An Identification of Abnormal Images using semi-supervised Segmentation Model

V.Manogajapathi, M.Frankline, R.Karthikeyen

Students-Dept. of Computer Science Christ College of Engineering and Technology Pondicherry, India

Abstract

The objective of semi-supervised image segmentation is to obtain the segmentation from a partially labeled image. By exploit the image numerous structure in labeled and unlabeled pixels, semi-supervised methods propagates the user labelling to unlabeled data, thus reducing the need for the user labeling. Several semi-supervised research mechanism have been proposed in the observation. In this paper, we consider the offending of segmentation of large collections of images and the classification of images by related diseases. We are detecting abnormal images by the process of segmentation and classification. The segmentation used in this paper has two advantages. First, user can specify their own values by highly controlling the segmentation. Another is, at initial stage this mechanism needs only minimum tuning of model parameters. Once initial tuning process is done, the setup can be used to automatic segment a large collection of images that are distinct but share identical features. And for classification of diseases, a numerous research method, called parameter-free semi-supervised local Fisher discriminant analysis is used. This method preserves the global structure of labelled samples in addition to separating unlabelled samples in various classes from each other. The semi-supervised method has a systematic form of the globally optimal solution, which can be computed efficiently by Eigen decomposition. Espousal experiments on various collections of biomedical images suggest that the proposed mechanism is effective for segmentation with classification and is computationally adequate.

Keywords

Biological image segmentation, semi-supervised segmentation, multiple imaging, ,microscopy images.

1. Introduction

Image segmentation is a challenging task and remains an open problem in image processing. Unsupervised methods explore the intrinsic data structure to segment the image into regions with different statistics. However, these methods often fail to achieve the desired result, especially if the desired segmentation includes regions with very different characteristics. On the other extreme, supervised image segmentation methods first learn a classifier from a labeled training set. Although these methods are likely to perform better, marking the training set is very time consuming. Semi-supervised image segmentation methods circumvent these problems by inferring the segmentation from partially labelled images. The key difference from supervised learning is that semi supervised methods utilize the data structure in both the labeled and unlabeled data points. Hence, the main advantage of semi-supervised image segmentation methods is that they take advantage of the user markings to direct the segmentation, while minimizing the need for user labeling. Image segmentation is used in many areas, including computer vision, computer graphics, and medical imaging, to name a few. Fully featured automated image segmentation has many fundamental difficulties and is still a very hard problem. Therefore, the segmentation problem is hostile modeled if no additional knowledge about the desired segmentation is given. In many applications, such as cell segmentation in organ segmentation images and organ microscopy images in medical images, the kind of segmentation and objects of interest are known in advance.

2. Materials and Methods

A. Segmentation

Detach the foreground object from the background object from a static image engages by determining the both partial and fulll pixel coverages, also known as extraction a matte. The input image to be pre-segmented into 3 sectors: foreground, background and unknown, which is called a trimap in the previous approaches. Inside the unknown region partial values are calculated for unknown regions. Images with large portion of semi-transparent foreground will fail in pre-segmentation based approach where the tri map is difficult to create even manual methods. In this paper we include both the matting and segmentation problem together and propose a unified optimization approach based on the Belief Propagation. In the image every pixels opacity value are estimated repeatedly, based on a small specimen of foreground and background pixels analysed by the user. Experimental outputs show that compared with previous methods, our mechanism is more efficient to extract high quality mattes for foregrounds of the images with significant semi-transparent regions. The observed image

$$I(z)\ (z = (x, y))\ldots\ldots \quad (1)$$

Foreground image- $F(z)\ldots\ldots$ (2)

Background image- $B(z)\ldots\ldots$ (3)

Equating all equations 1,2and 3……

Alpha map: $I(z) = \alpha z F(z) + (1 - \alpha z)B(z)$

Where, αz can be any value in [0, 1].

However, even knowing the background color is still insufficient to fully constrain the problem thus some

Manuscript received February 5, 2015
Manuscript revised February 20, 2015

simple constraints are made, which require expert level parameter tuning and can fail on fairly simple foregrounds. The technique by introducing statistical methods based on representative foreground and background samples.

