to have a strong foundation before going into the major development areas such as training and testing which also involves providing equal or more importance and focus should be given to the preliminary steps that include cleaning and segmenting of the datasets.

Segmentation is a significant stage in disease detection that is considered as an initial step of dataset processing. An accurate segmentation can add to the performance of the entire model [5]. The procedure is to partition the preprocessed image into numerous segments based on the similarities in their pixel intensities. These partitions are designed to make it simple for the machine to comprehend the image and function as intended. In many of the existing models segmentation is done as an extended part of preprocessing or otherwise it is excluded. This variation in the order and usage of segmentation is due to the requirements and objectives of each study [6]. When the major goal of the research activity is to isolate and inspect the diseased areas or to remove undesired portions from the image, for example, segmentation can be done prior to feature extraction. In some cases, to reduce the computational burden it may be more efficient to perform segmentation before feature extraction [7]. This is done when dealing with large datasets or complex segmentation algorithms. It's significant to keep in mind that there is no segmentation strategy for disease detection systems that is universally applicable. Depending on the setting, goals, and limitations of the research, the segmentation process might take several forms [8]. To determine the best method for a particular study, researchers may test out various approaches. The ultimate objective is to diagnose diseases with accuracy and reliability while taking into accounts the particular needs and constraints of the study.

In this paper, segmentation is done mainly to eliminate the unwanted elements from the image. The presence of leaves and other objects in the background can affect the feature extraction stage which can result in the poor performance of the system [9]. Therefore, it is possible to help the machine in extracting important information without any confusions by segmenting or eliminating the background regions from the image. Here, it is done using a hybrid method that combines an enhanced grabcut algorithm with a deep learning algorithm to accurately segment the leaf image from its background. Segmentation using grabcut technique became prominent since 2004 [10]. Due to its ease of use and effectiveness, it is still an extensively utilized procedure. However, this approach has a few limitations, and in order to fix them, we improved the algorithm using deep learning. The proposed algorithm operates by fine tuning the background mask generated by enhanced grabCut algorithm using deeplabv3+. The primary goals of this study is to improve segmentation accuracy, particularly in cases where traditional methods struggle, and to offer a method that balances performance with computational efficiency.

The key contributions of the paper are highlighted below:

- A thorough study on recent developments in the automated agricultural crop leaf disease detection and classification has been done.
- The use of Machine Learning and Deep Learning algorithms in various fields of leaf disease identification and classification are discussed.
- Effective conventional segmentation techniques and their challenges are studied.
- An enhancement to the conventional grabcut algorithm using deep learning based deeplabV3 is proposed.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- The architecture of the proposed method is discussed in detail.
- Several assessment metrics such as pixel accuracy, recall and precision, intersection of union and F1 score are used to calculate the performance of the proposed method over the existing methods.

This paper is divided into five sections: Section 2 is related works that gives an overview on recent developments in the area of agricultural leaf image segmentation. Section 3 is methodology, explaining the methods and algorithms used in the proposed method. Section 4 describes the experiment and results of the proposed method and section 5 is conclusion and discussions on future works.

2. LITERATURE REVIEW

This section gives an overview of various studies that have been done in leaf image segmentation using machine learning and deep learning algorithms. Based on their methodologies these research works are grouped into two categories: 1) leaf image segmentation models using conventional methods 2) leaf image segmentation models using advanced methods.

Leaf image segmentation models using conventional methods: The traditional image segmentation methods and algorithms that are used for dividing an image into sections or objects based on attributes like color, texture, intensity, shape or other image qualities are referred to as conventional image segmentation [11]. These techniques don't involve deep learning or neural networks for their functioning. Conventional segmentation methods are widely used and are popular even now due to their various applications [12]. The authors of [13] use a combination of two conventional segmentation algorithms, otsu threshold and color vegetative indices to segment the unhealthy leaves from cruciferous crops. K means clustering is another popular conventional segmentation algorithm [14] that segment the image into various clusters based on the similarities and characteristics of the pixel values. In this study, the min-max hue histogram approach was employed in conjunction with k means clustering to pinpoint the cluster that comprises diseased regions. Fuzzy clustering is another kind of clustering algorithm that has been widely utilised in image segmentation tasks [15]. These models are famous for the effective handling of situations where clusters overlap or when data points exhibit characteristics of multiple clusters simultaneously. But the inability to withstand tainted photos is their biggest flaw. A hybrid algorithm using fuzzy c means is proposed in [16]. This total generalized variation fuzzy C means (TGVFCMS) algorithm identifies and separates the diseased areas from the input image and uses a CNN for disease classification. [17] Describes a combination of k mean clustering with super pixel clustering to segment the diseased color image. The results shows that the suggested model outperforms the other state-of-the-art techniques. [18] Proposes a particle swarm optimization algorithm to segment the diseased parts from sunflower leaf image dataset. The results were obtained with less computational efforts but the increased search time is a major limitation in these types of segmentation models. [19] Uses an HSV color space to segment the tomato leaf images. Here, segmenting just the leaf portions of the image before feature extraction shows that training time may be reduced in half and results can be more accurately produced. Whereas, in [20] k-mean clustering based segmentation is performed to separate the diseased areas from the healthy parts of the leaves. This helps in extracting the features accurately which is proved using the experimental results.

