Original Article

Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma

Qingtao Qiu[1,2*], Jinghao Duan[2*], Guanzhong Gong[2], Yukun Lu[1,2], Dengwang Li[1], Jie Lu[2], Yong Yin[2]

[1]Shandong Province Key Laboratory of Medical Physics and Image Processing Technology, Institute of Biomedical Sciences, School of Physics and Electronics, Shandong Normal University, Ji'nan 250358, China; [2]Department of Radiation Oncology, Shandong Cancer Hospital Affiliated to Shandong University, Ji'nan 250117, China

Contributions: (I) Conception and design: J Duan, Y Yin; (II) Administrative support: Y Yin; (III) Provision of study materials or patients: Q Qiu, J Duan, Y Yin; (IV) Collection and assembly of data: Q Qiu; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*These authors contributed equally to this work.

Correspondence to: Professor Yong Yin, PhD. Department of Radiation Oncology, Shandong Cancer Hospital Affiliated to Shandong University, Ji'nan 250117, China. Email: yinyongsd@126.com.

Background: The reproducibility of radiomic features is a critical challenge facing radiomic models of tumor prediction or prognosis. The aim of this study is to evaluate the reproducibility of radiomic features with the GrowCut and GraphCut semi-automatic tumor segmentation methods in hepatocellular carcinoma (HCC) CT images.

Methods: Computed tomography (CT) data sets (arterial enhanced phase) of 15 patients with HCC were randomly selected in this study. To acquire the gross tumor volume (GTV), semi-automatic segmentation with the GrowCut and GraphCut methods was implemented in 3D Slicer software by two independent observers. Meanwhile, manual delineation of the GTV was implemented by five abdomen radiation oncologists in this study. We divided the sample into three groups: the GrowCut group, the GraphCut group and manual group. Radiomic features (including tumor intensity histogram-based features, textural features and shape-based features) were extracted using the Imaging Biomarker Explorer (IBEX) software. The intra-class correlation coefficient (ICC) was applied to assess the reproducibility of all radiomic features.

Results: The radiomic features in the GrowCut group (ICC =0.87±0.19) showed higher reproducibility compared with the radiomic features in the GraphCut group (ICC =0.82±0.24, P<0.001) and the manual delineations group (ICC =0.80±0.21, P<0.001), respectively. For intensity histograms, semiautomatic segmentation tools can yield more reproducible features. GLCM features were more robust in the GrowCut group segmentations. Furthermore, no statistically significant difference in the remaining feature categories was observed between the manual method and two semiautomatic methods.

Conclusions: Our study reveals that variations exist in the reproducibility of quantitative imaging features extracted from tumor regions segmented using different methods. The 3D Slicer can serve as a better alternative to the manual delineation, and care must be taken when selecting segmentation tools to draw the tumor region.

Keywords: Hepatocellular carcinoma (HCC); radiomics; reproducibility; semiautomatic segmentation

Submitted Jun 06, 2017. Accepted for publication Sep 22, 2017.
doi: 10.21037/tcr.2017.09.47
View this article at: http://dx.doi.org/10.21037/tcr.2017.09.47

Introduction

Hepatocellular carcinoma (HCC), the sixth most common cancer worldwide, is the main type of cancer in liver parenchymal cells (1). Medical imaging techniques, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission CT (PET-CT), play important roles in oncology. Especially in radiotherapy, imaging dominates treatment planning and response monitoring (2). Several publications have shown that quantitative image features have potential applications in providing consistent, nonbiased descriptors to the tumor research.

For the past several years, as an emerging individualized precision medical technology, radiomics has applied advanced computational methodologies to transform the image data of the regions of interest into high dimensional feature data. Next, quantitative and high-throughput analysis of feature data is completed to probe tumor phenotype (3-6). Radiomics utilizes noninvasive imaging to provide more comprehensive information about the entire tumor and can be used in diagnosis, prognosis and prediction (6,7). For patients with colorectal liver metastases, relative differences of CT textural features occurring after treatment were better than RECIST in predicting and assessing the pathological response to chemotherapy (8). Another, more recent HCC radiomics study showed that CT-based radiomics signature was a powerful predictor for preoperative estimation of early recurrence (9).

The radiomics features must be reproducibility, non-redundancy and informative (10). Reproducibility, the most basic and essential problem in radiomics, refers to measurements of radiomic features performed using different equipment, different methods or observers, or at different sites and times (10,11). The reproducibility may be influenced by many factors, such as imaging devices (12), repeat CT scans (13-15), tumor volume definition (16-18) and feature extraction (19,20). For acquiring accurate results, the producible features should be selected in building prognostic or predictive models.