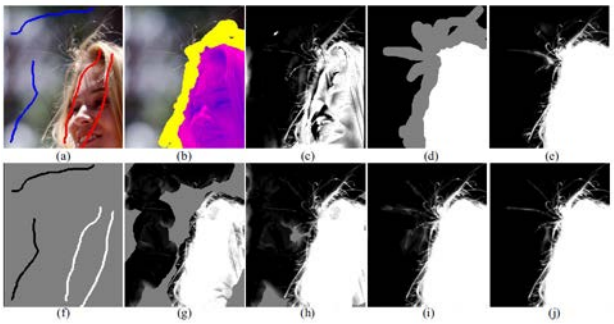

Figure-1.a Images for segmentation

The mechanism produces segmentations of images by classifying each image pixel as non-vessel or vessel, based on the pixel's featured vector. Vectors are composed of the pixel's intensities and continuous two-dimensional Morlet wavelet transform replies taken at the multiple scales. The Morlet wavelet is capable of tuning them to unique frequency, thus allowing vessel enhancement and noise filtering in a unique step.

ALGORITHM:

Step 1: Calculate X1,1(autonomous of ג and μ).

Step 2: Calculate Y1,1 (reliant on (ג

Step 3: Resolve the linear system

{Val – ST Ω1 SΩ1Y1,1} a1|n

Where, n = 1,2,…,N-1.

Step 4: Calculate X2,2 and X2,1(Autonomous of ג and μ)

Step 5: Calculate Y2,1 and Y2,2(reliant on ג and μ)

Step 6: Resolve the linear system

{Val – ST Ω2 SΩ2Y2,2}a2|n = ST Ω2 SΩ 2Y2,1a1|n

Where n = 1,2,…,N-1.

Algorithm 1: Algorithm for Segmentation

A. Similarity Measures

Two types of similarity measures namely, photometric and geometric are considered. The photometric is based on pixel locations, whereas the geometric is based on color features.

For each pixel i ∈ γ^s, its geometric neighbor G_i^(s,s)⊂γ^s is defined as,

G_i^(s,s):={j ∈ γ^s:0<‖i-j‖_∞≤r_g }

where r_g>0 is a constant controlling the size of the window, and ‖.‖_∞ is the vector maximum norm. We often set the r_g=1 so that a 3×3 window around pixel i is used. Note that i∉G_i^(s,s) and the geometric neighbor is not defined across two images.

The geometric similarity g_(i,j)^(s,s) is defined as

g_(i,j)^(s,s): ={(ce^(-(‖i-j‖_2^2)/(σ_i^2)), if j∈ @ 0: otherwise)┤ g_(i,j)^(s,s)

where c is a normalization constant such that ∑ ⟦j ∈ γ^s ⟧ g_(i,j)^(s,s) ≡1, and σ_i^2 is computed as the sample variant of the geometric location within G_i^(s,s).

For each pixel ∈ γ^s , let F_i be its feature vector. We use the RGB values over a 3×3 window around the pixel to construct a featured vector of dimension 27. Then, the within-image photometric neighbor ⟦ p⟧ _i^(s,s)⊂γ^s is defined to be the top 4 pixels within the 17×17 window around pixel (excluding pixel itself), whose featured vectors are nearer to F_i (in Euclidean norm). Using a larger window size allows us to connect photo metrically similar pixels that are further apart. However, doing so it will increase the computational cost. The choice of the size 17 × 17 is a balance between both extremes. The within-image photometric similarity is defined as p_(i,j)^(s,s): ={(ce^(-(‖F_i-F_2 ‖_2^2)/(σ_i^2)), if j∈ @0 ,otherwise)┤ P_(i,j)^(s,s)

where c is a normalization constant such that ∑ ⟦j ∈ γ^s ⟧ p_(i,j)^(s,s) ≡1, and ρ_i^(2)is computed as the sample variance of the photometric features within P_i^(s,s).

c. Optimization

Each labeled pixel of the image the unlabeled pixels are connected through a sequence of directed edges, each of which labeled or unlabeled connects a pixel to one of its neighbors in the same image or it may be different images.