Leaf image segmentation models using advanced methods: These are the segmentation techniques that use deep learning and other powerful algorithms to accurately generate output from complex models. They use large datasets to train models that can learn complex patterns and features in leaf images which lead to accurate segmentation results. In comparison to traditional methods, these cutting-edge techniques offer more sophisticated and precise strategies for leaf image segmentation. In various research and application domains, they provide accurate analysis and characterization of leaves since they can manage complicated leaf shapes, variations in appearance, and difficult backgrounds. A deep learning algorithm Mask- R CNN is used in [21] for segmentation of multiple overlapping leaf regions from complicated backgrounds. Here the ground truth data has been created artificially using the foreground and background pixels. The algorithm's performance was assessed using Misclassification Error (ME), and it performed better than more widely used, effective algorithms like the otsu algorithm and grabcut. [22] Proposes a modified U-net architecture based on multi-scale feature extraction. The main drawback here is the manual creation of ground truth mask using an application. A CNN based architecture is implemented in [23] to effectively segment and classify the diseases affecting cotton plant leaves. As a result of less dataset many data augmentation methods were used

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to enhance the efficiency of the model. Utilization of more advanced methods can avoid such complexities.[24] Proposes a model that uses conditional random fields to upgrade the segmentation performance of the SegNet architecture. The conditional random fields make the SegNet network sharp which eliminates the issue of low accuracy. In [25] an attention mechanism is fused to the DeeplabV3+ network to improve the segmentation accuracy and the overall performance of the model. They also have done a comparative study to strengthen their points. Authors of [26], [27] and [28] also discuss 4 different types of modifications to the Mask RCNN architecture. These improvements are made at the backbone networks and experimental results depicts the high performance of the models.

3. MATERIALS AND METHODS

3.1. Image Dataset

In this work, a set of 1500 images of different crop leaves are used. Few of the crop leaf images are collected from online repositories such as Kaggle and TensorFlow. The others are taken using phone camera from different farmlands. This variation in the sources of dataset is mainly done to show the difference in segmenting both sets of images. Generally, the images that are collected from the online repositories are easier to preprocess and segment accurately as they are uploaded for such similar purposes. These types of images contain less noise and don't contain much distraction in the background. Therefore, using simple preprocessing techniques we can clean them easily. We can also remove the background from such images without any difficulties. Whereas, the offline images are of different resolutions and they contain a lot of other things along with the leaf parts. They are also prone to noises and other distortions. A large amount of resizing, filtering, de-noising etc has to be done in the case of offline image datasets. The given figure 1 shows few sample images of the input dataset.

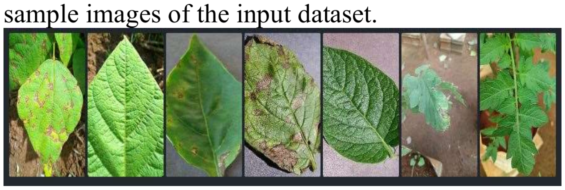

Figure 1. Sample Leaf Image Dataset

3.2. Image Pre-processing

Preprocessing of data is one of the basic steps in an automated detection and classification system. It is the process of preparing the dataset by performing a set of techniques and operations to make it suitable for further processes. The main aim of preprocessing is to improve the quality of the dataset which can influence the efficiency of the system. There are many preprocessing techniques such as filtering, resizing, noise removal, contrast enhancement etc and the usage of these techniques vary from one model to another as it depends largely on the assigned task and dataset characteristics. In this paper we are using a median filtering technique and otsu threshold to preprocess the dataset which is discussed in the following subsections.

3.2.1. Median Filtering

A median filtering technique is adapted to remove the unwanted noise from the input images. It is a non linear filtering method which is commonly used as a preprocessing technique to clean the input image without any data loss. The widespread usage of median filtering is due to its ability to protect the edges while removing the noise. The working of a median filter is by moving from one pixel value to the adjacent pixel value in a digital image while replacing the pixel values with the median value of its neighboring pixels [29]. A pattern known as window is used to calculate the median value of the neighboring pixels. This window can be a matrix of 3x3 and it slides over the entire digital image, replacing each value with the median of its neighboring pixel values. Here, the pixel values of the input images that contain noise are replaced with the median values. The given figure 2 shows the input images after performing preprocessing techniques such as de-noising and resizing.

Figure 2. Input images after performing median filtering and resizing.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.2.2. Otsu Threshold Algorithm

This is one of the widely used conventional segmentation algorithm which works based on a given threshold value [30]. This threshold value divides pixel values of the input image into two classes; foreground and background. As mentioned in the previous section, we have both infield and online images. The images which are collected directly from the farmlands are more complex and different from the images that are taken from online repositories even after undergoing various preprocessing techniques. Figure 3 shows the input images after performing otsu threshold algorithm. The entire background is removed only in image 3 as it belongs to the dataset taken from the online repositories. Therefore, much advanced algorithms are required to enhance the accuracy of segmentation on real time images as it highly depend on factors such as image quality, lighting, image clarity, focus etc.

