Parameter-free Image Segmentation Based on Extreme Learning Machine

Hongwei Zhang[1] and Liuai Wu[2,†] and Yanchun Yang[3]
[1] Lanzhoujiaotong University, Lanzhou, China
E-mail: dieice2030@163.com Tel: +86-13919105278
[2] Lanzhoujiaotong University, Lanzhou, China
E-mail: computer158@163.com Tel: +86-13321210898
[3] Lanzhoujiaotong University, Lanzhou, China
E-mail: yangyanchun102@sina.com Tel: +86-13088707606

Abstract— **For the problem of spending much time on adapting parameters, a parameter-free image segmentation method based on extreme learning machine (ELM) is proposed. Firstly, each image is segmented as superpixels by simple linear iterative clustering (SLIC) with different parameters. Secondly, each superpixel segmentation result is combined with some rules, and initial segmentation results are obtained. Each initial segmentation result is evaluated, and the parameter with the best performance is selected as its class. Thirdly, in order to construct the training sets of ELM, the cooccurrence of each image is constructed, and some of its attributes are calculated as its features, and a parameter-free framework is learned by ELM. The experimental results show that the proposed method in this paper gets better segmentation results, which is closer to human annotation than other methods.**

I. INTRODUCTION

Image segmentation is that an image is classified into some parts according to similarity criterions [1]. Image segmentation is widely used in many fields including medicine, military, logistics and so on.

There are kinds of popular image segmentation algorithms based on threshold, region, graph, energy and so on. The algorithms based on threshold is simple and easy to implement. They perform well when the foreground is much different from the background in the image, but perform pool when not. The algorithms based on region like watershed algorithm are good at boundary response but sensitive to noises. The algorithms based on graph perform well in kinds of situation, but they are complex. The algorithms based on energy like level set algorithm have good effects in theory but are difficult to implement. Lots of scholars do many works about image segmentation to get better performance. Daniel Gómez, et al.[2] proposed a fuzzy image segmentation based upon hierarchical clustering to clarify what should be understood by fuzzy image segmentation. Yubing Li, et al.[3] proposed a grab cut image segmentation algorithm based on image region and improved performance in the complex background. Wang, B, et al.[4] proposed a novel image segmentation method based on pulse-coupled neural network, applied in image segmentation successfully. However, most algorithms need to adapt parameters by human which takes a lot of time. Inspired by F. Boemer, et al.[5], who proposed a parameter-free superpixel segmentation method , a parameter-free image segmentation algorithm based on extreme learning machine is proposed in this paper.

II. RELATED WORK

A. *Efficient graph-based image segmentation*

Efficient graph-based image segmentation (EGBIS) is an image segmentation algorithm proposed by Felzenszwalb, P. F. and Huttenlocher, D. P, et al. [6]. EGBIS is simple and fast while some of the boundary information can be lost. EGBIS treats each image as a graph $G=(V,E)$ where each node $v_i \in V$ represents a pixel in the image, and each weighted edge $e(v_i,v_j)\in E$ corresponds to the difference between each pairs of neighboring pixels. The weight of each edge is calculated according to some rules. And then EGBIS structure a minimum spanning tree T_{MS}, denoted by $G'=(V,E')$. Next, the edges $e'(v_i,v_j)\in E'$ ($E'\subset E$) is generated in non-descending order and combined based on some strategies. At last, the graph $G''=(R,E'')$($E''\subset E'$) is generated as the result of steps above where R represents the final combined region as well as the segmentation result.

B. *Simple linear interactive cluster*

Simple linear interactive cluster (SLIC) is a superpixel algorithm proposed by Achanta, R et al.[7]. Superpixels refers to pixel blocks of certain semantics which have similar characteristics such as brightness, color, texture and other features. It has less algorithm complexity and more memory efficiency by using superpixel. Meanwhile, SLIC is a simple and fast superpixel algorithm with good boundary response.

