



Plant Classification in Images of Natural Scenes Using Segmentations Fusion

N. Nikbakhsh^a, Y. Baleghi Damavandi^{*a}, H. Agahi^b

^a Department of Electrical & Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

^b Department of Mathematics, Faculty of Basic Science, Babol Noshirvani University of Technology, Babol, Iran

PAPER INFO

Paper history:

Received 25 April 2020

Received in revised form 30 June 2020

Accepted 22 August 2020

Keywords:

G-calculus

Image Segmentation Fusion

Mutual Information

Plant Classification

Tsallis Entropy

ABSTRACT

This paper presents a novel approach to automatic classifying and identifying of tree leaves using image segmentation fusion. With the development of mobile devices and remote access, automatic plant identification in images taken in natural scenes has received much attention. Image segmentation plays a key role in most plant identification methods, especially in complex background images. Where there are no presumptions about leaf and background, segmentation of leaves in images with complex background is very difficult. In addition, each image has special conditions, so parameters of the algorithm must be set for each image. In this paper, image segmentation fusion is used to overcome this problem. A fast method based on maximum mutual information is used to fuse the results of leaf segmentation algorithms with different parameters. Applying Tsallis entropy and *g*-calculus, generalized mutual information equations are derived and used to obtain the best consensus segmentation. To evaluate the proposed methods, a dataset with tree leaf images in natural scenes and complex backgrounds is used. These images were taken from Pl@ntLeaves dataset and modified so that they do not have a presumption. The experimental results show the use of Tsallis entropy and *g*-calculus in image segmentation fusion, improves plant species identification.

doi: 10.5829/ije.2020.33.09c.07

1. INTRODUCTION

Classification and identification of plant species are essential for plant extinction prevention, the use and development of plant resources, herbal medicines and separation of products from undesired plants in automation systems. However, identification of plants requires expert knowledge, and is a challenging and tedious task. Automatic plant identification through digital images helps the classification and identification of plant species according to time and human energy consumption. Unfortunately, despite the widespread biodiversity of plants, current processes of identification and classification are both error prone and slow [1]. In recent years, with the development of mobile devices and remote access, automatic plant identification in images taken in natural scenes has received much attention. Compared to other plant organs, leaves are the most widely used part for automatic plant identification using computer vision techniques because they are accessible throughout the year and easy to analyze. Automatic tree

leaf classification is very accurate in images with white and uniform light background, but very difficult in natural scene images that are acquired in uncontrolled conditions.

To increase the accuracy of the classification, the classifier can use several features such as leaf color, overall shape, edge and vein structure. In most approaches, the accuracy of identification depends on the accuracy of the segmentation, which is the most difficult task in a complex natural scene [2, 3]. Because there are no specific conditions for image acquisition in natural scenes, an unsupervised segmentation algorithm with constant parameters leads to poor results for some images. In general, there is no best algorithm for segmentation of various images. In this paper, fusing the results of different algorithms or different parameters is used to overcome this problem. The clustering ensemble combines the results of different clustering algorithms to obtain a more accurate final clustering which is called consensus clustering. Mutual information is one of the most common parameters used in clustering ensemble. It

*Corresponding Author Institutional Email: y.baleghi@nit.ac.ir
(Y. Baleghi Damavandi)

shows the amount of dependency between the two variables. However, if clusters have a large amount of data, fusion of them using criteria such as maximum mutual information requires a fast methods.

Plant leaf identification has been a hot research topic in recent years [4, 5]. Although a number of methods have been designed to identify plant species, new methods are still needed to improve the identification accuracy and speed, especially in complex background images. Most of tree leaves datasets have images with plain and uniform background. In addition, most of image datasets for tree leaves with natural scene use defaults such as color, position or orientation of leaf in the image. PI@antLeaves dataset [2, 6] contains tree leaf images in the real world. These images are divided into three categories that were taken under different conditions. The scan category includes images with a white background, and the scan-like category includes images with a uniform background. There is no occlusion or overlap in both of them. However, the photograph category includes photos directly taken from the trees in nature. These photographs have a non-uniform background with optical distortions, shadows, color and luminance variations, and other problems such as overlapping. The most important advantage of this dataset is that it contains images that are very close to what a smartphone user may take from nature.

Most plant identification methods require images with simple background, or assumptions about leaf color or shape, and many of them need user interaction. Very few works handle the automatic segmentation and identification of plants in images with complex background [3]. A system called LeafSnap [7] was introduced to identify plant species in the mobile system that uses leaf shape in plain background images, but does not work well in complex and natural background images. In [8], the authors employed the semi-supervised Fuzzy C-Mean (FCM) classifier on different features to classify a number of images in the PI@ntLeaves dataset (scan, scan-like, and photo) In [4] the authors used the Support Vectors Machine (SVM) classifier with the color, shape and texture characteristics to classify tree leaf images. They reached an accuracy of about 61% for a number of scan and scan-like images on the PI@antLeaves dataset, but this accuracy was only 8.5% for photo images. In [9] using Tsallis entropy and truth table, a new fast method was proposed for segmentation of tree leaves in images with complex background. In [10] the authors obtained 63.4 % of accuracy for 70 classes of species with simple background images (the scan and scan-like images). In [11], an approach was proposed to identify tree leaves using hand-crafted features, however, this method requires the leaf image without any occlusion and with a uniform background.

