Computers and Electronics in Agriculture 157 (2019) 102–109

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

SLIC_SVM based leaf diseases saliency map extraction of tea plant

Yunyun Sun[a], Zhaohui Jiang[a,*], Liping Zhang[b], Wei Dong[b], Yuan Rao[a]

[a] School of Information and Computer Science, Anhui Agricultural University, Hefei 230036, China
[b] Institute of Agricultural Economics and Information, Anhui Academy of Agricultural Sciences, Hefei 230036, China

ARTICLE INFO

Keywords:
Tea plant
Leaf disease
SLIC
SVM
Saliency map

ABSTRACT

For the purpose of improving the extraction of tea plant leaf disease saliency map under complex backgrounds, a new algorithm combining SLIC (Simple Linear Iterative Cluster) with SVM (Support Vector Machine) is proposed in this paper. Firstly, super-pixel block is obtained by SLIC algorithm, significant point is detected by Harris algorithm, and fuzzy salient region contour is extracted by employing convex hull method. Secondly, the four-dimensional texture features of super-pixel blocks in salient regions and background areas are extracted, and then the classification map is obtained by classifying the super-pixel blocks with the help of SVM classifier. Lastly, the morphological and algebraic operations are implemented for repairing classified super-pixel blocks. As a result, one accurate saliency map of tea plant leaf disease image is obtained. Through testing based on 261 diseased images, the quality evaluation index, the accuracy, precision, recall and F-value are 98.5%, 96.8%, 98.6% and 97.7%, respectively. It demonstrates that the proposed method performs better than the other three SLIC-based algorithms in visual effects and quality assessment index. Such conclusion can be drawn that the proposed method can effectively extract tea plant leaf disease saliency map from complex background. Consequently, this research is expected to lay a good basis for the study of tea plant leaf disease identification. Last but not the least, the proposed method has good potential that extracts saliency map of crops or plants disease.

1. Introduction

Tea plant is an important economic crop. It is capable of providing the function of health care, hygienism, environmental protection and safety (Ghosh et al., 2013; Cabrera et al., 2006). It is well known that tea is one of the closest to natural green healthy drinks, and tea industry belongs to the group of green industry. China is the original area of tea plants and has a long history in tea production. All over the world, about half of tea orchards are located in China with the total area of 35.29 million square hectometers, which covers 18 provinces of China (Majumder, 2011). Unfortunately, there are various diseases, insects, grasses and other harmful factors in the process of planting and production of tea plant, which seriously affects tea production and quality and causes loss of economic benefits. The tea plant diseases have reached about 380 types all over the world. About one hundred kinds of tea plant diseases occur in China, more than 40 species of them are common. Leaf disease accounts for a large proportion of tea diseases. Usually, the yield reduction caused by leaf diseases of tea plants is not noticeable, it is difficult to grasp the degree of yield loss through direct observation (Gao et al., 1996; Chen et al., 1982). Detect tea plant leaf diseases accurately is an extremely important task during tea cultivation. Machine vision is the most common disease detection method (Ma et al., 2017).

ROI (Region of Interest)-based saliency map extraction always plays an important role in image processing and analysis (Cheng et al., 2015). The extraction method of threshold, edge and region images based is popularity, Yanan Ma segmented the ROI of hyperspectral images with 98.8% correct rate by automatic threshold method (Ma et al., 2014). However, the method ignores image space and textures features. Therefore, if the gray values difference of image is not obviously, the effect is poor (Otsu, 2007). Edge is the basic feature of image, which is the junction between an attribute and another attribute area, the traditional edge detection algorithm is Prewitt, Sobel and Canny, etc. Mathew J have extracted crop diseases saliency map by edge detection (Mathew et al., 2015). However, the problems of edge discontinuity, edge information loss or edge blur etc. in the division of complex images edge is common, it is difficult to ensure the integrity of boundary information (Ziou et al., 1998). Based on the principle of pixel color similarity in the same area, determine a seed area firstly, and then the target area is continuously merged with similar pixels in the seed area to realize target segmentation is the process of region-based segmentation algorithm (Yang et al., 2008; Zhao et al., 2011). However, if you want to obtain a good segmentation results, you should combine this method with other algorithms, but the complexity also increased (Bischof, 2002). Therefore, these methods are not suitable for the extraction of leaf disease of tea plant saliency map in complex background.

* Corresponding author.
E-mail address: jiangzh@ahau.edu.cn (Z. Jiang).

https://doi.org/10.1016/j.compag.2018.12.042
Received 11 April 2018; Received in revised form 18 December 2018; Accepted 23 December 2018
Available online 27 December 2018

As we know, it helps in greatly reducing the complexity of image processing when grouping pixels based on similarity among pixel features, then super-pixel blocks is used to represent lots of pixels. Usually, this method is employed as a preprocessing step of the segmentation algorithm. The simple linear iterative algorithm (SLIC) is one of the super-pixel segmentation algorithms, which has simple theory and is easy to be understood. The advantages of high operating efficiency and flexible parameter adjustment results in its being widely used as one good saliency map extraction candidate (Ren and Malik, 2003). Fu K combined SLIC algorithm with normalized pattern cutting (Ncut) for the detection of significant object frame, which significantly improved detection accuracy and effectively suppressed the background (Fu et al., 2015). B J used Markov chain for saliency detection after the SLIC's processing, and used equilibrium distribution in traversal Markov chain to reduce the absorption time of smooth background regions (Bowen Jiang et al., 2013). Moreover, they conducted extensive tests based on four benchmark data sets, which demonstrated that the proposed method has better robustness and efficiency. Lu H proposed a region-based color modelling for joint crop and maize tassel segmentation. They used graph-based segmentation algorithms and SLIC to generated salience regions. More specifically, they proposed to model colors with ensemble neural networks specific to achieve robustness to illumination, finally, the proposed method achieved corn tassel saliency map extraction efficiently and rapidly (Lu et al., 2016).

