A General Approach to Liver Lesion Segmentation in CT Images

Li Cao[a,b], Jayaram K. Udupa[*a], Dewey Odhner[a], Lidong Huang[a,c], Yubing Tong[a], Drew A. Torigian[a]
[a]Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, PA; [b]Science and Technology on Multi-spectral Information Processing Laboratory, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China; [c]College of Information and Communication Engineering, Beihang University, Beijing 100083, China

ABSTRACT

Lesion segmentation has remained a challenge in different body regions. Generalizability is lacking in published methods as variability in results is common, even for a given organ and modality, such that it becomes difficult to establish standardized methods of disease quantification and reporting. This paper makes an attempt at a generalizable method based on classifying lesions along with their background into groups using clinically used visual attributes. Using an Iterative Relative Fuzzy Connectedness (IRFC) delineation engine, the ideas are implemented for the task of liver lesion segmentation in computed tomography (CT) images. For lesion groups with the same background properties, a few subjects are chosen as the training set to obtain the optimal IRFC parameters for the background tissue components. For lesion groups with similar foreground properties, optimal foreground parameters for IRFC are set as the median intensity value of the training lesion subset. To segment liver lesions belonging to a certain group, the devised method requires manual loading of the corresponding parameters, and correct setting of the foreground and background seeds. The segmentation is then completed in seconds. Segmentation accuracy and repeatability with respect to seed specification are evaluated. Accuracy is assessed by the assignment of a delineation quality score (DQS) to each case. Inter-operator repeatability is assessed by the difference between segmentations carried out independently by two operators. Experiments on 80 liver lesion cases show that the proposed method achieves a mean DQS score of 4.03 and inter-operator repeatability of 92.3%.

Keywords: computed tomography (CT), liver lesions, segmentation, fuzzy connectedness

1. INTRODUCTION

As liver cancer is among five cancers causing the most deaths worldwide, and metastatic lesions are common in the liver, a precise analysis of liver lesions is needed to exactly evaluate lesional properties and burden and to guide possible applicable therapies[1]. However, to segment liver lesions in computed tomography (CT) images is a challenging task due to a number of factors, including the irregularity of lesion shape, the inhomogeneity of lesion tissues, and the similarities between the image characteristics of lesions with a variable number of surrounding normal tissues. A common approach is for a trained operator to manually segment the lesions, which is time-consuming and subjective due to the differences of skill, expertise, and experience among operators[2]. Since the diversity and complexity of liver lesions make it inadequate to perform lesion segmentation automatically[3], the semi-automatic approach is more feasible and suitable to accomplish the task.

In the past ten years, many semi-automatic methods[1-6] have been developed for segmentation of various types of liver lesions from CT scans by applying some basic image segmentation algorithms, such as thresholding, region growing, level set, and watershed. By combining adaptive thresholding with model-based morphological processing, Moltz *et al.*[4] present a semi-automatic algorithm for segmentation of liver metastases in CT scans. By using the watershed algorithm first for liver contour extraction, and region growing and level-set based surface smoothing methods for segmentation refinement, a semi-automatic technique based on a minimum cross-entropy multi-thresholding algorithm[1] is developed for liver tumor segmentation. Based on 2D region growing with knowledge-based constraints, Wong *et al.*[2] propose a semi-automatic method to segment liver tumors from constituent 2D slices obtained from 3D CT images. Based on a marker-controlled watershed transformation, a semi-automatic method developed by Yan *et al.*[5] can accurately segment 3D liver metastases in volumetric CT images. In addition, for the segmentation of complicated liver lesions, a local

Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling,
edited by Robert J. Webster, III, Ziv R. Yaniv, Proc. of SPIE Vol. 9786, 978623
 · doi: 10.1117/12.2217778

multiphase C-V level set algorithm integrated with multilevel Otsu threshold selection method is proposed[6]. Although these methods achieve good segmentation performance, they are effective only for the specific types of liver lesions on which they have been validated. Most existing methods also require a region of interest (ROI) to be specified around the lesion or prior information about the boundary of the liver.