In recent years, another approach called the convolutional neural networks (CNN) has been used to

identify plants, in which the leaf features are directly represented [6, 12]. In [13], the authors achieved a classification accuracy of 71% by testing a variety of CNNs on a combination of scan, scan-like and photo images. However, there are still many problems. In addition to hardware restrictions, this method requires a large number of well-categorized data which makes it impossible to use it in many cases, especially in datasets with a limited number of samples. These methods are not sensitive to image details. Also they are not accurate in overlapping images, and the relationship between pixels is not fully considered [14].

The purpose of the present work is to introduce a novel approach to automatic classification and identification of tree leaves using image segmentation fusion. The proposed method based on the two different methods for image segmentation fusion, and the dataset are described in Section 2. The experimental results are reported in Section 3, and the conclusion is given in Section 4.

2. MATERIALS AND METHODS

Automatic identification of plants in images taken from natural scenes is very difficult when prior information about the leaf or background is not available. Leaf image segmentation is the most important step in plant identification in leaf images with complex background. However, each segmentation algorithm requires special settings for each image. In this paper, image segmentation fusion is used to overcome this problem. In Figure 1, the flow diagram of the proposed method is shown.

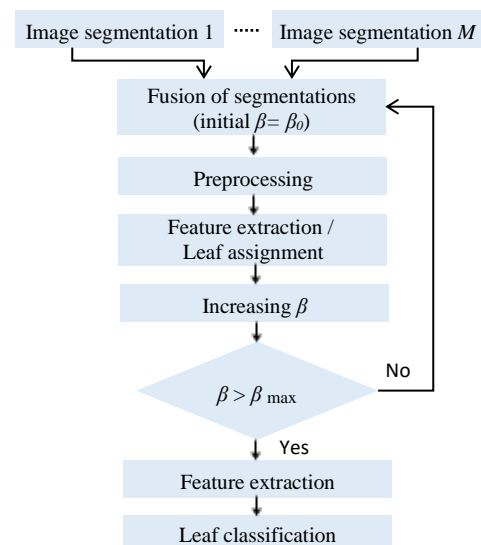


Figure 1. Flow diagram of the proposed method

Image segmentation fusion is usually a two-step algorithm. In the first step, the base clusterings are generated by different clustering algorithms or an algorithm with different parameters. Then, using a consensus function, the base clusterings are combined to find a new clustering so that it is most similar to the base clusterings. In this paper, fusion of unsupervised segmentations is done using maximum mutual information and a g -function [15] or Tsallis entropy.

2. 1. Image Segmentation Fusion

Suppose X is an N -pixel image, and there are M different unsupervised segmentation algorithms. Each algorithm $S^j = \{S_i^j, [i]_1^N\}$, divides the image into background and object segments (labeled 0 and 1). Unlike supervised classification, the labels produced by these algorithms do not match. Table 1 shows an example of the labels produced by the different segmentation algorithms. A set of all segmentation results for the pixel i , provides a feature vector for representation of this pixel, $x_i = \{S_i^j, [j]_1^M\}$. Image segmentation fusion, combines the base segmentation results into S^* , and gets the best segmentation (S_{best}^*) which divides the image pixels into two parts so that this segmentation is most compatible with all the base segmentation results.

If the labels of the two segmentation results are in full correspondence, the mutual information between them is maximum. The mutual information between X, Y is defined as:

$$I(X, Y) = H_X + H_Y - H_{X,Y} \tag{1}$$

where H_X and H_Y are the classical entropy, and $H_{X,Y}$ is the joint entropy. The classical Shannon entropy is defined as $H = -\sum_x P(x) \ln P(x)$, and is a measure of uncertainty of the random variable X with probability distribution $P(x)$. Mutual information between S^* and all of the base segmentation results (S) is obtained as follows:

$$I(S^*, S) = \sum_{j=1}^M I(S^*, S^j) \tag{2}$$

The results of the different base segmentation algorithms for each pixel, are considered as a new feature vector for it. In an image with N pixels and M segmentation algorithms as in Table 1, there are only 2^M different

TABLE 1. The results of different segmentation algorithms

| X | S^1 | S^2 | S^3 | ... | S^M | S^* |
|-------|-------|-------|-------|-----|-------|---------------------------------------|
| x_1 | 1 | 0 | 0 | ... | 1 | $S_1^* S_N^* S_N^* S_1^* S_1^* S_1^*$ |
| x_2 | 0 | 1 | 1 | ... | 0 | $S_2^* S_2^* S_2^*$ |
| ... | ... | ... | ... | ... | ... | ... |
| x_N | 1 | 0 | 0 | ... | 1 | $S_N^* S_N^* S_N^* S_N^*$ |

feature vectors. A state table can be created which includes all of the different feature vectors or 2^M state vectors [9, 16]. This table has 2^M rows or state vectors (SV). The number of repetitions for each of the state vectors in all pixels from the image is computed as f_i . The first row in this state table is selected as b_1 , and the vector with maximum distance with b_1 is selected as b_2 .