Let u_s for s=1,2 be two given multichannel images. Their sizes are not necessarily the same. Let ⟦ γ ⟧ ^s be the group of all pixels in image u^s. Let Ω^s be the group of all unlabeled pixels in image u^s. Let Γ^s be the set of pixels in image u^s labeled to one of the classes M by the user. Thus ⟦ γ⟧ ^s=Ω^s∪Γ^s Here, we allow both images to contain unlabeled and labeled pixels for the sake of generality. The set of labeled pixels Γ^s is divided into ⟦ Γ⟧ ^(s⁄1),….,Γ^(s⁄m) , where is the group of pixels are labeled with class m, for m=1,….,M.

Let s^'=2 if s=1, and let s^'=1 if s=2, so that s^' is an index reference to an image differ from the image indexed by s. For each pixel i ∈ γ^s and each pixel j ∈ γ^t , let ω_(i,j)^(s,t)≥0 be a similarity score between the pair of pixels, for s,t=1,2. When t=s, the similarity ω_(i,j)^(s,t) is computed within image ⟦ u⟧ ^s; when t=s^', the similarity is computed across two images. For each i ∈ γ^s , it is assumed that the similarity values are normalized such that

$$\sum_{j \in \gamma^s} \omega_{i,j}^{s,s'} + \sum_{j \in \gamma^{s'}} \omega_{i,j}^{s,s'} \equiv 1$$

For each pixel i ∈γ^s, let N_i^(s,t) ⊂ γ^(t)be a set of pixels in image u^t, which is called the neighbor of i in u^t. The within-image neighbor and the across the image neighbor are defined respectively by N_i^(s,s) ：=G_i^(s,s)∪P_i^(s,s) and 〖 N〗_i^(s,s') :=P_i^(s,s'). Presumably, these pixels have high similarity scores with i. The basic idea of the method is that the memberships of identical pixels should be identical. For each unlabeled pixel i∈Ω^s, the membership to class m inferred from its neighbors is the weighted average, i.e.

$$\sum_{j \in N_i^{s,s}} \omega_{i,j}^{s,s} \alpha_j^{s/m} + \sum_{j \in N_i^{s,s'}} \omega_{i,j}^{s,s'} \alpha_j^{s'/m} \equiv 1$$

3. RESULTS AND DISCUSSION

In this sector, we evaluate the performance of the proposed multiple-image segmentation model tentatively using three image data sets. All algorithms are developed in MATLAB 2013 with a Core i7 3.07-GHz machine.

B. Image Data sets

In the first analysis, a set of 56 breast cancer cell images of various sizes in is used. Overall, all these images have identical features. However, some contain benign cells, whereas some contain both malignant and benign cells, so that there are some variant among them. The manual segmentation of the cell nuclei which is used as the ground truth to evaluate the accuracy of our derived method. In the second analysis, a group of 20 retinal images of size292x283, which is called DRIVE, is used. Some images visual some clue of early diabetic retinopathy, whereas others do not in the third experiment, a set of 30 cross sections of a human retina with size 256x384in is used to experiment. Color exaggeration due to staining and shape variation of the layers makes the segmentation task very assert.

Method: The 15 simulated images are segmented together at a time using the scheduled multiple-image model, with the image in a fully labelled specimen.

Results: We observe that the scheduled mechanism is able to differentiate much of the alien object as retina, and the accuracy is not lowered at all. This is mainly requisite to the maximum colour contrast between the yellow alien object and the blood vessels and due to the presence of the yellow optic disc on the right in the training data. Observe that the accuracy drops between Image4 and Image5. The minimization in the accuracy is mainly due to the misclassification of pixels in the reddish circular disc. For the first four images, the scheduled mechanism is able to classify much of the alien object as retina. It shows that the proposed method has a certain degree of tolerance. However, although the intensity of the alien object becomes too close that of the blood vessel pixels (i.e., Image5 to Image 10), the object is classified as blood vessel by the properties of the scheduled mechanism.