Figure 3: Images after Otsu segmentation

3.3. GrabCut Algorithm

GrabCut is an effective image segmentation algorithm that uses rectangular bounding boxes to separate foreground objects from the background of the input image. Here the pixels that are outside the bounding box will be considered as background. GrabCut is a type of graph-cut technique in which the end user has to choose the Region of Interest (RoI) by labelling either a bounding box or lazzo on the image. This technique has a reduced user interaction when compared to that of the original graph cut technique which is the result of two enhancements: iterative estimation and incomplete labeling. Initially the algorithm uses Gaussian Mixture Model (GMM) to estimate the color distribution of the foreground and background. It is nothing but a probabilistic model that sums the weighted distributions of various Gaussian distributions to describe the probability distribution of a dataset. Here the GMM uses two components to represent both the foreground and background. Inorder to have a flexible modeling of correlation between the features, each Gaussian component has its own covariance matrix with k components which represent a separate cluster in the data. An extra k vector, k = {k_1,….,k_n,…..k_N} is added with, k_n € {1,…K} to make the GMM computationally manageable. This assigns a distinctive GMM element to each pixel in the image, where one component will be either from background or foreground model based on $\propto_n = 0 \; or \; 1^1$. Here, if $\propto_n$ equals 0, it means the pixel is assigned to the background model, and if $\propto_n$ equals 1, it means the pixel is assigned to the foreground model.

Based on the GMM component variables k, Gibbs energy is defined using T, the smoothness term and data item U, as given below:

$$E(\alpha,k,\theta,z) = U(\alpha,k,\theta,z) + T(\alpha,z) \quad (1)$$

The data point U determines the pixel label α by using k which is the Gaussian variable, to z the image data and grey histogram θ. This can be modified using Eq. (2)

$$U(\alpha,k,\theta,z) = \sum_n D(\alpha_n,k_n,\theta,z_n) = -\sum_n (\log p(z_n|\alpha_n,k_n,\theta) + \log \pi(\alpha_n,k_n)) \quad (2)$$

Here p is the Gaussian probability distribution of pixel z_n to the current pixel α_n, Gaussian pixel k_n andθ. π is regarded as mixed weight coefficient, and D is expressed as,

$$D(\alpha_n,k_n,\theta,z_n) = \frac{1}{2}\log det \sum(\alpha_n,k_n) - \log \pi(\alpha_n,k_n) + \frac{1}{2}[z_n - \mu(\alpha_n,k_n)]^T \sum(\alpha_n,k_n)^{-1} [z_n - \mu(\alpha_n,k_n)] \quad (3)$$

Gaussian model parameter is expressed as,

$$\theta = \{\pi(\alpha,k),\mu(\alpha,k),\sum(\alpha,k), \alpha = 0,1, k = 1,\dots,K\} \quad (4)$$

Here, the values such as μ and $\sum$ shows the mean and covariance of the 2K Gaussian components which is evaluated for the background and foreground distributions. Also the term for smoothness is defined as,

$$T(\alpha,z) = \gamma \sum_{(m,n\in C)} d_E\,(m,n)^{-1}[\alpha_m \neq \alpha_n]\exp\,(-\beta(z_m - z_n)^2) \quad (5)$$

Where γ is a constant, C is a collection of neighbouring pixel pairs and d_E represents the Euclidean distance of adjacent pixels.

The overall operation and significant steps of grabCut algorithm are discussed below:

a) Initialization:

The foreground and background GMM models are initialized using the label$\alpha_n = 0$ when the pixel $n \in T_B$ and $\alpha_n = 1$when the pixel doesn't belong toT_B.

b) GMM model processing:

The pixels values of target area are assigned and evaluated using the given equations:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$k_n = arg\min_{k_n} D_n(\alpha_n, k_n, \theta, z_n) \qquad (6)$$

$$\theta = arg\min_{\theta} U(\alpha, k, \theta, z) \qquad (7)$$

c) Graph creation:

Based on the data value and smoothness term an undirected graph is created using the labeled background and foreground pixels.

d) Segmentation:

Minimum cut algorithm can be used to solve the energy function's minimal value E, which is depicted in equation 8.

$$\min_{\{\alpha_n : n \in T_{CP}\}} \min_{k} E(\alpha, k, \theta, z) \qquad (8)$$

We have used an annotation tool *Labelbox* since the conventional grabcut algorithm requires manual interventions to create ground truth mask of the input image. This is a time consuming process as we have taken a dataset of 1500 images. Also it uses a rectangular box to extract the target information where the coordinate values are given manually. This again resulted in poor performance due to the variations in selecting background portion using the rectangular box. This is depicted in figure 4 and 5. The given figure 4 shows the online input image along with the ground truth mask. Even though the model employed a precise ground truth mask image, the data loss is clearly visible in the segmented output image. Some approaches that combine grabcut algorithm with conventional segmentation techniques like otsu threshold, K mean clustering, etc. have been presented to address this problem. These models produce results that are somewhat accurate. However, they might not function effectively in situations like enormous datasets with complicated images.

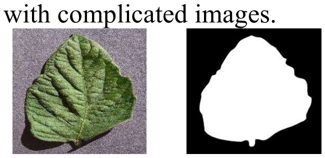

Figure 4: Input image and its corresponding ground truth mask.