SVM is an excellent learning and classification method as claimed by Vniknik (Vapnik et al., 2002). It has the several well-known advantages, such as, high precision, fast speed, good robustness and strong comprehensiveness etc. especially, it show unique advantages in solving small sample classification, nonlinear and high-dimensional, and was widely used in saliency map extraction (Cortes et al., 1995). Prominent object locations were estimated by using Trimap, then the whole salient object were segmented by the trained SVM classifier (Bai et al., 2014). ROA (region of attention) were created based on the saliency map and corner points, subsequently, classifier trained through learning the color and texture information of objects, finally, saliency map was precisely segmented by FSVM (Zhao et al., 2010). It is apparent that SVM performed excellent when dealing with saliency map extraction. However, automatic selection of training set and parameters still need to be explored and improved continuously. The combination of SLIC and SVM has been conducted in a few fields. In medicine, it is discovered that SLIC and SVM based on classification method to extract saliency map is superior to QS both in terms of accuracy and efficiency (Maghsoudi, 2017). Wu Z attempted to detect remote sensing image unsupervised change using SLIC and SVM. And it has been shown that this combination method can accurately extract remote sensing saliency maps (Wu et al., 2012). Based on the above performances, the combination of SLIC and SVM is a novel method for extracting saliency map, and it is reliable and effective for saliency map extraction. Theoretically, it could perform well when applied to tea plant leaf disease saliency maps.

For the purpose of exploring the accurate extraction of tea plant leaf disease saliency map in complex background, it is worth evaluating the feasibility of SLIC_SVM in tea plant leaf disease saliency map extraction. The proposed algorithm goes like the following. Firstly, we divide the image into small blocks by using SLIC super-pixel algorithm, and then extract the salient region fuzzy contours. Secondly, according to the characteristics of salient region and the background region super-pixel blocks, automatically extracting feature vectors of sample texture is implemented from multiple directions to train SVM. Finally, the disease saliency map can be extracted.

2. Data acquisition and preprocessing

Tea plant leaf disease images were collected in multiple directions under the tea plantation environment. After being identified and classified them by three plant protection experts, 1308 samples with five common diseases including Tea anthracnose (Ta), Tea brown blight (Tbb), Tea netted blister blight (Tnbb), Exobasidium vexans Massee (Evm) and Pestalotiopsis theae (Pt) were randomly selected for experimentation. The image capture and disease classification were completed by Institute of Agricultural Economics and Information, Anhui Academy of Agricultural Sciences. More specifically, a digital SLR camera is directly used to collect disease images in the open field under the conditions of natural light and flash light. And the backgrounds of natural environments include plant leaves, soil and sky. Image processing software helps in conducting metadata indexing, cutting, manual classification, and others pretreatment preprocessing. All image samples were infected with varying degrees of disease, the type and distribution of disease image samples shows in Table 1.

Table 1
Number and kind of disease samples.

Disease type	Train sample set	Test sample set	Total
(T a) Tea anthracnose	240	119	359
(T bb) Tea brown blight	74	36	110
(T nbb) Tea netted blister blight	283	141	424
(E vm) Exobasidium vexans Massee	224	56	280
(P t) Pestalotiopsis theae	108	27	135
Total	1047	261	1308

3. Methodologies

3.1. SLIC_SVM algorithm

After going through the characteristics analysis of tea plant leaf images, super-pixel and classifier algorithm were combined for the extraction of saliency map. More specifically, super-pixel block algorithm extracted fuzzy contour of saliency area, subsequently classifier classifies salient region super-pixel blocks and background super-pixel blocks.

The process of fuzzy contour extraction and classification is illustrated in Algorithm1. In the initial step, five kinds of disease sample images (Ta, Tbb, Evm, Pt and Tnbb) were chosen. Next, images were divided by Simple Linear Iterative Cluster (SLIC) super-pixel algorithm, and then significant points were detected by Harris algorithm. It is apparent that Steps 2 and 3 are precedent process of carrying out Step 4. In step 4, fuzzy contour of images is extracted using the convex hull technique. These four steps constitute phase 1 whereas, the saliency objects fuzzy are generated. In Step 5, we identify the features to distinguish salient region and background super-pixel blocks and extract it by GLCM function. In Steps 6 and 7, the texture features (energy, entropy, contrast and correlation) of super-pixel blocks are used to train Support Vector Machine (SVM), when training was completed the saliency map super-pixel block can be separated from images. Last but not the least, high-quality RGB saliency map can be obtained using simple morphological processing and logical operators.

Algorithm 1 (*: Basic steps of SLIC_SVM algorithm***).**

1 Tea plant leaves disease image acquisition
2 Super-pixel algorithm to divide image into several bocks
3 Harris Algorithm detects significant point
4 The convex hull encloses significant point to draw the fuzzy salient contour
5 Calling the GLCM function to calculate the features
6 Build training data set
7 Use the SVM classify super-pixel block
8 Morphological processing and logical operators to obtain saliency map

3.2. Fuzzy contour extraction

As for extraction of salient region, edge detection is popular among the canonical methods. However, when the boundary is complex or blurry, it is difficult to accurately detect the boundary. Super-pixel algorithm has been widely used in computer vision, such as image segmentation, saliency detection, target detection and so on, the good performance of which in boundary contour extraction have been tested by scholar. Considering the aforementioned discussion, we decide to combine the super-pixel block, corner detection and convex hull algorithm to extract salient region super-pixel blocks fuzzy contour. Simple Linear Iterative Cluster (SLIC) algorithm is good at boundary division, and has short running time. As a result, it is the most suitable algorithm to extract boundary in all super-pixel block algorithm (Achanta et al., 2012).

SLIC was proposed by Achanta in 2012, which is one of the super-pixel segmentation algorithms, based on color and distance similarities (Achanta, 2010). SLIC implements clustering in *CIELAB* color space, it is required to set initial clustering number k before clustering. SLIC algorithm goes as follows. Suppose that $M*N$ is the size of image and i is the number of pixels. Initialize k cluster centers $Ci = [li\ ai\ bi\ xi\ yi]^T$, which contain color space $[l\ a\ b\ ]^T$ and pixel location $[x,\ y]^T$. Then sampling and generating approximately equal sized super-pixel blocks at a regular network with S ($S = \sqrt{\frac{M*N}{k}}$) pixels apart. Later, a similar pixel search is performed in a region of $2S \times 2S$ and centered on the cluster centers, the theory is shown in Fig. 1. The distance D of other cluster from this center is calculated using Eq. (1). As a result, associate pixel i to the nearest cluster center and update the cluster center. Subsequently, the new cluster center is determined as the average of updated class pixel vector $[l\ a\ b\ x\ y]^T$. Once the clustering is completed, the residuals is expected to be converged to a certain value. Now, $k0$ super-pixel blocks is obtained.

$$D = \sqrt{dc^2 + ds^2} \tag{1}$$

$$dc = \sqrt{(lm - ln)^2 + (am - an)^2 + (bm - bn)^2} \tag{2}$$

$$ds = \sqrt{(xm - xn)^2 + (ym - yn)^2} \tag{3}$$

where D is the distance of five-dimensional space (*labxy*), dc, ds which are color and position distances expressed in Euclidean Distance, were calculated by Eqns. (2) and (3), respectively. m, n is the pixel number.