The goal of this paper is to develop a general strategy to segment all (or most) types of liver lesions. The premise of the approach is that by classifying the lesions into groups based on their clinically employed visual attributes and the attributes of their background, and by determining the group with which a lesion is identified, they can be segmented by using an effective delineation engine based on parameters preset for the group. This study is an instance from a larger project to develop a similar general strategy to segment lesions body-wide based on standardized grouping of lesions employing clinically used visual attributes.

2. METHOD

The group-based method for lesion segmentation consists of two main steps: classifying lesions into groups and segmenting the grouped lesions using an Iterative Relative Fuzzy Connectedness (IRFC) algorithm as the delineation engine[7]. The segmentation performance is assessed by accuracy and repeatability with respect to seed specification.

2.1 Lesion classification

An overview of the lesion classification approach is shown in Figure 1. Consider N subjects whose livers contain M lesions, labeled L_1, L_2, …, L_M. Generally, the number of lesions M is much larger than N. The clinically used visual attributes commonly considered for lesion description can be categorized into two classes – **C1**: those that are geometric (spherical, ovoid, polygonal, etc.), morphological (extensive, small, etc.), spatial/geographical (central, peripheral, etc.), and topological (solitary, multiple, etc.), and **C2**: those that are image appearance related (focal, diffuse, homogeneous, heterogeneous, gas-filled, fatty, cystic, etc.). Since many lesions have the same attributes (within the same organ or anywhere in the body), M lesions can be classified and merged into T groups, where T is much smaller than M. Since properties in **C1** matter much less than those from **C2** from the viewpoint of lesion delineation, we use the lesion appearance property, denoted *lap*, as the primary basis for grouping. In addition, considering the fact that background objects can be quite diverse, we take the lesion occupation property, denoted *lop*, as the secondary consideration for classification. To keep the grouping concise, we consider only two situations. The simple case of the lesions being situated in the organ interior proper is represented by *lop* = 0, which means that the possible background co-objects or tissues for the lesion are just the normal tissues in that organ. For lesions in the vicinity of the organ boundary, there are other background co-objects or tissues adjacent to the organ, such as bone, muscle, and fat, which play an important role in deciding the segmentability of the lesions. For this case we assign *lop* = 1.

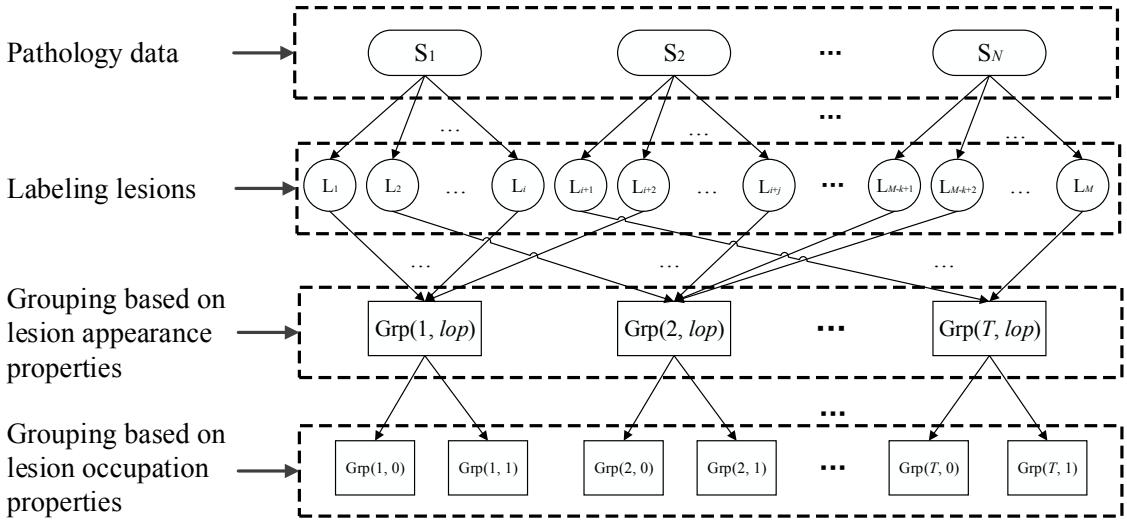

Figure 1. A schematic representation of the lesion classification approach.