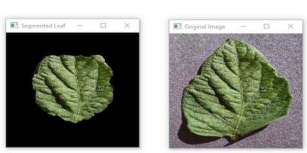

Figure 5: Segmented output image using original grabcut algorithm

Recent studies show that these types of complex images can be segmented or extracted using deep learning approaches. All the deep learning algorithms that perform well in leaf image segmentation process belong to the semantic segmentation category. Semantic segmentation is an image segmentation technique where each pixel is assigned a particular label such as fruit, flower, book etc. It is the most appropriate segmentation form that can be followed in the case of a leaf image dataset. Some effective deep learning based semantic segmentation algorithms are FCN, UNet, DeepLabV3+, Mask-R CNN, PSPNet etc. In this work, we will be using a deeplabV3+ algorithm to enhance the performance of the grabcut algorithm.

3.3.1. Deep learning based GrabCut Algorithm

In this paper, a deep learning based model is used to enhance the segmentation performance of the original GrabCut algorithm. Although, GrabCut is a powerful variant of Graph cut optimization algorithm it is a lengthy and tedious process as it requires high manual interactions. These shortcomings are fixed in the improved model using a deepLabV3+ network. The block diagram of the proposed method is given here..

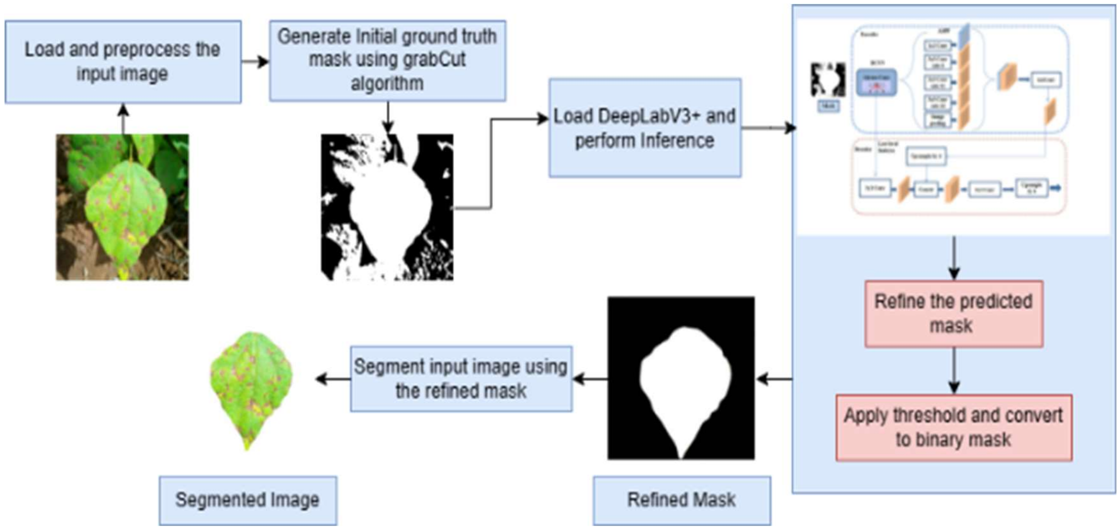

Figure 6: Working of deep learning based enhanced grabCut algorithm

3.3.2. DeeplabV3+ model

A deeplabV3+ model is an efficient variant of deeplabV3 semantic segmentation architecture. The modification is made by adding a simple decoder module to the deeplabV3 architecture inorder to improve the segmentation accuracy. This encoder – decoder model uses a ResNet50 backbone network to perform the initial feature extraction process. The given figure 7 shows the architecture of a deepLabV3+ network with a ResNet50 backbone. ResNet50 is deep convolutional neural network architecture with 50 layers that capture complex features and enable better performance in tasks such as semantic segmentation, object detection and classification. Another advantage of ResNet50 is the presence of skip connections that eliminate the problem of vanishing gradient during training. This is done by transmitting the information from the earlier layers directly through the skip connections to the following layers which results in effective training of the neural networks.

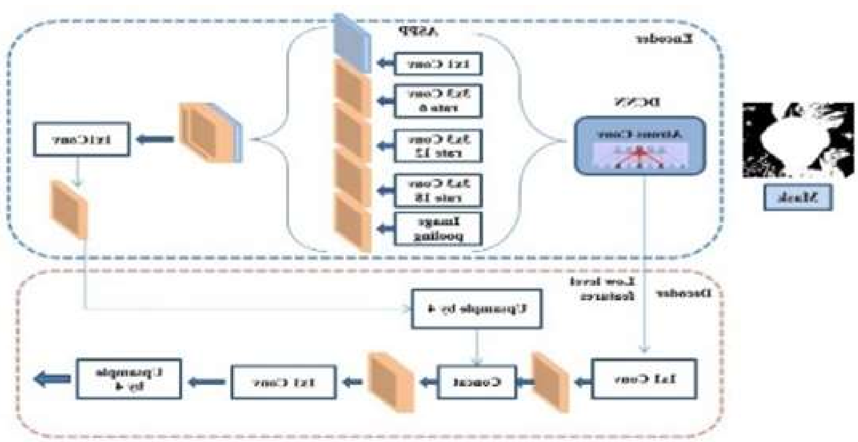

Figure 7: Architecture of DeepLabV3+

Journal of Theoretical and Applied Information Technology
15th October 2024. Vol.102. No. 19
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.3.3. Enhancing GrabCut using DeeplabV3+

The proposed model utilises a DeepLabV3+ network to improve the efficiency of the grabCut algorithm. Here, two possible modifications are done to the original grabCut algorithm. Firstly, this algorithm is modified to automatically generate the ground truth mask from the given input image instead of loading the manually prepared ground truth mask using annotation tool. Once this is done successfully, the generated imperfect mask is loaded to a deep learning model to process it into a refined mask. This is then used to segment the images accurately. The pseudocode of the proposed model is given below:

Pseudocode:

Input: An RGB colored leaf image
Output: Segmented image

1. The necessary libraries such as cv2, numpy, torch, torchvision.transforms are imported.
2. Input image is loaded using cv2.imread() and resized using cv2.resize()
3. An Initial ground truth mask is generated automatically using cv2.grabCut()
4. The background and foreground models are initialized and reshaped.
5. A pre-trained deepLabV3+ model (deeplabv3plus_resnet50) is loaded to refine the initial ground truth mask generated using grabCut algorithm.
6. The model performs inference on the mask and the predicted mask is generated using torch.argmax(). And then it is converted to a numpy array.
7. The predicted mask is refined by replacing the pixel values that are equal to 1 with 255 and the rest with 0.
8. Now this refined mask is applied to the image using cv2.bitwise_and()
9. Display the segmented output image using cv2.imshow()

4. RESULTS AND DISCUSSIONS

4.1. Experimental Setup

NVIDIA T4 Tensor Core GPU was utilized to perform segmentation on algorithms such as segnet, UNet and proposed method with python programming as the programming language and pytorch as the framework. All these algorithms were tested on two datasets namely: leaf images from PlantVillage dataset and similar manually collected leaf images. Initially, the images are preprocessed and resized to undergo segmentation. Number of iterations and the threshold value are the parameters that are considered in the case of traditional segmentation methods such as Otsu threshold, Otsu threshold and GrabCut algorithm and Grabcut algorithm. The parameters used in neural network based algorithms such as SegNet, UNet and the proposed method are given in Table 1.

Table 1: Parameter comparison in algorithms

Algorithm	Backbone	Inference speed	Epoch	Batch size
SegNet	ResNet50	0.0287	50	16
UNet	ResNet50	0.0355	50	16
Proposed method	ResNet50	0.0215	50	16

4.2. Discussion

This section explains the various experimental works done to evaluate the performance of the model. A 2x2 confusion matrix is used to evaluate the predictions of the generated output mask of the proposed model to that of the actual ground truth mask. The True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) values generated from the confusion matrix are used to evaluate the assessment parameters such as Pixel Accuracy (PA), recall, precision, F1 score and Intersection of Union (IoU). Using these parameters we compare the values with the existing models such as otsu threshold, combination of otsu threshold and grabcut, original grabcut algorithm and with few deep learning techniques that are commonly used for semantic segmentation such as FCN, SegNet and UNet architectures. The visualization of outputs for each model is depicted in the given figure 8.

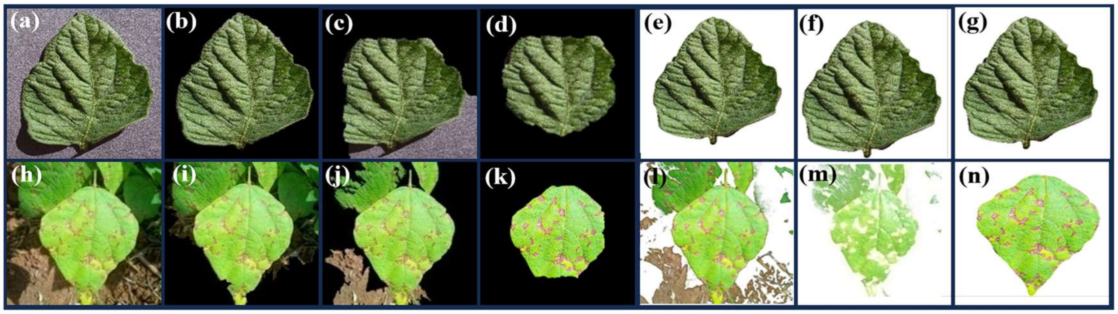

Figure 8: (a) Sample input data taken from Plant village dataset and correponding output images obtained after segmentation using (b) Otsu Threshold algorithm (c) Otsu threshold and GrabCut algorithm (d) Grabcut algorithm (e) SegNet (f) UNet (g) Proposed method; (h)Real time image of sample input data and corresponding output images obtained after segmentation using (i) Otsu Threshold algoritham (j) Otsu threshold and GrabCut algorithm (k) Grabcut algorithm (l) SegNet (m) UNet (n) Proposed method

Here, we are evaluating the performance of the model using a 2x2 confusion matrix which is given in figure 9. The TP, FP, FN and TN values obtained here are 71520, 4085, 4206, 126797 respectively. True positive (TP) the number of leaf pixels that are correctly identified as it is. False Positive (FP) depicts the number of background pixels that are mistakenly identified as leaf pixels. False negative (FN) are the leaf pixels incorrectly identifed as background pixels and True Negative (TN) is the number of background pixels correctly identifed as it is. Using these values we can calculate the following assessment parameters:

a) Pixel Accuracy:

This metric gives the count of accurately classified pixels when compared to the ground truth in the output image of segmentation model. It is calculated using the given formula,

$$\text{Pixel Accuracy} = (TP + TN) / (TP + FP + FN + TN)$$

$$= (71520 + 126797) / (71520 + 4085 + 4206 + 126797) \approx 0.9599$$

b) Intersection of Union (IoU):