Significant point detection is the first step of saliency map extraction (Wang et al., 2015). It is discovered that corner detection is more effective than other related methods. In order to minimize background interference, the abnormal values are adjusted using the average ± 3 times standard deviation of significant point. After the regular significant points were obtained, Graham scanning was used for searching minimum convex hull. Usually, salient point is on or within the convex hull. The algorithm is executed as follows:

Step1: Find the leftmost of significant point that has smallest y, find the significant point with smallest x instead if there are more than one smallest y;
Step2: Set P_0 as the origin of coordinate by coordinate translation;
Step3: Argument α is calculated by decreasing order according to Eq. (4). If there are equal α values, put those closer to P_0 in the front. Assume that we have ordered it in $[P0,\ P1, \cdots, P(n'-1)]$, the sorted result shown in Fig. 2.

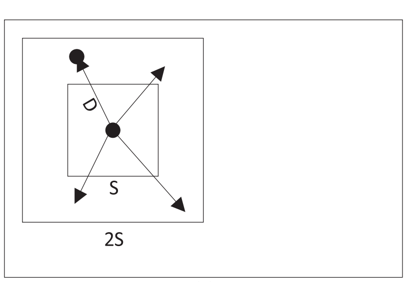

Fig. 1. SLIC clustering search graph.

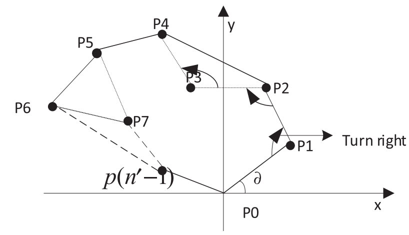

Fig. 2. Convex hull schematic diagram.

$$\partial = \arctan \frac{y}{x} \tag{4}$$

where x, y are the coordinates of significant point.

Step4: Create a stack with P_0 and P_1 in bottom, and then push $[P0,\ P1, \cdots, P(n'-1)]$ into the stack in order. If the two points on the top of the stack are not formed a 'turn to the left' relationship with P_0 (as shown in Fig. 2), push to stack until no point needs to be popped after the current point into the stack.
Step5: After completing the process of all the points, the points within stack are exactly the salient points on convex hull, and these points are also the points on fuzzy contour of salient region.

3.3. Accurate saliency map extraction

After the process of SLIC super-pixel block and corner detection, we can obtain the salient area contour. Subsequently, all training samples are imported into the SVM classifier to extract tea plant disease saliency map. For SVM-based saliency map extraction, the selection of training samples features has the greatest impact on the extraction results. For existing methods, manual selection or mark regional training samples is common, it causes that the segmentation result have great dependence on manual selected or marked samples. To avoid these problems, algorithm is designed for training sample features automatic and extract saliency map. During training, salient region training set is composed by super-pixel block feature data which in the contour, and background training set is composed by super-pixel block feature data which out of fuzzy contour. For the super-pixel blocks that is not completely in contour, classifying them as a part of background data set.

Feature extraction is conducted according to the characteristics of super-pixel block of images. For the images with significant color and texture difference, it is feasible to implement model training by extracting the color and texture of images. It is common extraction methods to extract the color and texture of images or combine them (Linker et al., 2012; David et al., 2010; Pourreza et al., 2012). In such experiment distinguishing salient region and background super-pixel block of tea plant leaf diseases, selecting texture features from Gray Level Co-occurrence Matrix (GLCM) algorithm as distinguishing features of significant regions and backgrounds (Ronge et al., 2014).

SVM classifier parameter selection includes classifier type and kernel function selection. Essentially, it is a dichotomy problem to apply support vector machine to significant image extraction. In SLIC_SVM saliency map extraction, Classifier type was C-SVC. C ($C > 0$) represents the penalty term, kernel function is polynomial kernel function, model architecture is shown in Fig. 3. The left is a part of the pixel blocks extracted by SLIC, where A is the salient region super-pixel blocks, B is the background super-pixel blocks, and they are separated by yellow outline. More specifically, number is the block serial number. $K(x, x')$ is the kernel function, $T = \{(x1, y1), \cdots (xi, yi), \cdots, (xl, yl)\}$ is the pixel block eigenvector set which to be trained, the right is the dichotomous output by SVM classifier.

The training set pixel block sample texture feature and label are input into SVM. Saliency map extraction step is as follows:

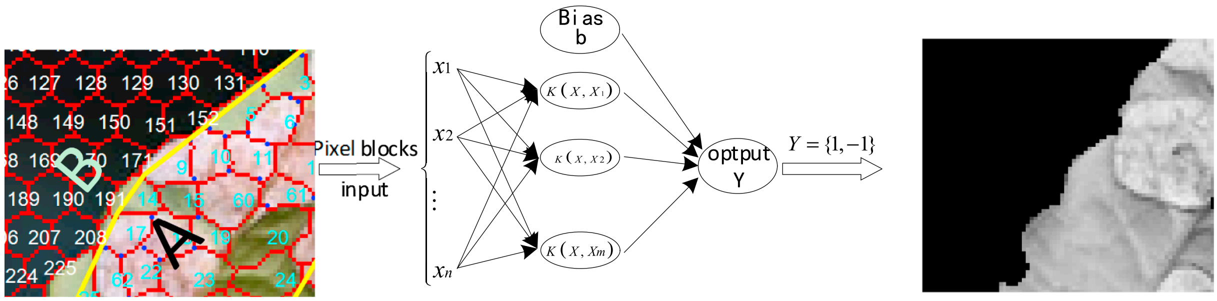

Fig. 3. SVM architecture diagram.