C. Examination Methods

For each experiment, the labeled image is served as one, and the remaining images are segmented using the proposed semi-supervised multiple-image model (MI). For the third experiment, only 50% of the labels are used for computing P2,1

(i.e., $|S| = 0.5M \min_{1 <= m <= M} | 1|m|$**).**

TABLE I. PARAMETER VALUES FOR TUNING MODEL

Parameters	*values*
λ	1
μ	1
α	0 or 0.5 or 1
Pigment Eigen Values	0.6

The parameters ƛ and μ for the one labeled image and one unlabeled image similarity measures are calibrated. Then it used throughout the whole image data set. Before segmenting the images, a color variation mark is applied to each unlabelled image so that its color histogram matches that of the labeled image. This is to correct the difference in contrast and brightness between images. To validate our method, we compare it with 4 other methods, namely, the classic support vector machine (SVM), the -nearest neighbor (KNN), the –means (KMeans), and the semi-supervised -means (SSK Means). The first two methods are supervised methods; the third is unsupervised; the fourth is semi-supervised. For the supervised and semi-supervised methods, the labelled pixels are used as the training set. The accuracy, which is computed as the no. of correctly classified pixels divided by the total number of pixels, and the –measure value, which is processed as the harmonic mean of precision and re-entry for the foreground object, i.e

$$F = \frac{2 \text{ precision x recall}}{2 \text{ precision + recall}}$$

are used for measuring of segmentation accuracy. The measure takes false positives, true positives, and false negatives into account. A higher F suggests a better performance. To further illustrate the developed model, we compare it with a state-of-the- art blood vessel segmentation algorithm MLVESSEL using the breast cancer cell and DRIVE data sets.

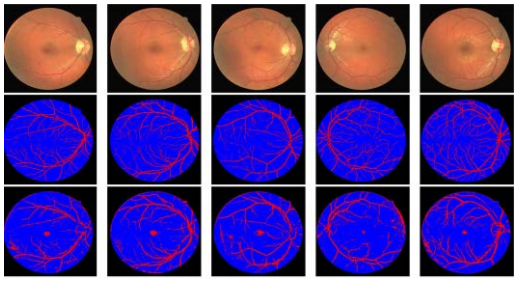

Figure-2.a Segmentation of Retinal eye blood vessel image

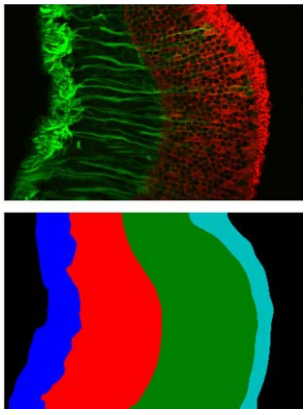

Figure:- 2.b Segmentation image with original and truth image

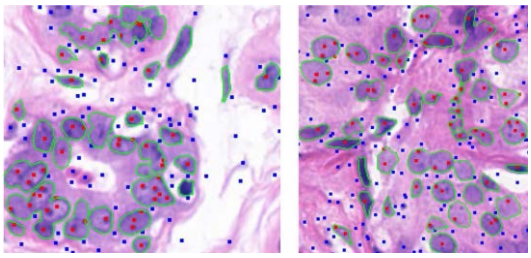

Figure:- 2.c Segmentation in Two breast cancer cell images(benign and malignant) by using scale invariant feature transform (SIFT) keypoint detection method.

D. Simulated output of Abnormal images

In the stationary of retinal disease, some hemorrhages and lesions, which are usually yellow and red objects, may be found in a retinal image. In this observation, we aim to test the robustness of the proposed system to such kinds of alien objects. The blood vessels are extracted as the primary object. Hence, it is trusted that these alien objects are put into the background and do not hinder the identification of the blood vessels.