SLIC is based on k-means clustering, which only calculates pixels at the cluster center of $2S\times 2S$ region ($S\approx\sqrt{N/k}$, where N is the number of pixels, and k is the number of cluster centers). In lab-color and xy-spatial space, the distance D' is computed as

$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}\ , \quad (1)$$

$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\ , \quad (2)$$

$$D' = \sqrt{(\frac{d_c}{N_c})^2 + (\frac{d_s}{N_s})^2}\ , \quad (3)$$

where N_c indicates the maximum color distance, N_s indicates the maximum spatial distance. As it is hard to calculate N_c, an approximate formula with a constant m is used in practice.

$$D = \sqrt{{d_c}^2 + (\frac{d_s}{N_s})^2 m^2}\ , \quad (4)$$

m is between 1 and 40 in practice.

C. Extreme learning machine

Extreme learning machine (ELM) is a machine learning algorithm proposed by Guangbin Huang[8] in 2004. ELM is a single hidden-layer feed-forward neural network with initial weights generated at random. As a result, ELM is fast to train and behave well in many circumstances with high accuracy.

ELM trains a model according to the following formulas.

$$\boldsymbol{H\beta} = \boldsymbol{T}\ , \quad (5)$$

$$\boldsymbol{H}_{ij}^{n\times L} = \sigma(w_j \bullet x_j + b_j)\ , \quad (6)$$

$$\boldsymbol{\beta}^{L\times m} = (\beta_1 \dots \beta_n)\ , \quad (7)$$

$$\boldsymbol{T}^{n\times m} = (\boldsymbol{t}_1 \dots \boldsymbol{t}_n)^{\mathrm{T}}\ , \quad (8)$$

σ refers to the activation function, w refers to the weights, b refers to the biases, L refers to the number of hidden neural, n is the number of training samples and m is the number of class. For multi-class classification, $\boldsymbol{T}$ refers to the classes of the data set, where columns equal to the number of classes and rows equal to the number of training samples.

ELM trains a model according to the following steps.

a) Randomly initialize w_j and b_j where $j = 1, 2, \dots\dots, L$;

b) Compute $\boldsymbol{H}$ according to $\boldsymbol{H}_{ij}^{n\times L} = \sigma(w_j \bullet x_j + b_j)$ where $i = 1, 2, \dots\dots, n$;

c) Compute $\boldsymbol{\beta}$ according to $\boldsymbol{H\beta} = \boldsymbol{T}$, $\boldsymbol{\beta} = \boldsymbol{H}^{\dagger}\boldsymbol{T}$, where $\boldsymbol{H}^{\dagger}$ is the Moore-Penrose pseudoinverse of $\boldsymbol{H}$;

d) Obtain trained ELM model.

III. METHOD

Traditional image segmentation algorithms need to set parameters by human and it is difficult to adapt parameters since different images are suitable for different parameters. Since then, a parameter-free graph-based image segmentation algorithm based on ELM is proposed in this paper and it is followed as the following steps.

a) Segment each image into different number of superpixels with SLIC

Each image is segmented into different number of superpixels with SLIC. There are two parameters needed in SLIC, which are the number of clusters k and the compactness constant m. We choose $m = 20$ and get k by ELM. To get better adaptation to different size of image, we use the average superpixel size k_s to distinguish the class of each image instead of k. SLIC gets similar result with similar k value so we choose k_s from the set $K_{a,b} = \{4n^2 : a \le n \le b, n \in \mathrm{N}\}$ where we choose $a = 5$, $b = 10$.

b) Combine each result by superpixel segmentation with advanced EGBIS and get the initial segmentation result

Combination strategy is executed for the superpixel segmentation result by using advanced EGBIS. Graph $G = (V, E)$ is constructed where the sets of superpixels is denoted by V and the sets of edges between each neighboring superpixels is denoted by E. The weights are calculated by the following formula.

$$w(e) = \sqrt{(l_1 - l_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2}\ , \quad (9)$$

The average brightness values of superpixel v_1 and v_2 are denoted by l_1 and l_2, where l_1 and l_2 range from 0 to 100. The average values from red to green of superpixel v_1 and v_2 are denoted by a_1 and a_2, where a_1 and a_2 range from -127 to 128. The average values from yellow to green of superpixel v_1 and v_2 are denoted by b_1 and b_2 where b_1 and b_2 range from -127 to 128. Then graph G is used to construct a minimum spinning tree $T_{\mathrm{MS}}(V, E)$. The edges are sorted in non-descending order, combined based on the following strategies when each region only contains one single superpixel at first.