Step1: GLCM algorithm extracts 4-dimensional texture features of each super-pixel block, and then defines super-pixel block labels to form super-pixel block feature vectors.

Step2: Salient region and background training sample blocks automatically is selected based on pixel block serial numbers, then they are fed into SVM architecture for conducting model training. The calculation formula of polynomial kernel function is Eq. (5), the final classification image is output by Eq. (6).

Step3: Morphological operations and logic operations are implemented after SVM image classification for obtaining saliency map (Zapotoczny et al., 2008).

$$K(x, x') = (\gamma x^T xi + r)^d \tag{5}$$

$$yi \in Y = \begin{cases} 1 & \text{label } 1 \\ -1 & \text{label } 0 \end{cases} \tag{6}$$

When the classification is finished, the saliency map will be extracted from the image using the SLIC_SVM algorithm.where $i = 1, 2, \cdots, l$, l ($l \le k_0$) is the number of sample pixel blocks within training set. (xi, yi) is the eigenvector of the i th super-pixel block feature vector x_i and label vector y_i in the training data, the Y value is corresponding to the tag vector.

4. Results and analysis

4.1. Extraction process of saliency map

To verify the algorithm effectiveness, we extract the tea leaf sample saliency map by using SLIC_SVM algorithm, and show the test results of one image. Firstly, SLIC algorithm blocked the sample image. As shown in Fig. 4, the diseased image is divided into several small pixel blocks, and the two different zone boundaries is completely divided. Then, significant points are detected by Harris Corner Detection Algorithm. The detection results are shown in Fig. 5, in which the blue dots denote the detected significant points. Afterwards, convex hull envelops the significant points outlined the salient region contours. Now, automatic selection of SVM training set is achieved. The salient region contours is yellow curve in Fig. 5. Notice that zone A is the super-pixels blocks within salient region, and B represents the super-pixels blocks within the background area.

After the salient region fuzzy contour is obtained, training set is constructed through selecting and extracting pixel blocks underlying features. GLCM algorithm is used to extracted the texture features from the four directions $\theta = 0°, 45°, 90°, 135°$ and $d = 1$, including the energy, entropy, contrast, and correlation four features value and their average is used as the feature vector of the training data set, the feature data points are shown in Fig. 6. The red spot is zone a feature value, and the blue dot is the texture feature value of area B, the number of data points corresponds to the number of pixel blocks in zones A and B, respectively. From the density and amplitude of data points in distribution figure, we can quickly distinguish the difference of zones A and B, the differences of eigenvalues' distribution is especially obvious in Fig. 6(c) and (d).

Fig. 4. SLIC super-pixel block image.

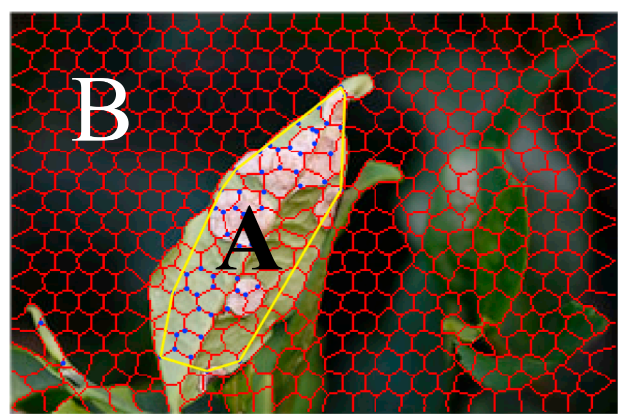

Fig. 5. Salient region fuzzy contour.

After the feature values extraction of training sample blocks of pixels in areas A and B, the feature vectors are input into the SVM classifier for extracting the initial saliency map. As shown in Fig. 7, most of the background and salient areas have been separated, but there are holes and a few connections in the saliency map. In order to eliminate these error blocks, the initial saliency map is repaired by applying morphological processing and algebraic operations. Fig. 8 presents the complete saliency map extracted.

4.2. Quality assessment of saliency map

In order to evaluate the extraction effects, the Accuracy, Precision, Recall, and F-Measure are calculated according to Eq. (7) for evaluating the quality of saliency map extraction (Ejaz et.al., 2013). The image segmented manually by Photoshop software is regarded as a standard. Fig. 9 depicts the quality assessment result of samples. Among the considered evaluation indices, Accuracy reflects the correctness ratio of

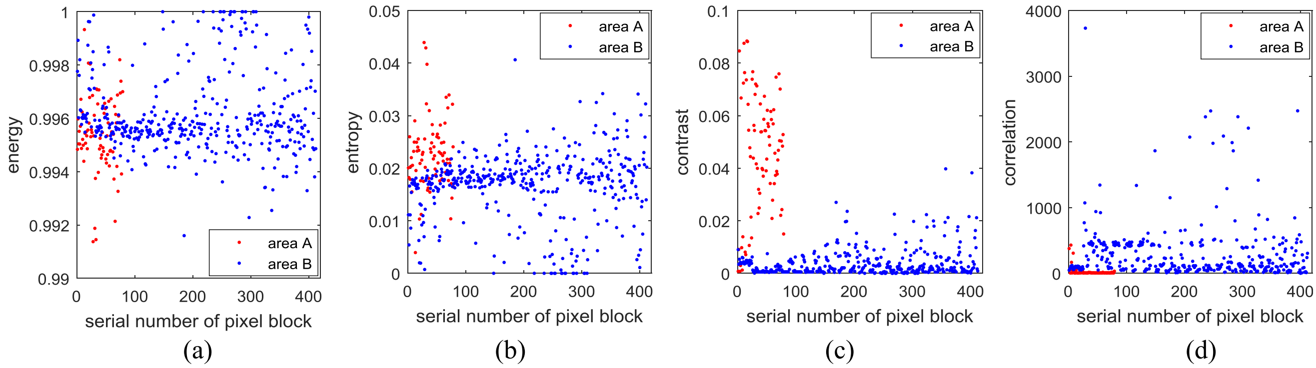

Fig. 6. Texture feature data distribution diagram where (a) is the data distribution of energy (b) is the entropy distribution (c) is the contrast distribution (d) is the correlation distribution.