Then all state vectors are grouped into two categories based on similarity to b_1 or b_2 . A new segmentation (S^*) is created by assigning the value of 0 or 1 to each state vector in this truth table. The value of each state vector is denoted by u_i^* , and is equal to 0 for vector b_1 and similar vectors, and 1 for vector b_2 and similar vectors. After generating this new segmentation (S^*), the mutual information between the new segmentation and the base segmentations is calculated. In the next step, the second vector in truth table is selected as the base vector b_1 , and the above process is repeated. This procedure is repeated for half of the state vectors, and the segmentation which has the maximum mutual information is selected as the best consensus clustering as Equation (3).

$$S_{best}^* = argmax_{S^*} I_{\beta}(S^*, S) \tag{3}$$

This method greatly reduces the complexity of computations. In an N -pixel image and M segmentation algorithms, the time complexity in the exhausting search is $O(2^{N-1})$, and in the Bayesian clustering ensemble method [17] is $O(2NT)$, where T is the number of iterations until the convergence. However; in the proposed method, the time complexity is $O(2^{M-1})$, and the mutual information is calculated only for $2^M/2$ cases.

In calculation of maximum mutual information, the use of Tsallis entropy or g -calculus can lead to better results because these two methods have an additional parameter. By varying this parameter from the initial value to the maximum, a better final consensus segmentation will be obtained. These two methods are described below.

Method 1: Mutual information based on Tsallis entropy (I_{β}) between S^* and S^j is calculated as follows:

$$I_{\beta}(S^*, S^j) = H_{\beta}(S^j) + H_{\beta}(S^*) - H_{\beta}(S^j, S^*) \tag{4}$$

$H_{\beta}(S^j)$ and $H_{\beta}(S^*)$ represent Tsallis entropies, and are calculated as follows:

$$H_{\beta}(S^j) = (1 - \beta)^{-1} (\sum_{k=0}^1 P_k(S^j)^{\beta} - 1), \tag{5}$$

where $P_0(S^j)$ is obtained by dividing the number of pixels in which $S^j = 0$ by the total number of pixels. $P_0(S^j)$ and $P_1(S^j)$ are obtained as follows:

$$P_k(S^j) = \frac{1}{N} \sum_{i=1}^M f_i \quad \text{if } SV_i^j = k, \tag{6}$$

where SV_i^j refers to the j th bit of SV_i (i th state vector) in the state table, and f_i is the frequency of each state vector in all pixels.

$$H_{\beta}(S^*) = (1 - \beta)^{-1} (\sum_{k=0}^1 P_k(S^*)^{\beta} - 1), \quad (7)$$

$$P_k(S^*) = \frac{1}{N} \sum_{i=1}^M f_i \quad \text{if } u_i^* = k, \quad (8)$$

where u_i^* refers to the value of SV_i (0 or 1).

$H_{\beta}(S^j, S^*)$ represents the joint entropy that is the entropy of a joint probability distribution for a pair of random variables, and is calculated as Equation (9).

$$H_{\beta}(S^j, S^*) = (1 - \beta)^{-1} (\sum_{k=0}^1 \sum_{l=0}^1 P_{kl}^{\beta} - 1), \quad (9)$$

where,

$$P_{kl} = \frac{1}{N} \sum_{i=1}^M f_i \quad \text{if } SV_i^j = k \text{ and } u_i^* = l \quad (10)$$

For example, P_{01} is obtained by dividing the number of pixels in which $S^j = 0$ and $S^* = 1$ by the total number of pixels.

In this case, the classical Shannon entropy for mutual information is obtained by $\beta = 1$.