The internal difference is defined as D_{INT}

$$D_{\mathrm{INT}} = \max(w(e)), e \in T_{\mathrm{MS}}(V, E)\ , \quad (10)$$

The difference between is defined as D_{DIF}

$$\begin{aligned} &D_{\mathrm{DIF}} = \min\left(w\left(v_s, v_t\right)\right), \\ &v_s \in V_i, v_j \in V_j, \\ &V_i \subset T_{\mathrm{MS}}\left(V, E\right), V_j \subset T_{\mathrm{MS}}\left(V, E\right) \end{aligned} \quad (11)$$

The minimum is defined as D_{MINT}

$$D_{\mathrm{MINT}} = \min\left(D_{\mathrm{INT}}(V_i)+\tau(V_i)\right), \quad D_{\mathrm{INT}}(V_j)+\tau(V_j) \tag{12}$$

where the threshold function is denoted by $\tau(V_i)$

$$\tau(V_i)=\frac{t}{|V_i|}, \tag{13}$$

where $|V_i|$ is the number of superpixels V_i and t is the threshold constant, and a new threshold function is proposed

$$\tau(V_i)=\frac{\sigma\bullet k_s}{|V_i|}, \tag{14}$$

where σ is a simple constant and in this paper $\sigma=2000$.

The combination rule is defined as the following formula.

$$D\left(V_i,V_j\right)=\begin{cases} true, D_{\mathrm{DIF}}\left(V_i,V_j\right)<D_{\mathrm{INT}}\left(V_i,V_j\right) \\ false, D_{\mathrm{DIF}}\left(V_i,V_j\right)\geq D_{\mathrm{INT}}\left(V_i,V_j\right)\end{cases}, \tag{15}$$

c) Calculate the variation of information of each initial segmentation result and choose the average superpixel size with the minimum variation of information as its class.

Each final segmentation result after combination is calculated with variation of information(VI) [9].

$$\mathrm{VI}\left(S,S'\right)=H\left(S\right)+H\left(S'\right)-2I\left(S,S'\right), \tag{16}$$

The smaller VI we get, the better segmentation result we get. And we choose the k_s with the smallest VI as the image's class.

d) Construct the co-occurrence matrix and calculate its energy, contrast, entropy and correlation for each image as its own feature

For each image, the co-occurrence matrix is constructed where $d=1$, $\theta\in\left\{0^\circ,45^\circ,90^\circ,135^\circ\right\}$ and calculate its attributes for each image as its own feature.

e) Learn a model using the training set constructed above with ELM and get the parameter-free framework

Then a model is learned using the training set constructed above. First, we initialize the weights and biases at random; then we calculate the initial matrix according to formula (6). Next, we calculate the output weights matrix according to formula (5). We can simply get the output weights matrix by the Moore-Penrose pseudoinverse. At last, with the obtained output weights matrix, we get the ELM model. And we can segment different images without adapting parameters by the trained ELM model.

IV. EXPERIMENTAL RESULT

In our experiment, we use the data set BSDS500(Berkeley Segmentation Data Set and Benchmarks 500)[10], which is provided by University of California, Berkeley. BSDS500 includes 500 different images from varieties of fields with ground truth by human annotations, which have kinds of difficulty to segment.

To verify our method, three methods are compared with the proposed method, which are Efficient Graph-Based Image Segmentation (EGBIS) algorithm proposed in the paper [6], Dynamic Color Gradient Thresholding (DCGT) algorithm proposed in the paper [11] and Fast and Robust Fuzzy C-Means (FRFCM) algorithm proposed in the paper[12].

Two results are chosen as the following.

Fig. 1 Experiment result. top left: Original image; top right: Segmentation result by human; middle left: Segmentation result in EGBIS; middle right: Segmentation result in DCGT; bottom left: Segmentation result in FRFCM; bottom right: Segmentation result in the proposed method