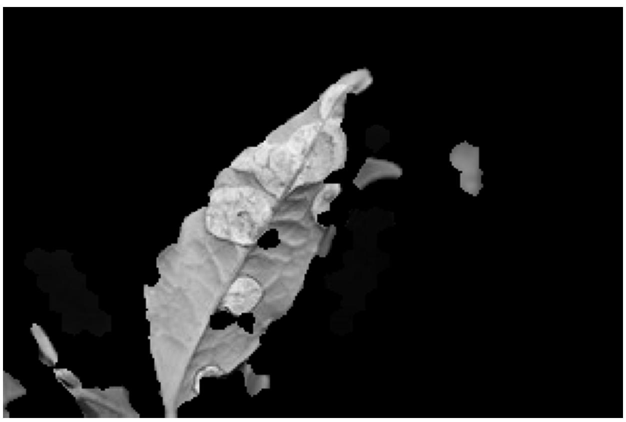

Fig. 7. classification map of SVM.

Fig. 8. accurate saliency map.

extracted salient region and background area pixels to original image, Precision reflects the correctness ratio of extracted salient region pixels to extracted salient region, Recall reflects the correctness ratio of extracted salient region pixels to original salient region, and F-Measure is the harmonic average of Precision and Recall.

$$\begin{cases} Accuracy = \frac{TS + TB}{TS + FS + FB + TB} \\ Precision = \frac{TS}{TS + FS} \\ \mathrm{Re}\,call = \frac{TS}{TS + FB} \\ F = \frac{2Recall * Precission}{Recall + Precission} \end{cases} \tag{7}$$

where TS is the correctness number of salient region pixels, TB is the correct number of background pixels, FS is the number of background pixels extracted to salient region, FB is the number of salient region pixels extracted to background.

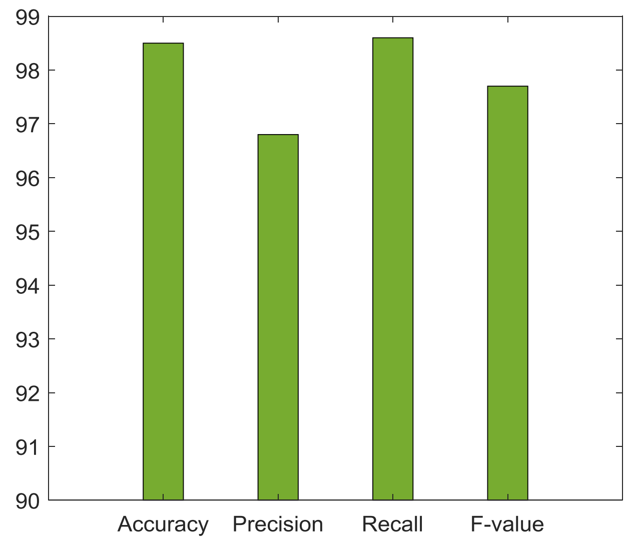

Fig. 9. Quality assessment results of samples.

4.3. Methods comparison

In order to verify the performance, we have compared the proposed algorithm with the typical saliency map extraction algorithms related to SLIC, such as, Saliency Detection with Multi-Scale Super-pixels (SPL_MS) (Jiang et al., 2013), Saliency Detection via Absorbing Markov Chain (Absorb_MC) (Tong et al., 2014), and Saliency Detection via Graph-Based Manifold Ranking (cvprode) (Yang et al., 2013). All the aforementioned algorithms are used to extract the saliency map of experimental set samples images with five common tea plant diseases.

Results of the example image are shown in Fig. 10. The calculation results of the quality evaluation index are shown in Table 2. Results indicate that the proposed saliency map extraction algorithm effectively completely distinguishes diseased leaves information from complex back- grounds. The proposed algorithm performed better than other three algorithms obviously. As illustrated in Fig. 9, Accuracy, Precision, Recall and F-value are 98.5%, 96.8%, 98.6% and 97.7%, respectively. Compared with other three algorithms, Accuracy of SLIC_SVM have increased by an average of 11.46%, Precision have increased by an average of 14.5%, Recall have increased by an average of 11.94%, and F-value have increased an average of 15.02%. The average standard deviation of SLIV_SVM algorithm is 0.021, improved approximately an order of magnitude than others algorithm.