Method 2: The g -calculus [15] is a development in Mathematics. Assuming that $-\infty \leq a < b \leq +\infty$ and $g: [a, b] \rightarrow [0, +\infty)$ is a continuous monotonic function, then \oplus (pseudo-addition), the generalization of the classical operation, is defined by using the generating function g as follows:

$$x \oplus y = g^{-1}(g(x) + g(y)). \quad (11)$$

The function used by g -calculus is called the g -function. Using a g -function, new mutual information equations are generated which we call them g -mutual information equations. Using pseudo-addition:

$$I_{\oplus}(S^*, S) = \oplus_{j=1}^M I(S^*, S^j) \quad (12)$$

If g is a strictly upward and continuous function such that $g(a) = 0$, then we get the following equation:

$$I(S^*, S) = g^{-1}[\sum_{j=1}^M g(I(S^*, S^j))]. \quad (13)$$

Assuming $g: [-\infty, +\infty] \rightarrow [0, +\infty)$, $g(x) = e^{\beta x}$, $\beta > 0$, then mutual information is obtained as follows:

$$x \oplus y = \frac{1}{\beta} \ln(e^{\beta x} + e^{\beta y}), \quad (14)$$

$$I_{\beta}(S^*, S) = \frac{1}{\beta} \ln \sum_{j=1}^M e^{\beta I(S^*, S^j)} \quad (15)$$

$$I_{\beta}(S^*, S) = \frac{1}{\beta} \ln \sum_{j=1}^M e^{\beta (H_{S^j} + H_{S^*} - H_{S^j, S^*})} \quad (16)$$

where $H_{\beta}(S^j)$, $H_{\beta}(S^*)$ and $H_{\beta}(S^j, S^*)$ are obtained as follows:

$$H_{\beta}(S^*) = \sum_{k=0}^1 P_k(S^*) \ln P_k(S^*), \quad (17)$$

$$H_{\beta}(S^j) = \sum_{k=0}^1 P_k(S^j) \ln P_k(S^j), \quad (18)$$

$$H_{\beta}(S^j, S^*) = \sum_{k=0}^1 \sum_{l=0}^1 P_{kl} \ln P_{kl}, \quad (19)$$

In these equations, $P_k(S^*)$, $P_k(S^j)$ and P_{kl} are obtained from Equations (6), (8) and (10).

2. 2. Preprocessing

The segmented image may contain undesirable noise, false classified pixels and connected regions. After the fusion of segmentations, morphological operations such as hole filling, opening and closing are used to improve the segmented image [9]. Morphological closing applies dilation process to an image, followed by an erosion process, while morphological opening is a reverse process. After these operations, boundary correction and separation of the overlapping parts that have different colors with the main part or the shadow are performed. By morphological dilation (or erosion) and subtraction of the source image, several pixels are selected around the borders. The pixels in this area, which are different in color from the object, are considered as background. The color dissimilarities of pixel $x(i, j)$ with the object and background (d_1, d_2), are obtained using Hue, Saturation and Value components in the *HSV* color space as follows:

$$d_1(x(i, j)) = \sum_{k=1}^3 \frac{(x(i, j, k) - \mu_1)^2}{\sigma_1^2} \quad (20)$$

$$d_2(x(i, j)) = \sum_{k=1}^3 \frac{(x(i, j, k) - \mu_2)^2}{\sigma_2^2} \quad (21)$$

where, the mean and variance of the object and background are assumed as $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$, respectively. If $d_1 > d_2$ then the pixel $x(i, j)$ is considered as the background. The example of these operations are shown in Figure 2.

2. 3. Shape Features Extraction

The accuracy of classifiers largely depends on the accuracy of the segmentation and feature extraction process. A feature is an object characteristic that is different from other objects. Leaf color may vary in different seasons and geographical locations. In addition, different plant species may have leaves of the same color. Shape is an important feature of image description. The accuracy of feature extraction from the shape depends greatly on the quality of the image segmentation. After the object is segmented from the image by fusion of the results of

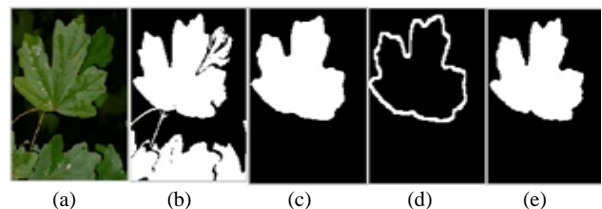


Figure 2. (a) Original image; (b) Result of fusion of segmentation algorithms; (c) After opening, selecting the largest segment, and closing; (d) The pixels around the borders; (e) After boundary correction

segmentation algorithms, it is necessary to calculate the similarity between the segmented shapes and the predefined ones. To do this, a feature vector including 6 digital morphological features [4, 9], 6 elliptical Fourier descriptors (EFD) [6] and the first Hu invariant moment [4, 6] is extracted from shapes.

Let the longest distance between two points on the border of the leaf be indicated by L , and the length of the longest line perpendicular to L by W (width). Also, A is the leaf area which indicates the number of pixels in the leaf and P is the leaf perimeter which counts the number of pixels at the leaf border. Then 13 features are obtained as follows.

Rectangularity feature, which shows the similarity of a leaf and its rectangle is calculated by $L.W/A$. Form factor, which shows the difference between a leaf and a circle, is defined by $4\pi A/P^2$. Perimeter to length ratio, is calculated by P/L . Perimeter ratio of length and width, is calculated by $P/(L+W)$. Aspect ratio shows the ratio of leaf length and leaf width (L/W). Vein feature, which defines the skeletal structure of the leaf, is calculated by dividing the total vein pixels by the total number of pixels in the leaf.