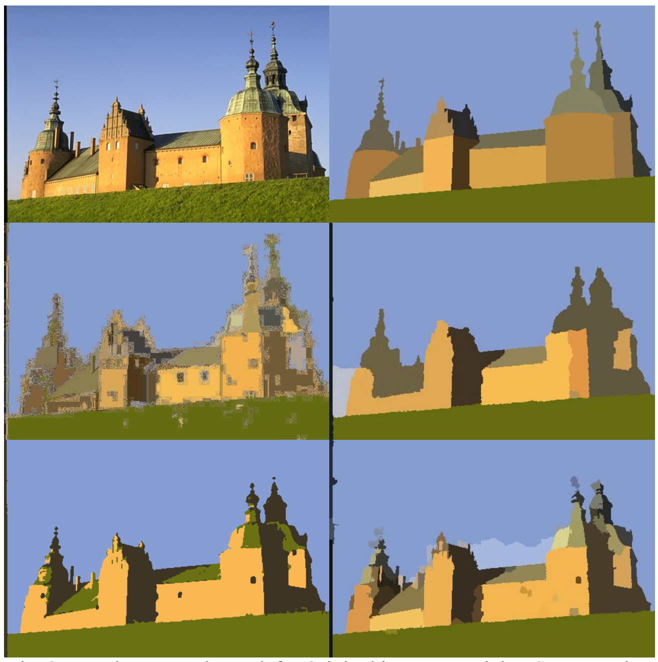

Fig. 2 Experiment result. top left: Original image; top right: Segmentation result by human; middle left: Segmentation result in EGBIS; middle right:

Segmentation result in DCGT; bottom left: Segmentation result in FRFCM; bottom right: Segmentation result in the proposed method

As we can see in the segmentation results, traditional EGBIS segments object well from the image, however with poor boundary accuracy. Compared with traditional EGBIS, DCGT works better on the boundary but it still has blurred boundary which is unacceptable. FRFCM performs well on the boundary but there exists some over-segmentation. Compared with other methods, the proposed method performs better, which is closer to the human annotation with good boundary response and little over-segmentation.

V. CONCLUSIONS

This paper adapts traditional EGBIS, combined with SLIC and ELM, and a parameter-free image segmentation method based on ELM is proposed. Superpixel is used instead of pixel and the combination strategy is changed which adapt the threshold to make it suitable for different images. ELM is used to learn a model that free human from adapting parameters and good performance is obtained. Most results perform well in vision, however some images with complex background don't get good results. To improve the performance of complex background is the direction in the future.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grants 61562057, 61162016 and 61462059), the Program for Changjiang Scholars and Innovative Research Team in University (IRT_16R36).

REFERENCES

[1] Jianbo Shi, Malik J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.

[2] Gómez Daniel, Yáñez Javier, Guada Carely, *et al.* Fuzzy image segmentation based upon hierarchical clustering[J]. Knowledge-Based Systems, 2015, 87: 26-37.

[3] Li Y., Zhang J., Gao P., *et al.* 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC).Grab Cut Image Segmentation Based on Image Region[C]//2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), 27-29 June 2018, 2018: 311-315.

[4] Wang B., Chen L. L., Wang M. Novel image segmentation method based on PCNN[J]. Optik, 2019, 187: 193-197.

[5] Boemer Fabian, Ratner Edward, Lendasse Amaury. Parameter-free image segmentation with SLIC[J]. Neurocomputing, 2018, 277: 228-236.

[6] Felzenszwalb Pedro F., Huttenlocher Daniel P. Efficient Graph-Based Image Segmentation[J]. International Journal of Computer Vision, 2004, 59(2): 167-181.

[7] Achanta R., Shaji A., Smith K., *et al.* SLIC Superpixels Compared to State-of-the-Art Superpixel Methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.

[8] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541).Extreme learning machine: a new learning scheme of feedforward neural networks[C]//2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), 25-29 July 2004, 2004, 2: 985-990 vol.982.

[9] Rand William M. Objective Criteria for the Evaluation of Clustering Methods[J]. Journal of the American Statistical Association, 1971, 66(336): 846-850.

[10] Arbelaez P., Maire M., Fowlkes C., *et al.* Contour Detection and Hierarchical Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.

[11] Balasubramanian Guru Prashanth, Saber Eli, Misic Vladimir, *et al.* Human Vision and Electronic Imaging XIII.Unsupervised color image segmentation using a dynamic color gradient thresholding algorithm[C]//Human Vision and Electronic Imaging XIII, March, 2008, USA. Washington:International Society for Optics and Photonics, 6806: 68061H.

[12] Lei T., Jia X., Zhang Y., *et al.* Significantly Fast and Robust Fuzzy C-Means Clustering Algorithm Based on Morphological Reconstruction and Membership Filtering[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(5): 3027-3041.