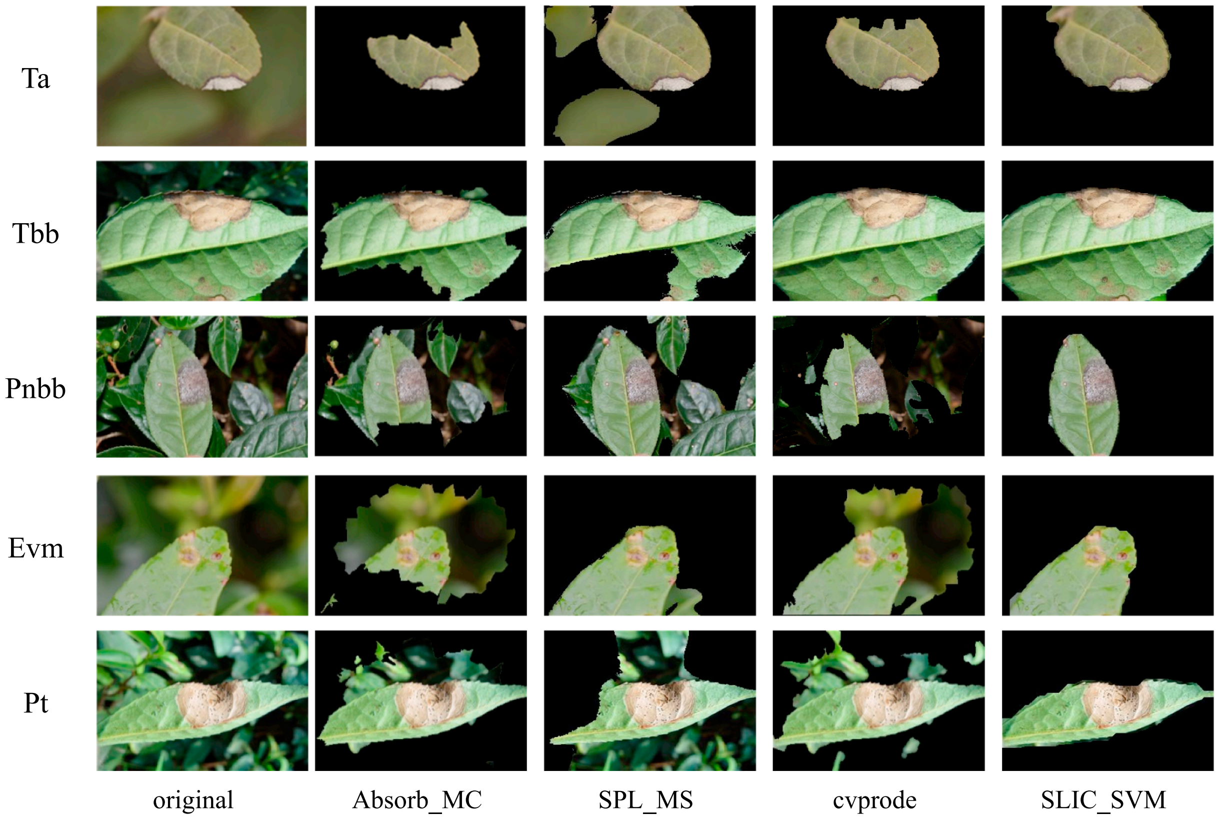

Fig. 10. Extracted images by different algorithms.

The above results demonstrate that SLIC_SVM algorithm obviously performed better than other schemes. This is because the proposed algorithm takes into consideration the characteristics of tea plants leaves disease images. The analysis shows that SLIC_SVM is very effective in extracting saliency map from complex background.

4.4. Parameters discussion

The parameters that gets involved in the proposed algorithm are SLIC initial clustering center k, SVM classifier penalty factor C, and polynomial kernel function formulas. It should be noted that kernel function uses three parameters of γ(g), r and d, the influence of parameters on the algorithm is investigated as follows.

4.4.1. Chosen of initial clustering center k

For the purpose of investigating the influence of the value of k, we set k as 100, 200, 300, 400, 500, 600, and 700, respectively. Then SVM was used to implement initial classification. The classification results are shown in Fig. 11, and the quality index calculations results are shown in Table 3.

In Table 3, the Accuracy, Precision, and F values are higher than other values when k is 500, but Recall is lower than other k values. Considering the actual meaning of Recall and Fig. 11, it can be seen that when $k = 500$, there are holes in salient region but the background and salient region have high resolution, the remaining k value has a small Precision value, moreover, the background and salient area have more adhesions in Fig. 11, which is not conducive to further extraction. Therefore, in order to ensure the segmentation accuracy and maximize the separation of the background and the salience region, it is reasonable if regarding $k = 500$ as the optimal initial clustering center number of the SLIC algorithm.

The number of initial cluster centers k influences and determines the number of final clustering super-pixel blocks, which directly affects the division of area blocks, and then affects the extraction of salient region contours. Therefore, when selecting k-values for SLIC algorithm to process images in blocks, it is the optimal option to follow the large evaluation index. At the same time, it is necessary to consider whether the connection between the salient region and the background or not is beneficial to the extraction of the final saliency map.

4.4.2. Model parameters definition of SVM

The SVM model mainly involves the determination of the SVM

Table 2
Extract quality comparison of different algorithm.

Algorithm	Accuracy	Precission	Recall	F-value
Absorb_MC	0.791 ± 0.162	0.756 ± 0.261	0.739 ± 0.240	0.717 ± 0.226
SPL_MS	0.853 ± 0.193	0.796 ± 0.223	0.875 ± 0.250	0.817 ± 0.221
cvprode	0.887 ± 0.138	0.812 ± 0.218	0.961 ± 0.084	0.86 ± 0.159
SLIC_SVM	0.985 ± 0.014	0.968 ± 0.028	0.986 ± 0.021	0.977 ± 0.022

Note: The data in table was the mean and standard deviation of evaluation indexes.

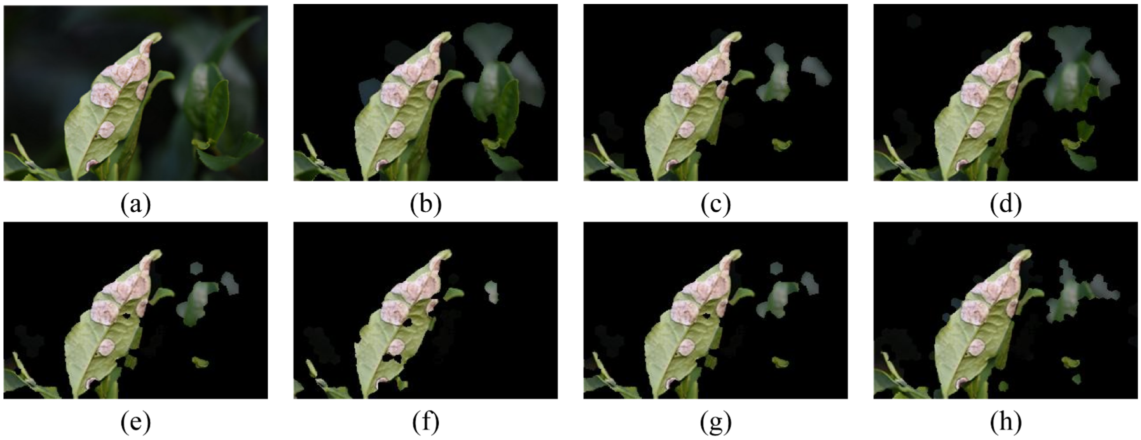

Fig. 11. The classification results of different k value where (a) original image (b) k = 100 (c) k = 200 (d) k = 300 (e) k = 400 (f) k = 500 (g) k = 600 (h) k = 700.