It is also known as Jaccard Index and compares the ratio between the intersection area to the union area of the predicted segmentation mask and ground truth mask to calculate the accuracy of the segmentation model. It is defined by,

$$\text{IoU} = TP/(TP + FP + FN)$$

$$= 71520 / (71520 + 4085 + 4206) \approx 0.9709$$

c) Precision:

It finds the ratio between the true positive values of the confusion matrix which are nothing but the correctly identified pixels to all the positive values. the formula to calculate precision is given here:

$$\text{Precision} = TP / (TP + FP) = 71520 / (71520 + 4085) \approx 0.9460$$

d) Recall:

This metric is known as true positive rate and it focuses on the ability of the model to find all positive values out of all the positive values. It is calculated as,

$$\text{Recall} = TP / (TP + FN) = 71520 / (71520 + 4206) \approx 0.9475$$

a) F1 Score:

It is the average obtained from precision and recall. The formula of F1 score is given as,

$$\text{F1Score} = 2 * (Precision * Recall) / (Precision + Recall) = 2 * (0.9460 * 0.9475) / (0.9460 + 0.9475) \approx 0.9478$$

These values are evaluated for each model and compared with the proposed model. The table 2 shows the comparative analysis to evaluate the strength and weaknesses of the proposed algorithm using PMI(Plus Minus Interesting Facts) framework. And the given table 3 depicts the assessment parameter values in percentage and we have also used a graphical representation of the same in figure 10. Another comparison of accuracy curves is given in figure 11, which is based on the Epoch values and pixel accuracy after training the algorithms. All these values and calculations point out the fact that our proposed model shows better evaluation results than all the other segmentation models which are taken here.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Comparative analysis using PMI framework

Segmentation Techniques	Plus	Minus	Interesting (Facts/Observations)
Otsu Threshold	This technique is simple and computationally efficient while compared to the other methods.	Lowest performance: IoU (72.02%), Pixel Accuracy (68.77%).	Struggles with complex images, highlighting its limitations for high-level segmentation tasks.
Otsu threshold and GrabCut	Strong performance: IoU (93.87%), Pixel Accuracy (90.31%). Improves results of standalone Otsu Threshold significantly.	Slightly lower than deep learning-based methods like Unet and the proposed method.	A hybrid method that performs surprisingly well, showing that combining basic techniques can yield competitive results.
GrabCut algorithm	Decent performance: IoU (85.37%), Pixel Accuracy (86.21%).	Falls behind deep learning models in handling complex image segmentation.	Shows that graph-based methods, while decent, are less effective than deep learning models in complex segmentation scenarios.
SegNet	Moderate performance with IoU (74.72%) and Pixel Accuracy (73.66%).	Lower performance compared to Unet and the proposed method.	Highlights the limitations of earlier deep learning models when compared to more advanced architectures.
Unet	Excellent performance across metrics: IoU (94.38%), Pixel Accuracy (94.28%), F1 Score (93.57%).	Slightly behind the proposed method in accuracy and boundary adherence.	Proves the strength of deep learning-based architectures, but shows that further refinements, as seen in the proposed method, can yield even better results.
Proposed method	Best performance: IoU (97.07%), F1 Score (94.78%), Pixel Accuracy (95.99%). Combines strengths of graph-based and deep learning methods for superior accuracy.	Computationally more demanding due to the integration of multiple algorithms.	Combining enhanced GrabCut with DeepLabV3+ shows novel improvements, making it highly effective for both complex and simple segmentation tasks.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

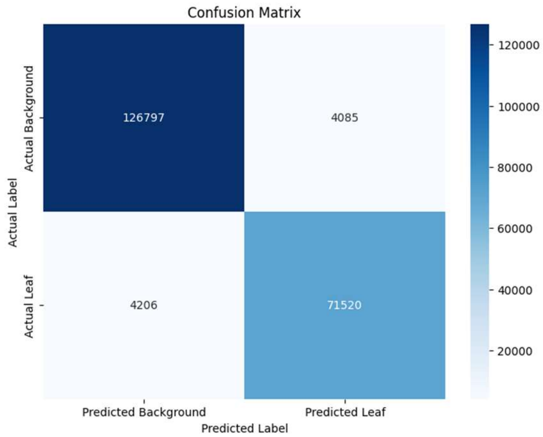

Figure 9: Confusion matrix of the proposed method

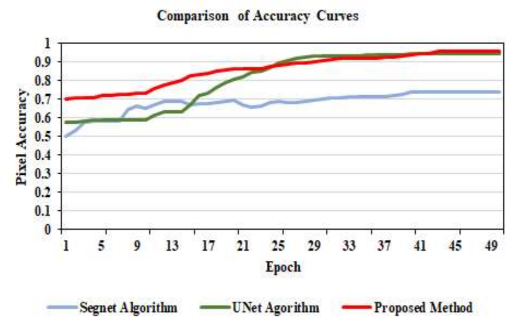

Figure 11: Comparison of Accuracy Curves based on Epoch value

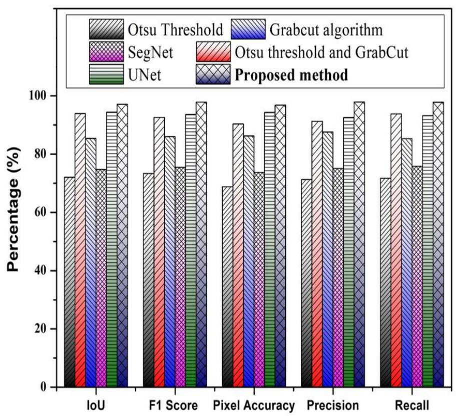

Figure 10: Graphical representation of the segmentation accuracies of different model