2.2 Lesion segmentation

For actual delineation, the IRFC engine is used, which constitutes a top-of-the-line algorithm in the fuzzy connectedness (FC) family. The framework we describe here may use any other effective delineation engine properly tuned to the framework in place of IRFC. IRFC relies on correctly identifying all relevant background tissue components that are distinct and that constitute topological neighbors of the lesion under consideration. These are the only background tissues that matter for delineation. At the core of IRFC (and any FC) method is what is referred to as an *affinity function*[8-10], which can be described completely by the mean (μ) and standard deviation (σ) of each background tissue component and the lesion itself. Since the number of distinct (normal) tissue types and their appearance patterns in CT imagery is small, lesions which have similar intensity attributes can be further grouped based on their background configuration. This strategy allows categorizing lesions based on their own image attributes into groups and subdivided into finer groups based on background tissues.

For lesions in any group Grp(*lap*, *lop*), *lap* =1, 2, 3, …, *T*, *lop*$\in$ {0, 1}, which have similar background properties, the background parameters (μ_{bi}, σ_{bi}) for each background tissue type *i* are estimated from a sufficient number of normal tissue samples taken from regions of tissue *i*. The background configuration for lesions in the organ interior is related to only the tissue type appearing in that organ. On the other hand, when the lesions are located near or at the organ boundary, the number of background tissue types increases due to other adjacent normal tissues. For lesions in any group Grp(*lap*, *lop*), *lap*$\in$ {1, 2, 3, …, *T*}, *lop* = 0, 1, which have similar lesion properties, the foreground parameters (μ_f, σ_f) are optimized by taking the median values of μ and σ in the set {(μ, σ): (μ, σ) are the mean and standard deviation of image intensities of lesions from a training set of lesions in the group}. Averages taken over all lesions for μ and σ are not as representative of the lesion characteristics as the median values. Generally, the number of lesion types *T* and the within-class distances between types are decided based on prior learning from a sufficiently large number of images containing lesion pathology.

For delineating a lesion in a given test image, first the correct group is identified. This is at present done manually by an operator who has been trained on the above lesion categorization process. The rest of the interactive operations required for completing the 3D delineation of the lesion are loading the foreground and background parameters for that identified group, and specifying at least one seed voxel in each background tissue region and the foreground region of the lesion. Delineation is then completed within a couple of seconds.

2.3 Segmentation evaluation

The above framework has been implemented in the CAVASS software system[11]. We performed two types of evaluations, one for testing the repeatability of delineations and another for testing their accuracy. The theoretical robustness property of IRFC (and all FC methods) to seeds guarantees the same delineation result if different seeds are placed in the "same" tissue regions and lesion. Since the interpretation of "same" regions can have inter-operator variations, we tested this inter-operator component by having two operators specifying seeds. If *A* and *B* represent binary delineations in two trials, repeatability or precision *r* is defined as $r = |A \cap B| / |A \cup B|$ where $|X|$ denotes volume of *X*.

Assessment of accuracy of the delineation of lesions is very challenging because of the inherent difficulty associated with establishing reliable surrogates of the ground truth idea that resides only in the collective qualitative wisdom of experts. Although several ideas have been published on this issue, no common method has evolved. Therefore, we took the approach of scoring the quality of the delineations on a 1-5 *delineation quality scale*, DQS. The higher the DQS score, the better is the segmentation result. A board certified radiologist (co-author Torigian) scored all tested lesions. To judge the quality of the delineations, each segmented lesion was visualized as a colored overlay over the CT image slice display with a mouse click "on" "off" option for helping to examine the segmentation as well as the underlying intensity pattern. A fixed preset gray-level mapping corresponding to soft tissues was used for all image displays in this visual quality scoring experiment.

