AUTOMATIC SEGMENTATION OF SKIN LESION IMAGES USING EVOLUTIONARY STRATEGY

Ning Situ[1]*, Xiaojing Yuan*[2]*, Nizar Mullani*[3]*, George Zouridakis*[1]

1. Computer Science Department, University of Houston
2. Engineering Technology Department, University of Houston
3. Translite LLC

ABSTRACT

Skin cancer has been the most common and represents 50% of all new cancers detected each year. If detected at an early stage, simple and economic treatment can cure it mostly. Accurate skin lesion segmentation is critical in automated early diagnosis system. It provides both the shape feature and the region of interest for texture analysis. In this paper, we propose an evolutionary strategy based segmentation algorithm and apply it to skin lesion. It can detect the lesion automatically without setting parameters manually. The method is validated by experiments and comparisons with manually segmentation by an expert and algorithms developed in [1, 2].

***Index Terms*—** Evolutionary Strategy, image segmentation, skin lesion detection, objective function

1. INTRODUCTION

Early detection of cancerous skin lesion has been agreed to be very important due to the wide spread of skin cancer as well as the economic and successful treatment if detected early. Malignant melanomas, the deadliest form of all skin cancers, has cure rate of higher than 95% when detected at an early stage[3]. Zouridakis, et al. [1, 2] developed an automatic malignancy detection system based on the size difference in skin lesion images from two imaging modalities: Cross-polarization Epiluminescence Microscopy (XLM) and Transillumination Epiluminescence Microscopy (TLM). Even though for most of the images, the systm achieves error rate less than 15% when compared with manual segmentation, for lesion images with high background noise or weak edge information, all algorithms give very poor performance(error rate higher than 40%). In addition, all segmentation methods performance also depend on initial values.

In this paper, we propose to use Evolutionary Strategy(ES) based segmentation for skin lesion detection. Like other evolutionary computation algorithms, such as genetic algorithm, ES has the property of seeking global optimum and getting out of local optimum automatically. Yuan et al. [4] applied ES to feature identification in natural and synthesis of images with multiple features. It has also been applied to image registration [5]. To apply the ES algorithm to skin lesion image segmentation, we formulated the segmentation problem as a search problem similar to [4]. The lesion area is segmented by an ellipsoid, whose parameters are optimized by ES algorithm with respect to the defined objective function. The method is validated by experiments on the same data set as used in [1, 2]. The experiments show ES based algorithm has a better performance than the algorithms in [1, 2].

The project is partially supported by UH-GEAR grant and funds from Translite LLC.

2. RELEVANT WORKS

Zouridakis, et al [1, 2] proposed an automatic skin lesion malignance detection system based on size difference of XLM image and TLM image. The XLM imaging modality captures surface pigmentation, while the TLM imaging modality visualizes both surface pigmentation and the increased blood volume and vasculature around a lesion if present. Based on skin physiology, for cancerous lesion more vascular activity can be observed, resulting in bigger lesion area for TLM images than that of XLM images [1, 2]. Each image undergoes same preprocessing procedures including masking, cropping, color converting and hair removal. After that, four segmentation techniques are employed to identify the lesion area: sigmoid, PCT, PCT+sigmoind, and fuzzy c-means. The segmentation results of these techniques are then selected by a scoring system, and the boundary of the selected one will be smoothed.

The performance of the malignancy detection system relies on these four segmentation methods. However, these four segmentation methods all have their own limitations when dealing with skin lesion images. The sigmoid method may fail if the histogram of the red channel can not be fitted by two Gaussian curves. The PCT is not based on any statistic property, and in some cases it can not generate enough contrast between the lesion and the background. The combined sigmoid and PCT produces better result, but it still suffers from the same problem as applying PCT only. Moreover, the threshold determined by this method is easily influenced by some

artifacts. The fuzzy c-mean can generally produce good performance for most lesion images. However, its performance degrades significantly for images with very small lesions. The poor segmentation introduces a lot of uncertainties when extracting precise lesion size for cancerous decision. A scoring system is designed to overcome the limitations of different segmentation methods by selecting the best segmentation based on majority vote mechanism. It looks at the differences between segmentation results and the edge strength and selects the one that has majority vote and the strongest edge.

In this paper, we present an ES-based segmentation algorithm for skin lesion detection. Because of the inherent properties of ES algorithm, the ES-based segmentation algorithm has three distinct advantages: (1)It is an unsupervised segmentation algorithm whose performance does not depend on initialization or threshold values; (2)robustness to artifacts and noise; (3)it is based on the statistical property of the image. Because of these properties, images fed into the ES-based segmentation algorithm do not need to go through the full pre-processing steps mentioned above. In specific, ES-based segmentation method does not need hair removal procedure. In addition, ES-based method does not require manually selected threshold and is robust to the initial values.