Elliptical Fourier Descriptors are used as a set of elliptical harmonics to approximate a closed edge. We select the first 6 harmonics as 6 features.

The first Hu invariant moment of an intensity function $f(x,y)$ is defined as:

$$M_1 = \left(\frac{\mu_{20}}{\mu_{00}}\right) + \left(\frac{\mu_{02}}{\mu_{00}}\right) \tag{22}$$

where μ_{00} , μ_{02} and μ_{20} are calculated as:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{23}$$

$$\bar{x} = \left(\frac{m_{10}}{m_{00}}\right), \quad \bar{y} = \left(\frac{m_{01}}{m_{00}}\right) \tag{24}$$

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y). \tag{25}$$

Thus, a feature vector containing 13 components is created.

In each value of β , after determining the S_{best}^* , which has the maximum mutual information, the segments of the leaf and background classes are examined. The features of these two segments are extracted and their similarities are calculated with the features of predefined shapes. The leaf class is determined with the maximum similarity by calculating the minimum distance with the stored features of the predefined shapes. The similarity is measured using the Euclidean distance between the shape features of these two classes (segments) and the features of the predefined shapes. Figure 3 shows the predefined leaf shapes for the 30 plant species used in this paper.

The new mutual information equations, made by Tsallis entropy or g -function, have an additional parameter, so that by changing this parameter the best result can be obtained for the fusion of segmentations.



Figure 3. Predefined leaf shapes for the 30 plant species

First, with initial parameter of β , the resulting segments in the leaf and background classes are examined, and using the Euclidean distance, their similarity with predefined shapes are calculated. In the next step, β increases, and this process is repeated until β reaches the maximum value. In each step, the Euclidean distance between the features of each resulting segments and features of the predefined shapes is calculated. Finally, a segmentation with one β is selected that makes the least Euclidean distance with one of the predefined shapes. In the fusion using Tsallis entropy, β can vary from 0.1 to 1, and if $\beta = 1$, then the fusion with the Shannon entropy method is obtained.

2. 5. Dataset

Most datasets used to evaluate plant classification algorithms have defaults that cannot be applied to real cases. For this reason, we generated a dataset of leaf images with natural scenes and without presumptions, to evaluate the segmentation and identification algorithms. This dataset consists of 200 tree leaf images with natural scenes extracted from PI@nTLeaves dataset with segmentation ground truth that we have extracted manually. It can be downloaded from “ftp://doc.nit.ac.ir/cee/electronic/baleghi.yaser/Plants_Dataset/“. Some of these photograph images have been modified so that the leaf is not always vertical or in the middle of the image. These images are classified into 30 species of plants and can be used to evaluate unsupervised algorithms for leaf segmentation and identification. Figure 3 shows an example of each class, and Figure 4 shows the sample images of this dataset. The images of this dataset are taken in natural conditions and have problems such as different shades and lighting, overlapping, different colors and defects.

2. 6. Performance Metric

Average top- n accuracies are usually used with $n=1, 5$ or 10 . Assuming that m is the number of evaluation samples, and y_i is the



Figure 4. Sample images from dataset

correct species class for input sample x_i , $i=1, \dots, m$, then:

$$\text{Top-}n \text{ accuracy} = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n I(f_j(x_i) = y_i) \quad (26)$$

where $j=1, \dots, n$, is the highest ranked species predictions according to their probability values, and $I(\cdot)$ is the indicator function which returns 1 for the true expressions and 0 for otherwise.

The other metric is $F1$ -measure that is defined as:

$$F1\text{-measure} = \frac{1}{C} \sum_{i=1}^C \frac{2 * TP_i}{2 * TP_i + FP_i + FN_i} \quad (27)$$

where, TP_i indicates the number of correct predictions on class i , FP_i indicates the samples that do not belong to class i but are predicted in this class, and FN_i indicates the samples that belong to class i but are not predicted in this class.

3. RESULTS AND DISCUSSION

3.1. Evaluation and Results

Since an algorithm with a fixed parameter is not suitable for segmentation of all images, the fusion of four different algorithms is used in this work. To do this, the FCM [18] and k -means algorithms with different parameters are used as individual image segmentation methods. In the k -means clustering algorithm, each pixel is assigned to the nearest cluster, but in the FCM algorithm each pixel is assigned to clusters with different degrees of membership. FCM uses a procedure to minimize the weighted summation of distances from the pixels to the M cluster centers as follows:

$$J = \sum_{i=1}^N \sum_{l=1}^M \mu_{li}^k \|X_i - C_l\|^2 \quad (28)$$

where X is an N -pixel image, k is a coefficient greater than 1, and $0 \leq \mu_{li} \leq 1$ is the degree of membership of image pixel X_i to the cluster l with center C_l .