Table 3
The evaluation index under different k.

k	Accuracy	Precision	Recall	F1_measure
100	0.772	0.381	0.989	0.549
200	0.898	0.582	0.978	0.729
300	0.742	0.352	0.997	0.52
400	0.917	0.632	0.976	0.767
500	0.964	0.828	0.935	0.878
600	0.922	0.649	0.976	0.779
700	0.883	0.547	0.993	0.705

classifier type and the choice of SVM training parameters. SVM types include C support vector classifier C-SVC, nu support vector classifier nu-SVC, single class support vector machine one class SVM, ε support vector regression e-SVR and n support vector regression nu-SVR. Since C-SVC is widely used in image segmentation, C-SVC classifier is chosen for extracting tea plant disease saliency map. When polynomial kernel functions is chosen, there are four parameters will influence the result, including the penalty factor C and γ(g), r, d in kernel function calculated by formula (5). Through extensive experiments, it is shown that the classification accuracy changes with parameters, the classification accuracy is quite different, as illustrated in Table 4.

As for the selection of SVM parameters, one always hopes to obtain the best parameters. Here, we discuss and explore the optimal parameters C and γ(g) that have the greatest impact on SVM. The idea of CV is taken into consideration to obtain the optimal parameters in a certain sense. The results are shown in Fig. 12, where (a) is the parameter selection contour map, (b) is the parameter selection 3D view. From the Fig. 12(a), the maximum accuracy displayed is 90.7%. From the Fig. 12(b), the accuracy is 91.0569%, corresponding to the value near the 90.7 in (a). Meanwhile, the optimal parameters value was C = 1, γ (g) = 0.08838. Compared with the results in Table 4, it is observed that the optimal value obtained by CV is the same as the value of C in Table 4, γ is suffering great difference but has a small difference in accuracy.

From the Table 4, the accuracy of training sample classification has no relationship with the size of the parameters, and there is no regularity at all. But the running time increases as the parameters increases. Therefore, the smaller parameter could be selected when the classification accuracy is the same. Compared to the manual method, the CV method saves a lot of time, and it can be one of the alternatives for determining the optimal parameters.

5. Conclusion

This paper proposes a method for extracting saliency map of tea plant leaf disease images. The characteristics of the algorithm are described as follows.

(1) The SLIC is used as a preprocessing step of the segmentation algorithm to extract tea plant leaf disease saliency maps, which can better separate the boundary of the region. Furthermore, it has capability of solving the problems of incomplete extraction of boundary extraction or loss of disease information in the extraction of existing disease saliency maps.
(2) The automatic selection of SVM algorithm training set could be realized as two steps. Firstly, fuzzy contour extraction is conducted a SLIC-based saliency region, and then the construction of training set data can be implemented by extracting the super-pixel block features in multiple directions with the help of GLCM algorithm.
(3) Through experiments on sample set image and comparison with other SLIC based algorithm, the method of combining SLIC with SVM is novel and effective, and it is one reliable and effective method for extracting image saliency maps. It is well applied in tea leaf disease saliency extraction.

The combination method of SLIC and SVM can accurately extract the saliency map of tea plant leaf disease, which is conducive to the accurate identification of tea leaf disease and lays effective foundation for rapid and accurate detection and disease prevention.

Table 4
The classification accuracy with different parameters.

	C	Accuracy		d	Accuracy		γ	Accuracy		r	Accuracy
d = 1	1	91.86	C = 1	1	91.86	C = 1	100	90.24	C = 1	0	89.87
γ = 300	10	89.43	γ = 300	2	87.8	d = 1	200	81.3	d = 1	4	84.95
r = 8	100	91.05	r = 8	3	32.52	r = 8	300	91.86	γ = 300	8	91.86
	200	91.86		4	32.52		400	81.3		10	89.02
	300	90.24		5	8.94		500	87.39		20	84.95

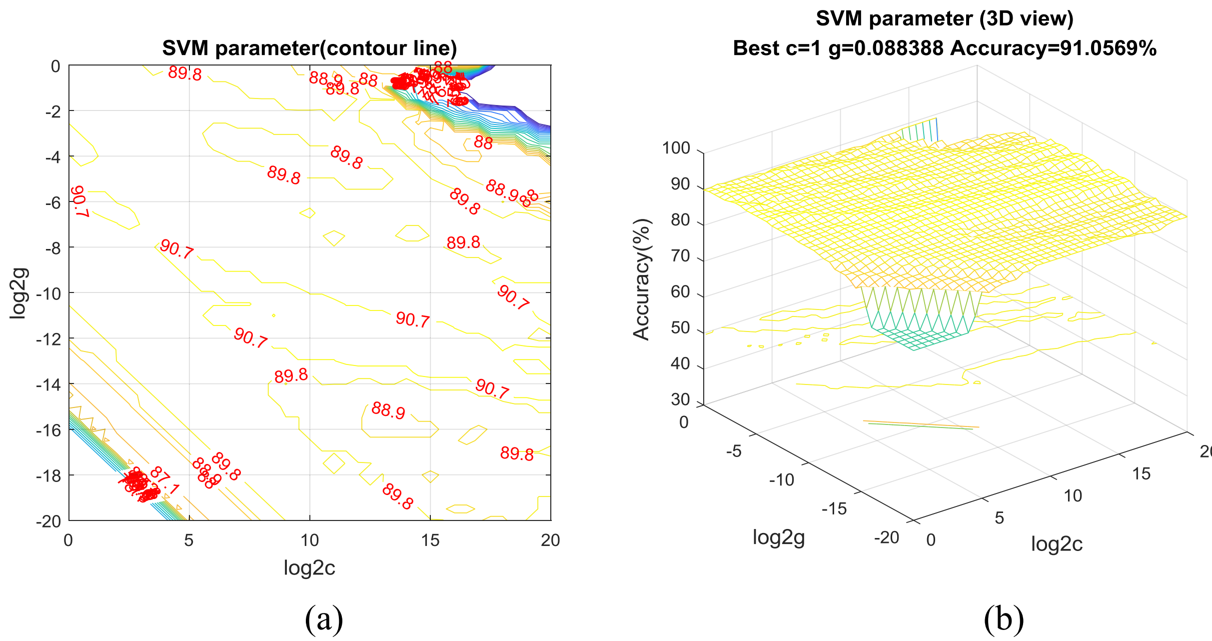

Fig. 12. C and γ(g) optimization results by algorithm.