Evolutionary Strategy (ES) is one kind of evolutionary computation that has been applied to various optimization problems. Genetic Algorithm(GA), another evolutionary computation technique, is the most popular and has already been used in the area of image segmentation [6, 7] and in the specific area of medical image segmentation[8, 9]. The major difference is that ES gene evolves in the real number domain, which avoids information loss due to the binary coding representation of GA. Yuan et al. [4] applied ES successfully to feature identification of natural and artificial images. It has also been applied to image registration [5].

To apply the ES algorithm to skin lesion image segmentation, we formulated the segmentation problem as a search problem similar to [4]. The lesion area is segmented by an ellipsoid, whose parameters are optimized by ES algorithm with respect to the defined objective function. The main reason we chose to use an elliptic template for segmentation is because it can be fully defined by five parameters. This makes it easy to implement an ellipsoid region based objective function.

3. ES ALGORITHM OVERVIEW

Evolutionary Strategy (ES) is a random search based optimization technique. We chose ES as our optimization methods because of its two properties. First of all, ES algorithm converge to global optimum instead of local optimum. Secondly, ES is formulated for optimization of real number functions.

The basic elements for using ES include: (1) A population (more than one) of candidate solutions; (2) A measure over each member of the population (or candidate solution) referred to as fitness/objective function; (3) A "SELECTION" operator that differentiates between members of a population based on the fitness value; (4) A "MUTATION" operator that makes random changes to a member of the population (corresponding to asexual reproduction in biology evolution); (5) A "RECOMBINATION" operator that generates a new organism (or individual solution) by combining "genetic material" from random selected members of the population. Fig.1 shows the evolution of candidate solutions (i.e., organisms) in one ES generation. From one generation to the next, the candidate solutions (organisms) evolves to give better and better fitness values. The gene pool stores the candidate solutions

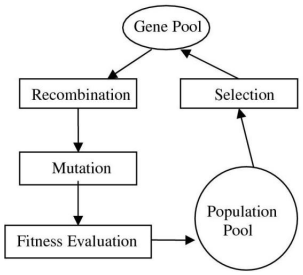

Fig. 1. Flowchart of ES

(μ)selected from population pool. They are used as parents to generate candidate solutions of the next generation. Each gene can be represented by their objective variables (which defines the dimension of the organism), and control variables (which defines the standard deviation and auto-correlation of the objective variables). The control variables are randomly generated Gaussian distribution. Each objective variables will go through mutation and recombination operation to generate λ organisms (candidate solutions), which are added to the population pool($(\mu + \lambda)$ organisms). The fitness of a newly generated offspring is evaluated by the objective function. μ organisms with good fitness values will be selected as parents for the next generation. This finishes one loop of one generation of the ES algorithm which is summarized in Fig.1.

4. APPLY ES TO SEGMENT SKIN LESION IMAGE

The skin lesion detection problem can be formulated into a numerical optimization problem by defining an ellipsoid structure that enclose the target lesion. The ellipsoid structure is defined by its center, major and minor axis and orientation, (X, Y, a, b, θ). As the result, the organism(candidate solution) in ES can be represented as $(X, Y, a, b, \theta; \delta_1, \delta_2, \ldots, \delta_5; \gamma_1, \gamma_2, \ldots, \gamma_{10})$, where the object variables defines the ellipsoid structure:

(X, Y): the center of an ellipse;
(a, b): the minor and major axis radius of an ellipse;
θ: the rotation angle of an ellipse.

The control variables, $\overrightarrow{\delta}$ and $\overrightarrow{\gamma}$, have the standard interpre-

tation of defining the hyper-ellipsoid that proscribes the mutation operator.

The objective function returns the fitness of an ES organism. We use a region-based objective function as defined in Equation (1) according to the skin lesion property: the lesion and the background skin are different statistically. Such an objective function favors an ellipse dividing the image into two homogenous areas with minimum variation in both regions.

$$F(X,Y,a,b,\theta) = \int_{\omega} |I(x,y)-c_1|^2 dxdy + \int_{\Omega\setminus\bar{\omega}} |I(x,y)-c_2|^2 dxdy \quad (1)$$

where $I(x,y)$ is the intensity value of the coordinate (x,y); ω is the area enclosed by the ellipse defined by (X,Y,a,b,θ); Ω is the area of the pixels whose intensity value is not zero; c_1 and c_2 represent the average intensity value of the pixels inside and outside ω respectively.