Table 3: Comparison between proposed method and other segmentation model

Segmentation techniques	Assessment parameters				
	IoU (%)	F1 Score (%)	Pixel Accuracy (%)	Precision (%)	Recall (%)
Otsu Threshold	72.02	73.28	68.77	71.23	71.67
Otsu threshold and GrabCut	93.87	92.56	90.31	91.20	93.78
GrabCut algorithm	85.37	86.01	86.21	87.54	85.26
SegNet	74.72	75.40	73.66	75.01	75.82
Unet	94.38	93.57	94.28	92.53	93.19
Proposed method	**97.07**	**94.78**	**95.99**	**94.60**	**94.75**

5. CONCLUSION

Segmentation is an important method that adds substantially to image processing and computer vision. A wide range of segmentation techniques are available today due to its effectiveness and simplicity. In this paper a hybrid approach is proposed to enhance the segmentation accuracy of the GrabCut algorithm. The original GrabCut algorithm is automated to generate masks which are then refined using a deeplabV3+ model to produce more accurate masks. The efficacy of this hybrid model is assessed by comparing it to a few powerful current segmentation models. The suggested method successfully overcame the limitations of conventional segmentation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

techniques and enhanced the overall accuracy of leaf image segmentation by using the capabilities of deep learning and integrating it into the GrabCut algorithm.

The findings of this work are of vital significance in today's environment, where segmentation accuracy is crucial for applications such as plant disease detection, autonomous systems, and medical imaging. The suggested hybrid approach not only increases segmentation performance, but it also provides a scalable solution that combines the effectiveness of classical methods with the resilience of deep learning. This work paves the way for future research on hybrid models that can adapt to a variety of image segmentation issues across domains.

AUTHOR CONTRIBUTIONS

Dr. P. J. Arul Leena Rose: Conceptualization, Validation, Writing – review & editing, Supervision.

Sreya John: Conceptualization, Methodology, Formal analysis and investigation, Writing – original draft.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

[1] Saranya, T., Deisy, C., Sridevi, S., & Anbananthen, K. S. M. (2023). A comparative study of deep learning and Internet of Things for precision agriculture. *Engineering Applications of Artificial Intelligence*,*122*,106034. https://doi.org/10.1016/j.engappai.2023.106034

[2] Yu, H., Song, J., Chen, C., Heidari, A. A., Liu, J., Chen, H., & Mafarja, M. (2022). Image segmentation of Leaf Spot Diseases on Maize using multi-stage Cauchy-enabled grey wolf algorithm. *Engineering Applications of Artificial Intelligence*, *109*, 104653. https://doi.org/10.1016/j.engappai.2021.104653

[3] Kim, T., Kim, H., Baik, K., & Choi, Y. (2022). Instance-Aware Plant Disease Detection by Utilizing Saliency Map and Self-Supervised Pre-Training. *Agriculture*,*12*(8),1084. https://doi.org/10.3390/agriculture12081084

[4] Kumar, A., & Patel, V. K. (2023). Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. *Multimedia Tools and Applications*, 1-27. – median filter https://doi.org/10.1007/s11042-023-14663-z

[5] Wang, Z., Lv, Y., Wu, R., & Zhang, Y. (2023). Review of GrabCut in Image Processing. *Mathematics*, *11*(8), 1965. https://doi.org/10.3390/math11081965

[6] Tassis, L. M., de Souza, J. E. T., & Krohling, R. A. (2021). A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images. *Computers and Electronics in Agriculture*, *186*, 106191. – coffee code https://doi.org/10.1016/j.compag.2021.106191

[7] Zhang, S., & Zhang, C. (2023). Modified U-Net for plant diseased leaf image segmentation. *Computers and Electronics in Agriculture*, *204*, 107511. https://doi.org/10.1016/j.compag.2022.107511

[8] Luo, Z., Yang, W., Yuan, Y., Gou, R., & Li, X. (2023). Semantic segmentation of agricultural images: A survey. *Information Processing in Agriculture*. https://doi.org/10.1016/j.inpa.2023.02.001

[9] Zhu, S., Ma, W., Lu, J., Ren, B., Wang, C., & Wang, J. (2023). A novel approach for apple leaf disease image segmentation in complex scenes based on two-stage DeepLabv3+ with adaptive loss. *Computers and Electronics in Agriculture*, *204*, 107539. https://doi.org/10.1016/j.compag.2022.107539

[10] Lu, J., Xiang, J., Liu, T., Gao, Z., & Liao, M. (2022). Sichuan Pepper Recognition in Complex Environments: A Comparison Study of Traditional Segmentation versus Deep Learning Methods. *Agriculture*, *12*(10), 1631. https://doi.org/10.3390/agriculture12101631

[11] Picon, A., San-Emeterio, M. G., Bereciartua-Perez, A., Klukas, C., Eggers, T., & Navarra-Mestre, R. (2022). Deep learning-based segmentation of multiple species of weeds and corn crop using synthetic and real image datasets. *Computers and Electronics in Agriculture*, *194*, 106719. https://doi.org/10.1016/j.compag.2022.106719