Since the $YCbCr$ and Lab color spaces have separate luminance and chrominance components, the RGB color leaf image is converted to $YCbCr$ and Lab color spaces. Segmentation Algorithms 1 and 2 divide the image into two parts using the fuzzy c -mean clustering algorithm and the a , b components in the Lab color space. Algorithm 1, uses the membership degree of less than 0.1, and Algorithm 2 uses the membership degree of less than 0.9. Algorithm 3 is k -means clustering using L , a , b components in the Lab color space, and Algorithm 4 is k -means clustering using Y , Cb , Cr components in the $YCbCr$ color space.

In Figure 5, the results of four segmentation algorithms for three sample images are shown in (b) to (e). In (f) and (g), the fusion results of these segmentations are shown using the classic Shannon entropy and the proposed method with Tsallis entropy. The results of the proposed method are obtained with $\beta=0.5$ for images 1-2, and $\beta=0.75$ for image 3 while the

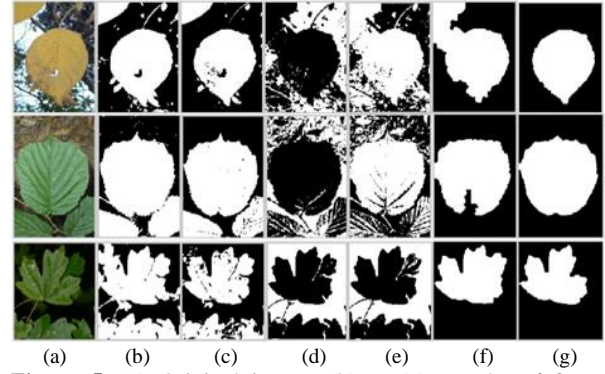


Figure 5. (a) Original image; (b) to (e) Results of four segmentation algorithms; (f) Fusion by $\beta=1$ (Shannon); (g) Fusion by proposed method (Tsallis entropy)

classical Shannon entropy is obtained with $\beta=1$. As shown in this figure, changing the parameter β , can lead to a good improvement in image fusion.

For the computation time comparison, these three images were tested in MATLAB on a PC with Intel core (i3) CPU, 3.7GHz processor and 8GB RAM. The average time required for image segmentation fusion by BISO [17] is approximately 94 seconds but with the proposed method is about 0.7 seconds. The computation time for fusion in the proposed method is greatly improved.

3.2. Comparison of Performance on the Whole Dataset

In Table 2, the average performances of classification algorithms are compared on all images of tree leaves from the dataset. To get these results, the average scores over 200 tree leaf images is calculated. As shown in this table, the proposed methods are better than all individual algorithms in terms of average classification accuracies (Top-1, Top-2, Top-5, and $F1$ -measure). The mean and standard deviation of accuracy ($F1$ -measure) over 5 runs for the two different proposed methods of fusion are approximately 0.63, 0.02, and almost similar.

TABLE 2. Comparison of average classification results based on the whole dataset

| Method | Top-1 accuracy | Top-2 accuracy | Top-5 accuracy | $F1$ -measure |
|-----------------------------------|----------------|----------------|----------------|---------------|
| FCM (Algorithm1, a, b) | 0.38 | 0.46 | 0.62 | 0.35 |
| FCM (Algorithm2, a, b) | 0.36 | 0.41 | 0.59 | 0.34 |
| k-means(Algorithm3, L, a, b) | 0.38 | 0.48 | 0.65 | 0.34 |
| k-means(Algorithm4, Y, Cb, Cr) | 0.40 | 0.49 | 0.64 | 0.38 |
| Fusion (Tsallis entropy) | 0.64 | 0.70 | 0.84 | 0.63 |
| Fusion (g-calculus) | 0.65 | 0.72 | 0.85 | 0.64 |

Each image needs to its own settings, so each segmentation algorithm only works well in some images. Due to the complexity of the background of these images, there is no algorithm that is suitable for all of the images. For this reason, the fusion of the results of different algorithms, leads to a good improvement in classification accuracy.

3.3. Comparison with Other Methods In Table 3, the proposed methods are compared with some other existing methods. This comparison is obtained using the average classification accuracies (Top-1, Top-2, Top-5, and $F1$ -measure) on all tree leaf images from the dataset.

In this table, Shannon fusion approach is obtained through the proposed method with Tsallis entropy and $\beta=1$. The results of MP+B/A+CS, FCM methods were reported in [8] using a number of images in the PI@ntLeaves dataset (Scan, Scan-like and Photograph categories) and a semi-supervised FCM procedure. The results of Table 3 show that the proposed fusion methods using Tsallis entropy and g -calculus are better in terms of classification accuracy than other methods. Mutual information equations using Tsallis entropy or g -calculus have a parameter which may lead to the best consensus clustering. These results indicate that the image fusion using Tsallis entropy and g -calculus, improves the average performance of image fusion with classical Shannon entropy. Also, these results indicate that the proposed methods, can overcome the problems in unsupervised identification algorithms.