1) Images: The two retinal images in are used. The first one is given as the labeled image. The second one is to generate 2 sets of 10 simulated abnormal images. In the first set of images, which are referred to as the yellow set, a circular disc with maximizing the degree of yellowness is just a posed to the centre of each image. Three of the simulated images in the yellowish set are shown in the first row in Fig.2.a [respectively Fig. 2. b]

2) Method: The fifteen simulated images are segmented one at a time using the newly proposed multiple image mechanism, with the image in Fig. 2 fully labelled sample.

3) Results: The segmentations of 3 of the simulated images in the yellow set are displayed in the second row in the output image. We observe that the scheduled mechanism is able to differentiate much of the alien object as the retina, and the accuracy is not affected at all. This is due to the maximum colour contrast between the yellow alien object and the blood vessels and due to the presence of the yellow optic disc on the right in the training data. The segmentations of two of the simulated images in the reddish group are displayed in the second row in output image. The accuracy of the segmentations is achieved efficiently. Observe that the accuracy drops between Image 4 and Image5. The decrease in the accuracy is mainly due to them is classification of pixels in the reddish circular disc. For the first four images, the proposed method is able to differentiate much of the alien object as retina. It shows that the newly proposed model has a certain degree of tolerance. However, when the intensity of the alien object becomes too close to that of the blood vessel pixels (i.e., Image 5 to Image 10), the object is classified as blood vessel by the properties of the proposed model. The figure (Figure 3.) give nbelow shows the experimental results of Multiple Images (Retina with focusing fovea region, Retinal Skin, Cell Image). In which, experiments are done in a simulation tool MATLAB 8.

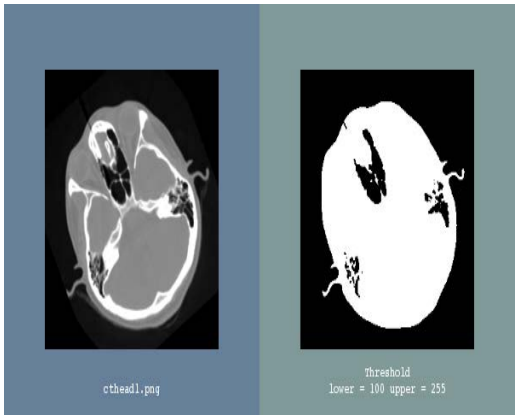

Figure:- 3.a Gray scle to binary conversion

Figure:- 3.b Labeled components

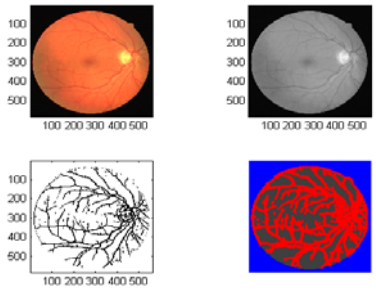

Figure:- 4 Refernce images for measuring similarity with input images

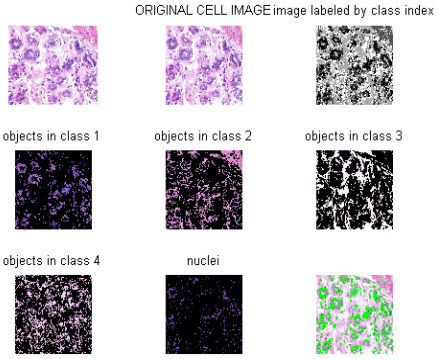

Figure:- 5 Simulated output image with labeled and unlabeled pixels to different alien object classes

4. CONCLUSION

In this paper, we have proposed and implemented a semiautomatic optimization model for segmentation of multiple images. The mechanism has a quadrate objective function and linear constraints. Due to the discrete maximum/minimum principles, the optimality conditions simply abscess down to solving linear systems (as opposed to the nonlinear Karush–Kuhn–Tucker systems for general quadratic programming problems). In our applications, the two parameters can be easily tuned. Once tuning is done at the initial stage, the setup can be used to segment all other images within the collection automatically soon enough. The quality of the results is also highly efficient. However, it relies on the logical assumption that the different object classes can be separated in the feature upcomings and that the user-supplied patterns can represent each class well. Experimental results on synthetic data and three well-known gene expression data sets demonstrate the effectiveness of the proposed pSELF algorithm. In this paper, the intrinsic structure preserved by pSELF is only the global structure of samples. Investigating that whether pSELF can preserve local structures together with unlabeled samples is an interesting future work. And also we can extend it as remote automatic diagnosis system.