[12] Wang, F., Rao, Y., Luo, Q., Jin, X., Jiang, Z., Zhang, W., & Li, S. (2022). Practical cucumber leaf disease recognition using improved Swin Transformer and small sample size. *Computers and Electronics in Agriculture*, *199*, 107163. https://doi.org/10.1016/j.compag.2022.107163

[13] Dutta, K., Talukdar, D., & Bora, S. S. (2022). Segmentation of unhealthy leaves in cruciferous crops for early disease detection

using vegetative indices and Otsu thresholding of aerial images. *Measurement*, *189*, 110478. https://doi.org/10.1016/j.measurement.2021.110478

[14] Trivedi, V. K., Shukla, P. K., & Pandey, A. (2022). Automatic segmentation of plant leaves disease using min-max hue histogram and k-mean clustering. *Multimedia Tools and Applications*, *81*(14), 20201-20228. https://doi.org/10.1007/s11042-022-12518-7

[15] Zhang, H., Li, H., Chen, N., Chen, S., & Liu, J. (2022). Novel fuzzy clustering algorithm with variable multi-pixel fitting spatial information for image segmentation. *Pattern Recognition*, *121*, 108201. https://doi.org/10.1016/j.patcog.2021.108201

[16] Krishnan, V. G., Deepa, J. R. V. P., Rao, P. V., Divya, V., & Kaviarasan, S. (2022). An automated segmentation and classification model for banana leaf disease detection. *Journal of Applied Biology and Biotechnology*, *10*(1), 213-220. http://dx.doi.org/10.7324/JABB.2021.100126

[17] Zhang, S., Wang, H., Huang, W., & You, Z. (2018). Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. *Optik*, *157*, 866-872. https://doi.org/10.1016/j.ijleo.2017.11.190

[18] Singh, V. (2019). Sunflower leaf diseases detection using image segmentation based on particle swarm optimization. *Artificial Intelligence in Agriculture*, *3*, 62-68. https://doi.org/10.1016/j.aiia.2019.09.002

[19] Nguyen, T. H., Nguyen, T. N., & Ngo, B. V. (2022). A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease. *AgriEngineering*, *4*(4), 871-887. https://doi.org/10.3390/agriengineering4040056

[20] Javidan, S. M., Banakar, A., Vakilian, K. A., & Ampatzidis, Y. (2023). Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. *Smart Agricultural Technology*, *3*, 100081. https://doi.org/10.1016/j.atech.2022.100081

[21] Yang, K., Zhong, W., & Li, F. (2020). Leaf segmentation and classification with a complicated background using deep learning. *Agronomy*, *10*(11), 1721. https://doi.org/10.3390/agronomy10111721

[22] Ngugi, L. C., Abdelwahab, M., & Abo-Zahhad, M. (2020). Tomato leaf segmentation algorithms for mobile phone applications using deep learning. *Computers and Electronics in Agriculture*, *178*, 105788. https://doi.org/10.1016/j.compag.2020.105788

[23] Singh, P., Singh, P., Farooq, U., Khurana, S. S., Verma, J. K., & Kumar, M. (2023). CottonLeafNet: cotton plant leaf disease detection using deep neural networks. *Multimedia Tools and Applications*, 1-26. https://doi.org/10.1007/s11042-023-14954-5

[24] Yue, Y., Li, X., Zhao, H., & Wang, H. (2020, October). Image segmentation method of crop diseases based on improved SegNet neural network. In *2020 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 1986-1991). IEEE. https://doi.org/10.1109/ICMA49215.2020.9233609

[25] Cai, M., Yi, X., Wang, G., Mo, L., Wu, P., Mwanza, C., & Kapula, K. E. (2022). Image segmentation method for sweetgum leaf spots based on an improved DeeplabV3+ network. *Forests*, *13*(12), 2095. https://doi.org/10.3390/f13122095

[26] Wang, D., & He, D. (2022). Fusion of Mask RCNN and attention mechanism for instance segmentation of apples under complex background. *Computers and Electronics in Agriculture*, *196*, 106864. https://doi.org/10.1016/j.compag.2022.106864

[27] Shen, L., Su, J., Huang, R., Quan, W., Song, Y., Fang, Y., & Su, B. (2022). Fusing attention mechanism with Mask R-CNN for instance segmentation of grape cluster in the field. *Frontiers in plant science*, *13*, 934450. https://doi.org/10.3389/fpls.2022.934450

[28] Cong, P., Li, S., Zhou, J., Lv, K., & Feng, H. (2023). Research on instance segmentation algorithm of greenhouse sweet pepper detection based on improved mask RCNN. *Agronomy*, *13*(1), 196. https://doi.org/10.3390/agronomy13010196

[29] Janani, M., & Jebakumar, R. (2023). Detection and classification of groundnut leaf nutrient level extraction in RGB images. *Advances in Engineering Software*, *175*, 103320. https://doi.org/10.1016/j.advengsoft.2022.103320

[30] Liu, L. (2023). Disease spot image segmentation algorithm with memory-based fruit fly optimization algorithm. *Multimedia Tools and Applications*, 1-29. https://doi.org/10.1007/s11042-023-15630-4