For our experiment, we adopted CMAES [10] to perform ES optimization. Our experiment results show that ES guarantees similar performance for different initial center points and ellipsoid size. This demonstrates the robustness of ES algorithm to the initialization. To further improve computational efficiency, we put the constraints in CMAES to specify a search area: $5 < a,b < 120$, and $1/9 \times N < X,Y < 8/9 \times N$, where N is the size of the image. These constraints are designed based on the fact that all lesion areas should not exceed the scope of the image, and always occupy significant amount of areas near the center of the image if not all. The fitness function of each individual is the same as Eq.(1).

The segmentation procedure can be illustrated by Fig.2. The first step is preprocessing which is similar to [1, 2]. The

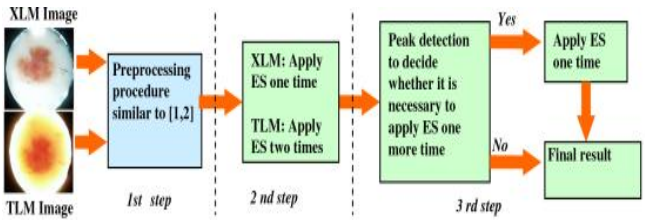

Fig. 2. Framework for the segmentation procedure

second step is applying ES to minimize Eq.(1) on the preprocessed image. For XLM image, we apply ES one time and then output the segmented region inside the ellipse for the third step. For TLM image, we apply ES two times. The result from the second step is already good enough for most images. But for images whose lesions are very small, we apply ES one more time to get a satisfactory result. In the third step, we first detect the peaks of the smoothed histogram of the image output by step two. Because lesion area always has lower intensity, the first peak of the histogram will represent the lesion area. If the first peak is lower than the maximum peak,it means that lesion area is not the dominant feature inside the ellipse. In this case, we apply ES one more time. Otherwise, we just output the result of step two.

The computational time complexity of the pre-processing steps is $O(N^2)$, where N^2 is the image size. Theoretical analysis of the computational time complexity of different evolutionary algorithms, including evolutionary strategy, is still ongoing research [11]. The time consumed for ES varied depends on how many generations it takes to converge. If it takes M generations to converge, and $\mu+\lambda$ organism for each generation, the computational complexity is $O(M \times (\mu+\lambda))$. As long as M is much less than N^2, the ES-based method is more efficient than pixel-based or region-based methods. We used Pentium(R) 4 CPU 2.26GHz, 1.50 GB of RAM to run the algorithm. The execution time for "easy" images which output results from step 2 is about 15 minutes and the execution time for "difficult" images which output results from step 3 is about 20 minutes.

5. EXPERIMENTS AND RESULTS

We applied the ES-based algorithm to the same skin lesion image sets used in [1, 2]. Among the 68 pairs of XLM and TLM images, only 51 XLM images and 60 TLM images were manually segmented by dermatologist since other images do not show pigmentation [2]. These are treated as true values, and we validate our ES based segmentation algorithm by comparing our results with the manually segmented results. Table(1) shows the error rate comparison between the ES-based method and four segmentation methods previously developed [1, 2] for "easy" images. The error rate is defined based on the overlapping area between the computer generated curve and the manual segmentation result from an expert[1, 2]. For most of the 110 skin lesion images, where the edge is well defined and the artifact and noise level are low, ES-based method achieves compatible result as the four segmentation methods in [1]. (see Table(1)).

Alg.	I	II	III	IV	ES
XLM	N/A	17.96%	15.58%	13.57%	14.82%
TLM	23.09%	N/A	19.22%	16.57%	16.71%

Table 1. Average error ratio for XLM and TLM images of the four algorithms in [1] and ES. I: Sigmoid, II:PCT, III: PCT plus Sigmoid, IV:Fuzzy c-mean.

Comparing with other four segmentation methods presented in [1, 2], the experiment results show that ES-based method performs better and is more robust when applying to images with higher noise level, very small lesion, or weak edge (as shown in Fig.3). In Fig.3, each row shows the segmentation results from three methods for one image. The first column shows the results from the ES based algorithm. The middle

column shows manual segmentation by a certified dermatologist. The last column shows the segmentation results for the same images best segmentation output from the scoring system in [1, 2].

Alg.	ES	Result from [2]
TLM#38	12.12%	22.35%
TLM#72	6.48%	42.38%
XLM#13	17.6%	45.05%

Table 2. Error ratio for images in Fig.3.