TABLE 3. Comparison of the proposed methods with other methods

| Method | Top-1 accuracy | Top-2 accuracy | Top-5 Accuracy |
|--------------------------|----------------|----------------|----------------|
| Mean shift [19] | 0.31 | 0.39 | 0.48 |
| Snakes [20] | 0.33 | 0.41 | 0.49 |
| FCM [18] | 0.38 | 0.46 | 0.62 |
| MP+B/A+CS, FCM [8] | 0.45 | 0.55 | 0.70 |
| Fusion (BISF) [17] | 0.41 | 0.49 | 0.58 |
| Fusion (Shannon entropy) | 0.44 | 0.51 | 0.68 |
| Fusion (Tsallis entropy) | 0.64 | 0.70 | 0.84 |
| Fusion (g -calculus) | 0.65 | 0.72 | 0.85 |

4. CONCLUSIONS

In this paper, we presented a new method for classifying plants in complex background images based on image segmentation fusion with maximum mutual information. Classification of plant leaf images with complex background is very challenging when there is no

presumption about the color or location of the leaves in the image.

The most important factor influencing the classification accuracy is the leaf segmentation. Segmentation algorithms require specific parameters and settings for each image. In this paper, we solve this problem by fusing the results of four different segmentation algorithms with different parameters. Experiments were performed on the PI@ntLeaves dataset. To get the best consensus segmentation we introduced new equations for maximum mutual information by using Tsallis entropy and g -calculus.

The evaluation results show that in general, the fusion of results is better than the result of a single algorithm. Each image needs its own settings, so each algorithm only works well on some images. The use of Tsallis entropy or g -calculus, results in a large improvement on the overall classification result, and offers a promising way to combine clusterings, especially in big data. It can be used to identify plants by a mobile phone as a terminal and is not just limited to leaf images. This method is fast and does not depend on the user's subjective judgment. This fast and simple method can help people to get to know the plant more quickly and better.

5. ACKNOWLEDGEMENT

The authors acknowledge the funding support of Babol Noshirvani University of Technology through Grant programs No. BNUT/370123/99 and BNUT/392100/99.

6. REFERENCES

1. Mata-Montero, E. and Carranza-Rojas, J., "Automated plant species identification: Challenges and opportunities", *IFIP Advances in Information and Communication Technology*, Springer, Vol. 481, (2016), 26-36, doi: 10.1007/978-3-319-44447-5_3.
2. Grand-Brochier, M., Vacavant, A., Cerutti, G., Kurtz, C., Weber, J. and Tougne, L., "Tree leaves extraction in natural images: Comparative study of preprocessing tools and segmentation methods", *IEEE Transactions on Image Processing*, Vol. 24, No. 5, (2015), 1549-1560, doi: 10.1109/TIP.2015.2400214.
3. Wäldchen, J. and Mäder, P., "Plant species identification using computer vision techniques: A systematic literature review", *Archives of Computational Methods in Engineering*, Vol. 25, No. 2, (2018), 507-543, doi: 10.1007/s11831-016-9206-z.
4. Wang, Z., Li, H., Zhu, Y. and Xu, T., "Review of plant identification based on image processing" *Archives of Computational Methods in Engineering*, Vol. 24, No. 3, (2017), 637-654, doi: 10.1007/s11831-016-9181-4.
5. Zhang, S., Zhang, C. and Wang, X., "Plant species recognition based on global-local maximum margin discriminant projection", *Knowledge-Based Systems*, Vol. 200, (2020), doi: 10.1016/j.knsys.2020.105998.
6. Zhang, S., Huang, W., Huang, Y.A. and Zhang, C., "Plant Species Recognition Methods using Leaf Image: Overview",