Acknowledgments

The authors would like to thank financial support provided by Open fund for Key Laboratory of Technology Integration and Application in Agricultural Internet of Things, Ministry of Agriculture, P. R. China (2016KL01) & the Introduction and Innovation of International Advanced Agricultural Science and Technology Program (Nos. 2015-Z44 & 2016-X34) & 2018 graduate innovation fund of Anhui Agricultural University (2018yjs-63).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2018.12.042.

References

Ghosh, P., Mishra, A.K., Bhattacharya, A.K., et al., 2013. Effect of tea in cerebrovascular disease and diabetes: a study. Int. J. Tea Sci. 43–47.
Cabrera, C., Artacho, R., Giménez, R., 2006. Beneficial effects of green tea—a review. J. Am. Coll. Nutr. 25 (2), 79–99.
Majumder, A.B., Bera, B., Rajan, A., 2011. Tea statistics: global scenario. Int. J. Tea Sci. 8 (1), 121–124.
Gao, X.H., 1996. On the integrated control measures of the tea plant disease. Progress of research on plant protection in China. In: Proceedings of the third national conference of integrated pest management, pp. 337–340.
Chen, T.M., Chen, S.F., 1982. Diseases of tea and their control in the People's Republic of China. J. Emerg. Med. 66 (1), 961–965.
Ma, J., Du, K., Zhang, L., et al., 2017. A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing. Comput. Electron. Agric. 142 (142), 110–117.
Cheng, M., Mitra, N.J., Huang, X., et al., 2015. Global contrast based salient region detection. IEEE Trans. Pattern Anal. Mach. Intell. 37 (3), 408–416.
Ma, Y., Huang, M., Yang, B., et al., 2014. Automatic threshold method and optimal wavelength selection for insect-damaged vegetable soybean detection using hyperspectral images. Comput. Electron. Agric. 106 (106), 102–110.
Otsu, N., 2007. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybernetics 9 (1), 62–66.
Mathew, J., Shinde, R.C., Patil, P.C.Y., 2015. Segmentation technique for soybean leaves disease detection. Int. J. Adv. Res. 3, 522–528.
Ziou, D., Tabbone, S., Ziou, D., 1998. Edge detection techniques – an overview. Int. J. Patt. Recogn. Image Anal. 8 (4), 537–559.
Yang, L., Dickinson, J., Wu, Q.M.J., et al., 2008. A fruit recognition method for automatic harvesting. In: IEEE, International Conference on Mechatronics and Machine Vision in Practice, pp. 152–157.
Zhao, D.A., Lv, J., Ji, W., et al., 2011. Design and control of an apple harvesting robot. Biosyst. Eng. 110 (2), 112–122.
Bischof, 2002. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16 (6), 641–647.
Ren, X., Malik, J., 2003. Learning a classification model for segmentation. IEEE Int. Conf. Comput. Vision Proc. 1, 10–17.
Fu, K., Gong, C., Gu, I.Y., et al., 2015. Normalized cut-based saliency detection by adaptive multi-level region merging. IEEE Trans. Image Process. 24 (12), 5671–5683.
Jiang, B., Zhang, L., Lu, H., et al., 2013. Saliency detection via absorbing markov chain. IEEE Int. Conf. Comput. Vision 1665–1672.
Lu, H., Cao, Z., Xiao, Y., et al., 2016. Region-based colour modelling for joint crop and maize tassel segmentation. Biosyst. Eng. 147, 139–150.
Vapnik, Vladimir N, 2002. The nature of statistical learning theory. IEEE Trans. Neural Network 8 (6), 1564.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.
Bai, X., Wang, W., 2014. Saliency-SVM: an automatic approach for image segmentation. Neurocomputing 136 (8), 243–255.
Zhao, Q., Hu, Y., Cao, J., 2010. Automatic image segmentation based on saliency maps and Fuzzy SVM. Iet Int. Commun. Conf. Wireless Mobile Comput. 121–124.
Maghsoudi, O.H., 2017. Superpixels based segmentation and SVM based classification method to distinguish five diseases from normal regions in wireless capsule endoscopy. Comput. Vision Patt. Recogn. 17, 1–13.
Wu, Z., Hu, Z., Fan, Q., 2012. Superpixel-based unsupervised change detection using multi-dimensional change vector analysis and Svm-based classification. Rem. Sens. Spatial Inform. Sci. I-7, 257–262.
Achanta, R., Shaji, A., Smith, K., et al., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34 (11), 2274–2282.
Achanta, R., Shaji, A., Smith, K., et al., 2010. SLIC superpixels. Epfl 1–15.
Wang, J., Lu, H., Li, X., et al., 2015. Saliency detection via background and foreground seed selection. In: Neurocomputing 152(C). pp. 359–368.
Linker, R., Cohen, O., Naor, A., 2012. Determination of the number of green apples in RGB images recorded in orchards. Comput. Electron. Agric. 81 (1), 45–57.
David, S., Murat, K., Chieri, K., et al., 2010. Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. Comput. Electron. Agric. 74 (2), 238–243.
Pourreza, A., Pourreza, H., Abbaspourfard, M.H., et al., 2012. Identification of nine Iranian wheat seed varieties by textural analysis with image processing. Comput. Electron. Agric. 83 (1), 102–108.
Ronge, R.V., Sardeshmukh, M.M., 2014. Indian wheat seed classification based On texture analysis using Ann. Int. J. Elect. Electron. Data Commun. 2 (7), 42–45.
Zapotoczny, P., Zielinska, M., Nita, Z., 2008. Application of image analysis for the varietal classification of barley: morphological features. J. Cereal Sci. 48 (1), 104–110.
Ejaz, N., Mehmood, I., Baik, S.W., 2013. Efficient visual attention based framework for extracting key frames from videos. Signal Process. Image Commun. 28 (1), 34–44.
Tong, N., Lu, H., Zhang, L., et al., 2014. Saliency detection with multi-scale superpixels. IEEE Signal Process. Lett. 21 (9), 1035–1039.
Yang, C., Zhang, L., Lu, H., et al., 2013. Saliency detection via graph-based manifold ranking. IEEE, Comput. Vision Patt. Recogn. 3166–3173.