3. RESULTS

This retrospective study was conducted following approval from the Institutional Review Board at the Hospital of the University of Pennsylvania along with a Health Insurance Portability and Accountability Act waiver. The utilized

pathological abdominal CT data sets consist of 40 subjects (age: 24-83 years), which include 22 male and 18 female subjects. The voxel size in these contrast enhanced images was about 0.83 x 0.83 x 5.00 mm^3.

Since the intensities of the normal abdominal tissues vary considerably in different subjects, we first classify the background objects into three intensity levels. (Based on our observation, this variation is much less in other body regions, especially the thorax, and hence the number of groups can be much smaller in those regions.) Due to the large range of intensities, the liver lesions from each intensity level are aggregated into 3 or 4 different types. Subsequently, we chose 5 lesions and 5 subjects as the training subset for estimating the affinity parameters for each lesion type and background objects in that intensity range. Note that the intensities of the liver vessels and the renal cortices are set as the same value due to their high similarity. Based on the above analysis, we arrived at 22 liver lesion groups in total, as shown in Table 1.

Table 1. Lesion and background types and the corresponding parameters expressed in Hounsfield Units.

Background intensity level 1	Lesion properties	(μ_{f1}, σ_{f1})	(-9, 17)
		(μ_{f2}, σ_{f2})	(11, 17)
		(μ_{f3}, σ_{f3})	(36, 16)
	Background properties	$\{(\mu_{b0i}, \sigma_{b0i})\}$	Liver (66, 17); liver vessels (122, 20).
		$\{(\mu_{b1i}, \sigma_{b1i})\}$	Bones (440, 188); peripheral tissue of kidneys (122, 20); liver (66, 17); muscle (38, 19); fat (-108, 16); gallbladder (9, 17); lung and air (-811, 109).
Background intensity level 2	Lesion properties	(μ_{f1}, σ_{f1})	(-3, 18)
		(μ_{f2}, σ_{f2})	(28, 17)
		(μ_{f3}, σ_{f3})	(47, 18)
		(μ_{f4}, σ_{f4})	(67, 11)
	Background properties	$\{(\mu_{b0i}, \sigma_{b0i})\}$	Liver (105, 19); liver vessels (179, 24).
		$\{(\mu_{b1i}, \sigma_{b1i})\}$	Bones (483, 195); peripheral tissue of kidneys (179, 24); liver (105, 19); muscle (48, 20); fat (-111, 17); gallbladder (9, 18); lung and air (-810, 109).
Background intensity level 3	Lesion properties	(μ_{f1}, σ_{f1})	(19, 22)
		(μ_{f2}, σ_{f2})	(52, 25)
		(μ_{f3}, σ_{f3})	(86, 20)
		(μ_{f4}, σ_{f4})	(128, 14)
	Background properties	$\{(\mu_{b0i}, \sigma_{b0i})\}$	Liver (153, 14); liver vessels (241, 19).
		$\{(\mu_{b1i}, \sigma_{b1i})\}$	Bones (503, 222); peripheral tissue of kidneys (241, 19); liver (153, 14); muscle (60, 16); fat (-97, 14); gallbladder (28, 12); lung and air (-795, 124).

After creating the above lesion group library, each liver lesion to be segmented is assigned to the most suitable lesion group. Figure 2 shows 6 common liver lesions and their corresponding delineation results. Note that for the case in column 3, we consider only the large homogeneous lesion. Their grouping information is listed in Table 2.

Clearly the lesions in columns (1) and (5) are segmented perfectly by the group-based IRFC method, their DQS score was given as 5. The lesions in columns (2) and (3) also can be segmented with good accuracy, despite the missing of a small part of lesion region with relatively large deviations. Their DQS score was set to 4. The segmentation of the lesions in columns (4) and (6) are not so good due to a slightly larger difference of lesion properties between the actual value and the most suitable parameters in lesion group library. In these cases, the score was 3.