- Neurocomputing*, (2020), doi: 10.1016/j.knosys.2020.105998.
7. Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C. and Soares, J.V., "Leafsnap: A computer vision system for automatic plant species identification", in European Conference on Computer Vision, Springer, (2012), doi: 10.1007/978-3-642-33709-3_36.
 8. Cerutti, G., Tougne, L., Coquin, D. and Vacavant, A., "Curvature-scale-based contour understanding for leaf margin shape recognition and species identification", International Conference on Computer Vision Theory and Applications (2013), doi: 10.5220/0004225402770284.
 9. Nikbakhsh, N., Y. Baleghi, and H. Agahi, "Maximum mutual information and Tsallis entropy for unsupervised segmentation of tree leaves in natural scenes", *Computers and Electronics in Agriculture*, Vol. 162, (2019), 440-449, doi: 10.1016/j.compag.2019.04.038.
 10. Le, T.-L., D.-T. Tran, and N.-H. Pham. "Kernel descriptor based plant leaf identification", in 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, (2014), doi: 10.1109/IPTA.2014.7001990.
 11. Saleem, G., Akhtar, M., Ahmed, N. and Qureshi, W.S., "Automated analysis of visual leaf shape features for plant classification", *Computers and Electronics in Agriculture*, Vol. 157, (2019), 270-280, doi: 10.1016/j.compag.2018.12.038.
 12. Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H. and Tekinerdogan, B., "Analysis of transfer learning for deep neural network based plant classification models", *Computers and Electronics in Agriculture*, Vol. 158, (2019), 20-29, doi: 10.1016/j.compag.2019.01.041.
 13. Hedjazi, M.A., I. Kourbane, and Y. Genc. "On identifying leaves: A comparison of CNN with classical ML methods", in 2017 25th Signal Processing and Communications Applications Conference, IEEE, (2017), doi: 10.1109/SIU.2017.7960257.
 14. Kamilaris, A. and F.X. Prenafeta-Boldú, "Deep learning in agriculture: A survey", *Computers and Electronics in Agriculture*, Vol. 147, (2018), 70-90, doi: 10.1016/j.compag.2018.02.016.
 15. Mehri-Dehnavi, H., H. Agahi, and R. Mesiar, "Pseudo-exponential distribution and its statistical applications in econophysics", *Soft Computing*, Vol. 23, No. 1, (2019), 357-363, doi: 10.1007/s00500-018-3623-x.
 16. Nikbakhsh, N. and Y. Baleghi. "A New Fast Method of Image Segmentation Fusion Using Maximum Mutual Information", in 2019 27th Iranian Conference on Electrical Engineering, IEEE, (2019), doi: 10.1109/IranianCEE.2019.8786371.
 17. Wang, H., Zhang, Y., Nie, R., Yang, Y., Peng, B. and Li, T., "Bayesian image segmentation fusion", *Knowledge-Based Systems*, Vol. 71, (2014), 162-168, doi: 10.1016/j.knosys.2014.07.021.
 18. Arora, J., K. Khatter, and M. Tushir, "Fuzzy c-means clustering strategies: A review of distance measures", *Software Engineering*, Springer, (2019), 153-162, doi: 10.1007/978-981-10-8848-3_15.
 19. Comaniciu, D. and P. Meer, "Mean shift: A robust approach toward feature space analysis", *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 5, (2002), 603-619, doi: 10.1109/34.1000236.
 20. Chan, T.F., B.Y. Sandberg, and L.A. Vese, "Active contours without edges for vector-valued images", *Journal of Visual Communication and Image Representation*, Vol. 11, No. 2, (2000), 130-141, doi:10.1006/jvci.1999.0442.

Persian Abstract

چکیده

در این مقاله یک رویکرد جدید برای طبقه بندی و شناسایی بدون نظارت برگ درخت با استفاده از همجوشی تصاویر قطعه بندی شده ارائه شده است. با توسعه دستگاه‌های تلفن همراه و دسترسی از راه دور، شناسایی اتوماتیک گیاهان در تصاویر گرفته شده با زمینه طبیعی مورد توجه بسیاری قرار گرفته است. در اکثر روش‌های شناسایی گیاه به ویژه در تصاویر با زمینه پیچیده، قطعه‌بندی برگ‌های گیاه نقش بسیار مهمی دارد. در جایی که پیش فرض‌هایی در مورد برگ و زمینه وجود ندارد، قطعه‌بندی برگ‌ها در تصاویر با پس‌زمینه پیچیده بسیار دشوار است. علاوه بر این، با توجه به این که هر تصویر شرایط ویژه‌ای دارد، برای هر تصویر باید پارامترهای الگوریتم تنظیم شود. در این مقاله برای رفع این مشکل از همجوشی تصاویر قطعه‌بندی شده استفاده شده است. نتایج چهار الگوریتم بدون نظارت قطعه‌بندی برگ با پارامترهای مختلف استخراج می‌شود. برای ادغام نتایج این قطعه‌بندی‌ها، از یک روش سریع و بر اساس حداکثر اطلاعات متقابل استفاده می‌شود. با استفاده از آنتروپی Tsallis و g -calculus، معادلات اطلاعات متقابل تعمیم یافته به دست آمده و برای بدست آوردن بهترین قطعه‌بندی استفاده می‌شوند. برای ارزیابی روش‌های پیشنهادی از یک پایگاه داده تصویر برگ درخت در صحنه‌های طبیعی و با زمینه پیچیده استفاده شده است. این تصاویر از پایگاه داده $PI@ntLeaves$ گرفته شده و طوری اصلاح شده‌اند تا هیچ پیش فرضی نداشته باشند. نتایج تجربی نشان می‌دهد که استفاده از آنتروپی Tsallis و g -calculus در ادغام نتایج قطعه‌بندی تصویر، شناسایی گونه‌های گیاهی را بهبود می‌بخشد.