Tumor segmentation is crucial for subsequent quantitative imaging extraction. Although manual delineation by experts is a common method considered as a 'gold standard', it is time-consuming and suffers from inter-observer variability. Recent studies have shown that 3D Slicer semiautomatic segmentation results were almost consistent with the manual contour by expert (21-23). 3D Slicer (23) is a free and open-source software package for medical image analysis in which many extensions are available for tumor segmentation on CT images. Since liver tumors have indistinct borders, there is high variability in radiologists' determination of tumor outlines, leading to increased variation in features extraction (17). However, to the best of our knowledge, few studies have investigated the stability of radiomic features extracted from tumor regions defined by different semiautomatic methods.

In this study, we evaluate the reproducibility of quantitative imaging features derived from tumor volume segmented using GraphCut and GrowCut interactive methods in 3D Slicer and to determine robustness of feature categories to propel clinical radiomics research of HCC patients.

Methods

CT imaging data of HCC patients

The CT imaging data set of 15 patients who have been diagnosed with primary HCC between December 2015 and May 2016 were randomly collected. All patients received abdominal enhanced CT scanning, and a Philips scanner (Holland, CT Lightspeed 16) was used with an imaging protocol of tube voltage 120 kV, cube current 300 mA, thickness 3 mm and in-plane resolution 0.97×0.97. Each patient has only one lesion (the volume range of tumor is 5–168 cm^3, median 19 cm^3). Arterial CT images were included in this study. This study was approved by the institutional review board (IRB) and ethics committee of Shandong Cancer Hospital Affiliated to Shandong University. The ID/number of ethics approval was 201606021.

Semiautomatic tumor segmentation

3D Slicer is an open-source, publicly available image analysis platform and was developed for segmentation, registration and three-dimensional visualization. In this study, GraphCut (24) and GrowCut (25) were implemented in 3D Slicer and tumor volume was defined twice by each of the two independent observers to determine intra-observer reproducibility. The run1 and run2 were first and second segmentations among different observers to assess inter-observer reproducibility.

GraphCut semiautomatic segmentation

The knowledge based star shape prior was incorporated in the graph cut algorithm in the 3D Slicer GraphCut

extension. Star shape prior is a generic shape prior that applies to a wide class of objects to achieve more robust segmentation. Graph cut turns the image segmentation into discrete graph optimization (min cut/max flow). First, this method builds an energy function after mapping the images to undirected weighed graphs; meanwhile, voxels in the images are treated as graph nodes. Next, the similarity of nodes in graph is calculated as the weight of connections between nodes. Finally, the min cut and energy function minimized strategy is employed to obtain the optimal segmentations.

Before GraphCut was activated in 3D Slicer, operators need to add four fiducials around the tumor after loading images. Two fiducials on the first slice and last slice where the tumor begins and ends to show to identify the start-end of the tumor and two fiducials on the middle slice where the tumor area is the largest at the diagonal corners of a rectangle, which can contain the tumor. Thereafter, 2D or 3D star shape constraints can be checked as needed.

GrowCut semiautomatic segmentation

GrowCut has better performance on accuracy and speed in tumor segmentation by using a competitive region growing algorithm (22,24). A set of initial labels needs to be given by users to mark foreground and background, and cellular automata automatically segments the remaining image using a weighted similarity score. The neighbor that results in the largest weight greater than the given voxel's strength confers its label to the given voxel. If there are two or more tumors in the image to be segmented, the corresponding class of initial labels was needed.

The GrowCut is executed as follows: first, it defines the tumor and non-tumor region with different label value; next, algorithm automatically computes a region of interest. After that step, GrowCut was activated to label iteratively all of the voxels in the ROI until all the voxels are labeled or until no voxel can change its label any more.

If not satisfy with the result, the foreground tumor region can be edited manually both in GraphCut and GrowCut.

Manual tumor delineation

Five experienced radiologists manually defined the gross tumor volume (GTV) of primary HCC in MIM software (www.mimsoftware.com) twice per radiologist using standard delineation protocol (window width: 200 HU, window level: 40 HU). The radiologists were blind to one another.

Quantitative imaging feature extraction

Seventy-one quantitative imaging features were extracted from the information contained in the voxels of the tumor region segmented by the three strategies. This process was implemented in IBEX (Imaging Biomarker Explorer, MD Anderson cancer center, USA), an open-source, easy to use radiomic software (26). These features were organized into three categories: (I) intensity histogram; (II) texture; and (III) shape. Seventeen first-order statistical features derived from the tumor intensity histogram reflect distribution of values of individual voxels without concern for spatial relationships. Thirty-eight textural features describe spatial arrangement of voxels were calculated from different parent matrices, including gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), neighbor gray-tone difference matrix (NGTDM). Sixteen shape based features provide the geometrical of tumor volume. To reduce the effect of noise on the textual features, all of the voxel intensity values within the ROI were rescaled to 8-bit images using a discrete resampling method before calculating the GLCM, GLRLM, and NGTDM features (27). In this study, GLCM features were the average of all 13 symmetric directions in 3D, GLRLM features were the average of values calculated from 2 directions in the 2D slice-by-slice, and the NGTDM was defined by neighborhood in the 3D.