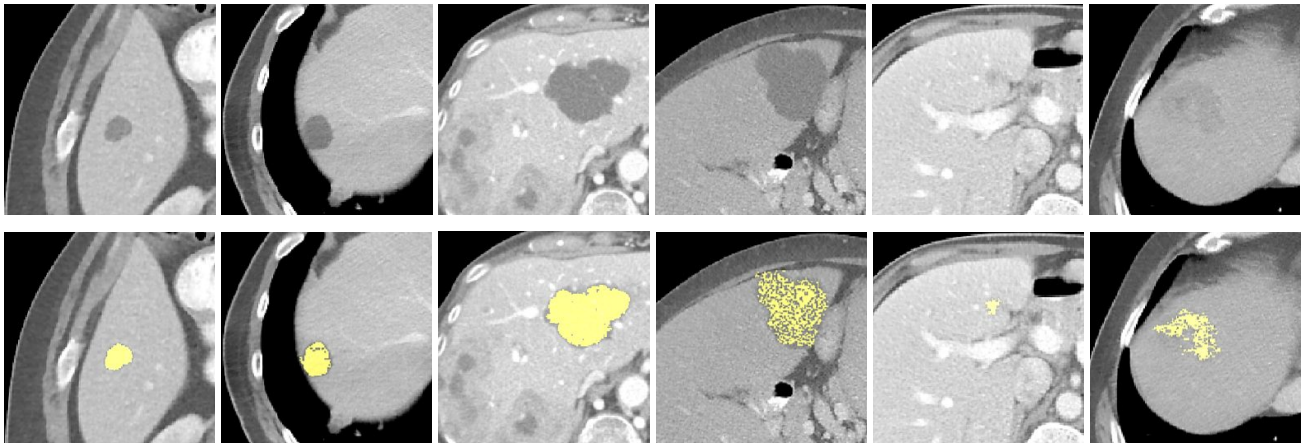

Figure 2. Examples of lesion delineations where the result is overlaid on the original CT slice for six lesions. Columns: 1, 2, 3, 4 – focal cysts, peripheral in columns 1, 2, and 4 and central in 3. Columns 5, 6 – focal, peripheral, heterogeneous, and malignant lesions.

Table 2. Group information of lesions in Figure 2.

Image	Background intensity level	Lesion group
Column 1	2	Grp(2, 0)
Column 2	2	Grp(1, 1)
Column 3	3	Grp(1, 0)
Column 4	1	Grp(1, 1)
Column 5	3	Grp(3, 1)
Column 6	2	Grp(3, 1)

Table 3. Accuracy assessment: DQS score distribution.

DQS score	Number of lesions	%
5	36	45.00
4	21	26.25
3	14	17.50
2	7	8.75
1	2	2.50

Table 4. Assessment of repeatability of delineations.

Lesion	$\|A \cap B\|$	$\|A \cup B\|$	Repeatability r	Lesion	$\|A \cap B\|$	$\|A \cup B\|$	Repeatability r
L_1	45	51	0.88	L_6	2707	2708	1.00
L_2	71	113	0.63	L_7	74	75	0.99
L_3	17	17	1.00	L_8	218	298	0.73
L_4	72	72	1.00	L_9	92	92	1.00
L_5	1667	1667	1.00	L_{10}	77	77	1.00

We randomly selected 80 liver lesions from 20 subjects to assess the accuracy via DQS score following the above scoring principle. As shown in Table 3, 88.75% of the lesions are segmented by the proposed method at or above a score of 3, and 71.25% of the lesions can be segmented at or above a score of 4. The method achieved a mean DQS score of 4.03 over the 80 lesions.

The assessment of repeatability is performed on 10 liver lesion segmentation cases. Table 4 shows that the proposed method has high repeatability with an average value of $r = 0.92$.