References

[1] Michael K. Ng, and Andy M. Yip, ”A Semi-supervised Segmentation Model for Collections of Images”, IEEE trans. on Image processing, vol.21,no.6, June 2012.

[2] Y. Boykov and V. Kolmogorov, “An experimental comparison of max-flow/min-cut algorithms for energy minimization in vision,” IEEE Transactions. Pattern Anal. Mach. Intell., vol. 26, no. 9, pp. 1124–1137 ,Sep. 2004.

[3] A. Protiere and G. Sapiro, “Interactive image segmentation via adaptive weighted distances,” IEEE Trans. Image Process., vol. 16, no.4,pp. 1046–1057, Apr. 2007.

[4] J. Guan and G. Qiu, “Interactive image segmentation using optimization with statistical priors,” in ECCV, Graz, Austria, 2006.

[5] M.K.Ng, G.Qiu, and A.M.Yip, “Numerical methods for interactive multiple-class image segmentation problems,” Int. J. Imag. Syst. Technol., vol. 20, no. 3, pp. 191–201, Sep. 2010.

[6] F. Wang, C. Zhang, H. Shen, and J. Wang, “Semi-supervised classification using linear neighborhood propagation,” in Proc. CVPR, 2006,pp. 160–167.

[7] O. Delalleau, Y. Bengio, and N. Le Roux, “Efficient non-parametrical function inductions in semisupervised learning,” in Proc. 10th Int.Workshop AI Stat., 2005, pp. 96–103.

[8] J. Guan and G. Qiu, “Modeling user feedback using a hierarchical graphical model for interactive image retrieval,” in Advances in Multimedia Information Processing, ser. Lecture Notes in Computer Science, H.H.-S. Ip,H.Leung, O. C. Au,M.-T. Sun,W.-Y. Ma, and S.-M.Hu, Eds. Berlin, Germany: Springer-Verlag, 2007, pp. 18–29.

[9] N. Houhou, X. Bresson, A. Szlam, T. F. Chan, and J.-P.Thiran, Semisupervised segmentation based on non-local continuous min-cut,” in Proc. SSVM, 2009, pp. 112–123.

[10] Y. Li, J. Sun, C. K. Tang, and H. Y. Shum, “Lazy snapping,”ACM Trans. Graph., vol. 23, no. 3, pp. 303–308, Aug. 2004.

[11] M. K. Ng, G. Qiu, and A. M. Yip, “Numerical methods for interactive multiple-class image segmentation problems,” Int.J. Imag. Syst. Technol., vol. 20, no. 3, pp. 191–201, Sep.2010.

[12] C. Rother, V. Kolmogorov, and A. Blake, “Grabcut: Interactive foreground extraction using iterated graph cuts,” ACM Trans. Graph., vol.23, no. 3, pp. 309–314, Aug. 2004.

[13] S.T. Roweis and L.K. Saul, “Nonlinear Dimensionality Reduction by Locally Linear Embedding,” Science, vol.290, no. 5500, pp. 2323-2326, 2000.

[14] A. Protiere and G. Sapiro, “Interactive image segmentation via adaptive weighted distances,” IEEE Trans. Image Process.,vol. 16, no. 4, pp. 1046–1057, Apr. 2007.

[15] Y.P. Li, X.H. Hu, H.F. Lin, and H.Yang, “A Framework for Semi-supervised Feature Generation and Its Applications in

[16] Biomedical Literature Mining,” IEEE/ACM Trans. Computational Biology and Bioinformatics, vol. 8, no. 2, pp.294-307, Mar./Apr.2011.