Statistical analysis

In this study, intra-class correlation coefficient (ICC) as defined by McGraw and Wong (28) was employed to assess the reproducibility of radiomic features derived from tumor volume segmented by three methods. ICC is a descriptive statistic between 0 and 1, where 0 and 1 indicate null and perfect reproducibility, respectively. A variety of algorithms were provided in the literature for the ICC calculation. To assess the reproducibility of radiomic features extracted from inter-observer segmentations, we used the definition of ICC(A,1), and variance estimates were obtained from two-way mixed effect model of analysis of variance (ANOVA), given by:

$$\mathrm{ICC(A,1)} = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$$

Additionally, we used the definition of ICC(C, 1) to

assess the reproducibility of radiomic features derived from intra-observer segmentations, variance estimates were obtained from one-way analysis of variance (ANOVA), with the following form:

$$\mathrm{ICC(C,1)} = \frac{MS_R - MS_W}{MS_R + (k-1)MS_W}$$

where MS_R = mean square for rows (observations, fixed factor), MS_W = mean square for residual sources of variance, MS_E = mean square error, MS_C= mean square for columns (observers, random factor), k and n represent number of observers and number of subjects, respectively. In order to help readers to better understand the model, we provide a table of measurements of GLCM-Contrast from all segmentations as one example. It can be found in *Table S1*.

To compare the differences of feature range between manual and two semi-automatic segmentations, Z-score normalization was applied to standardize the radiomic features because different features would have various ranges. The Z-score normalization was defined as follows:

$$z = \frac{x - \mu}{\sigma}$$

where μ and σ were the mean value and standard deviation of radiomic features, respectively.

Wilcoxon rank-sum tests were used to compare the differences of ICC between manual delineation and two semiautomatic methods. P<0.05 was considered to be significant. All data were expressed as the mean ± SD. SPSS version 22.0 (SPSS, Chicago, IL, USA) was used for ICC and Wilcoxon rank-sum test computation.

Results

The ICC values of the 71 quantitative imaging features across tumors segmented by three methods was presented in *Figure 1*. Noticeably, radiomic features derived from GrowCut-based segmentation (ICC =0.87±0.19) had significantly higher ICC values compared to features extracted from GraphCut-based segmentation (ICC =0.82±0.24, P<0.001) and manual delineation (ICC=0.80±0.21, P<0.001). The statistically significant difference was observed in ICC values for features-based GraphCut and manual segmentation (P=0.036). For ICCs of manual, GraphCut and GrowCut methods, the confidence intervals were (0.608, 0.954), (0.774, 0.938) and (0.752, 0.967), respectively. Overall, 53 of the radiomic feature showed higher ICC for GrowCut, and 47 showed higher ICC values for GraphCut segmentation sets compared to the manual method (P<0.001). Also, comparing GrowCut to GraphCut, 52 features with higher ICC values were extracted from GrowCut segmentations.

For tumor intensity histogram features, no statistically significant change was observed in GraphCut segmentation sets (ICC =0.77±0.29) compared to the manual method (ICC =0.76±0.26) (P=0.332), but GrowCut (ICC =0.90±0.14) showed significantly higher reproducibility (P<0.001). For GLCM features in the textural category, GraphCut (ICC =0.89±0.11) and GrowCut (ICC =0.91±0.12) has significantly higher reproducibility than the manual method (ICC =0.77±0.23) (P=0.010, P=0.004, respectively). For GLRLM, NGTDM features in textural category and shape-based features, no statistically significant difference was observed between the manual and the two semiautomatic methods. All of the features were divided into four groups according to their ICC values: excellent (0.73≤ ICC ≤1), good (0.6≤ ICC <0.75), fair (0.4≤ ICC <0.6) and poor (ICC <0.4) reproducibility (29). The number of features in each group is presented in *Table 1*. The excellent reproducibility of radiomic features for manual, GraphCut, and GrowCut segmentations were 73% [52], 77% [55], and 81% [58] of total, respectively. These features can be found in *Figure S1*.

To evaluate the effect on robustness with multiple algorithmic initializations, we analyzed the ICC of features extracted from inter-observer and intra-observer segmentations. In *Figure 2A*, we observed higher ICC values in GrowCut inter-observer segmentation groups (average ICC =0.87±0.18). In *Figure 2B*, higher ICC values were also observed in GrowCut intra-observer segmentations (average ICC =0.90±0.11). There are distinct differences of ICC in GraphCut for both inter- and intra-observer segmentation (P<0.001, P=0.008, respectively). *Figure 3* depicts the Z-score normalized feature range of all of the 18 segmentation sets (10 manual, 4 GraphCut and 4 GrowCut). Overall, the range of feature based GrowCut segmentations was smaller than that of GraphCut (P<0.001) and manual (P<0.001). GraphCut showed no significant difference compared to manual method (P=0.062). These data are available in *Figure S2*.