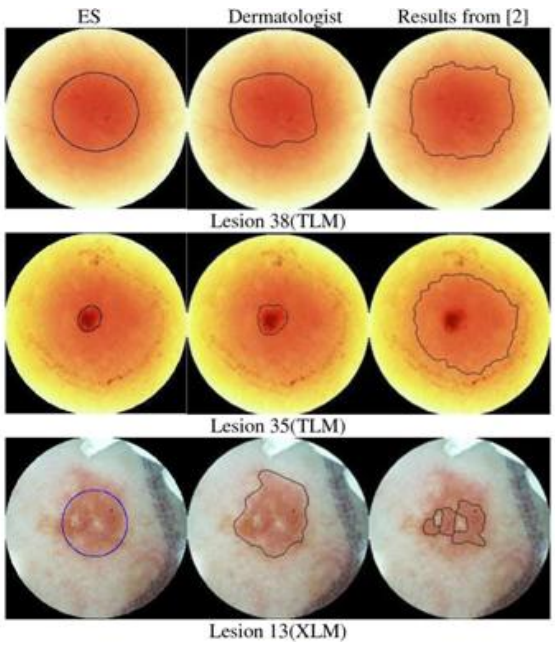

Fig. 3. Comparing ES, dermatologist , and [2]'s results.

Results shown in Fig.3 demonstrate better segmentation results from ES-based algorithm for Lesion 38(TLM), 35(TLM), and 13(XLM). The edge of the lesion 38 is very vague; the lesion 30 is very small; and lesion 13 is lesion clusters with holes in between. Based on the objective function in Eq.(1), ES-based method is not affected by the edge strength (lesion 38), size of the lesion(lesion 35), and holes in the middle of the lesion (lesion 13)(see Table(2)). In specific, under close examination, for lesion 35, the proposed ES-based methods can identify smaller region than the algorithms in [1, 2]. This is because we incorporated more detailed region information using the three-step procedure as described in Section 4.

6. DISCUSSION AND CONCLUSION

In this paper, we present an ES-based segmentation method developed for automatic skin lesion images segmentation. Experiments were done for 60 TLM and 51 XLM images. Results demonstrate that ES-based algorithm is more robust and cannot be easily affected by artifacts. It does not require any user input parameters, such as threshold, and its performance does not depends on initial values. Our ES-based segmentation method is flexible to adopt other fitness functions. In the future, we plan to incorporate some edge and texture information to further improve the segmentation results and computational efficiency.

7. REFERENCES

[1] M. Doshi G. Zouridakis and N. Mullani, "Early diagnosis of skin cancer based on segmentation and measurement of vascularization and pigmentation in nevoscope images," in *Proceedings of the 26th Annual International Conference of the IEEE EMBS*. IEEE, 2004, pp. 1596–1593.

[2] M. Doshi, "Automated segmentation of skin cancer images," M.S. thesis, University of Houston, May 2004.

[3] Anonymous authors, "Cancer facts & figures 2002," American Cancer Society, Atlanta, Georgia, 2002.

[4] Xiaojing Yuan, Jian Zhang, Xiaohui Yuan, and Bill P. Buckles, "Multi-scale feature identification using evolution strategies," *Image Vision Comput*, vol. 23, no. 6, pp. 555–563, 2005.

[5] Xiaohui Yuan, Jian Zhang, and Bill P. Buckles, "Evolution strategies based image registration via feature matching.," *Information Fusion*, vol. 5, no. 4, pp. 269–282, 2004.

[6] S. M. Bhandarkar and H. Zhang, "Image segmentation using evolutionary computation," *IEEE Trans. on Evolutionary Computation*, vol. 3, no. 1, pp. 1–21, 1999.

[7] Cor J. Veenman, Marcel J. T. Reinders, and Eric Backer, "A cellular coevolutionary algorithm for image segmentation," *IEEE Transactions on Image Processing*, vol. 12, no. 3, pp. 304–316, 2003.

[8] Payel Ghosh and Melanie Mitchell, "Segmentation of medical images using a genetic algorithm," in *GECCO*, Mike Cattolico, Ed. 2006, pp. 1171–1178, ACM.

[9] Chris McIntosh and Ghassan Hamarneh, "Genetic algorithm driven statistically deformed models for medical image segmentation," in *GECCO*, Mike Cattolico, Ed. 2006, ACM.

[10] *http://www.bionik.tu-berlin.de/user/niko/cmaes.m*, Oct. 2006.

[11] X. Yao, J.E. Rowe, and J. He, "Computational complexity analysis of evolutionary algorithms," *http://gow.epsrc.ac.uk/ViewGrant.aspx?GrantRef=EP/C520696/1*.