4. CONCLUSIONS

Lesion segmentation has remained a challenge in different body regions. Generalizability is lacking in published methods as variability in results is common, even for a given organ, such that it becomes difficult to establish standardized methods of disease quantification and reporting. This paper makes an attempt at a generalizable method based on classifying lesions with their intensity and location properties, which are usually expressed by visual qualitative attributes in CT images as well as the characteristics of background tissue regions in their neighborhood. To our knowledge, such an effort has not been made to date. The idea is demonstrated and implemented on liver lesions in CT images. Based on preset parameters derived for each liver lesion group from a few training samples, the IRFC-based delineation method can segment most of the liver lesions effectively wherein the operator identifies seeds as per group protocol to identify the lesion and background tissue regions. Our experimental evaluation shows that the proposed method can be used to segment liver lesions in this standardized manner with a mean DQS score of 4.03 and inter-operator repeatability of 92.3%.

Though there are still some complicated cases which cannot be handled effectively, such as multiple liver lesions with high density in the same CT slice, and highly heterogeneous liver lesions, other lesion attributes may be usefully added for auxiliary segmentation in such cases. In addition, refinement of the parameters of the lesion group library and automation of the current manual group selection operation may reduce false negative regions (under-segmentation problem) and improve the implementation efficiency. We are also extending this methodology body-wide to other organs.

ACKNOWLEDGEMENTS

The training of Ms. Li Cao in the Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, for the duration of one year was supported by China Scholarship Council.

REFERENCES

[1] Choudhary, A., Moretto, N., Ferrarese, F. P., and Zamboni, G. A., "An entropy based multi-thresholding method for semi-automatic segmentation of liver tumors," Proc. MICCAI Workshop 3-D Segmentat. Clin. : A Grand Challenge II (2008).

[2] Wong, D., Liu, J., Fengshou, Y., Tian, Q., Xiong, W., Zhou, J., Qi, Y., Han, T., Venkatesh, S., and Wang, S., "A semi-automated method for liver tumor segmentation based on 2D region growing with knowledge-based constraints," Proc. MICCAI Workshop 3-D Segmentat. Clin. : A Grand Challenge II-Liver Tumor Segmentat. (2008).

[3] Stawiaski, J., Decenciere, E., and Bidault, F., "Interactive liver tumor segmentation using graph-cuts and watershed," Proc. MICCAI Workshop 3-D Segmentat. Clin. : A Grand Challenge II (2008).

[4] Moltz, J., Bornemann, L., Dicken, V., and Peitgen, H., "Segmentation of liver metastases in CT scans by adaptive thresholding and morphological processing," Proc. MICCAI Workshop 3-D Segmentat. Clin. : A Grand Challenge II (2008).

[5] Yan, J., Schwartz, L. H., and Zhao, B., "Semiautomatic segmentation of liver metastases on volumetric CT images," Med. Phys. Papers 42(11), 6283-6293 (2015).
[6] Luo, Z., and Ou, S., "Segmentation of complicated liver lesion based on local multiphase level set," Proc. 3rd Int. Conf. Bioinformatics and Biomedical Engineering, 1-4 (2009).
[7] Ciesielski, K. C., Udupa, J. K., Falcão, A. X., and Miranda, P. A. V., "Fuzzy connectedness image segmentation in graph cut formulation: a linear time algorithm and a comparative analysis," J. Math. Imaging Vis. Papers 44(3), 375-398 (2012).
[8] Udupa, J. K., and Samarasekera, S., "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," Graphical Models and Image Processing, Papers 58(3), 246-261 (1996).
[9] Ciesielski, K. C., and Udupa, J. K., "Affinity functions in fuzzy connectedness based image segmentation II: Defining and recognizing truly novel affinities," Comput. Vis. Image Understand. Papers 114 (1), 155-166 (2010).
[10] Udupa, J. K., Odhner, D., Zhao, L., Tong, Y., Matsumoto, M. M. S., Ciesielski, K. C., Falcao, A. X., Vaideeswaran, P., Ciesielski, V., Saboury, B., Mohammadianrasanani, S., Sin, S., Arens, R. and Torigian, D. A., "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," Medical Image Analysis, Papers 18, 752-771 (2014).
[11] Grevera, G., Udupa, J.K., Odhner, D., Zhuge, Y., Souza, A., Iwanaga, T. and Mishra, S., "CAVASS: A computer-assisted visualization and analysis software system," J. Digital Imaging, Papers 20, 101-118 (2007).