Discussion

Medical imaging is now routinely used and is playing an

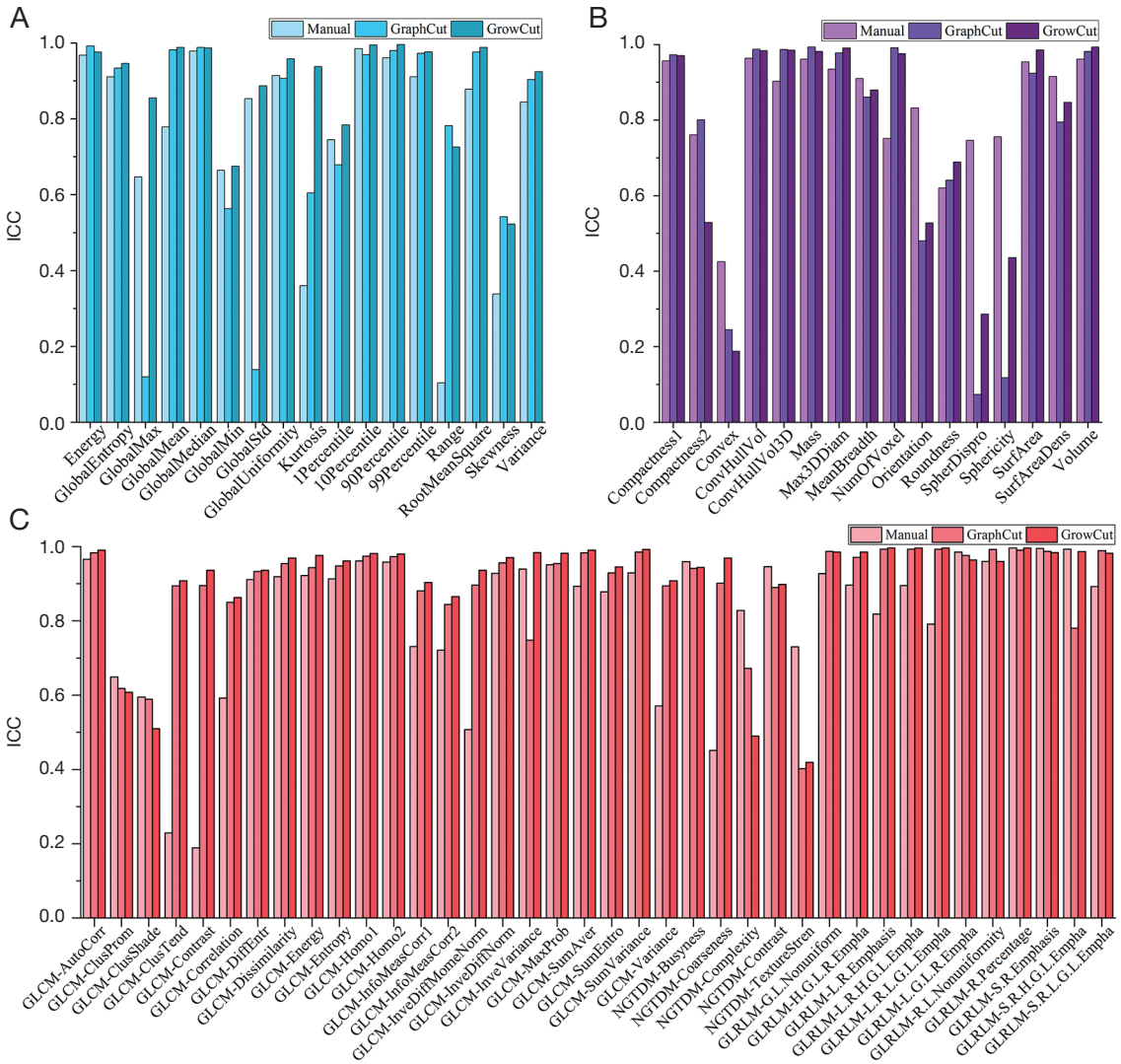


Figure 1 Feature comparison of intra-class correlation coefficient (ICC) between manual delineation and two semi-automatic segmentations. (A) Intensity histogram based features; (B) shape based features; (C) textural features.

Table 1 Number of features in four reproducibility groups across three segmentation strategies

Reproducibility groups	Manual	GraphCut	GrowCut
Poor (ICC <0.4)	7	5	2
Fair (0.4≤ ICC <0.6)	4	5	7
Good (0.6≤ ICC <0.75)	8	6	4
Excellent (0.75≤ ICC ≤1)	52	55	58

ICC, intra-class correlation coefficient.

essential role in clinical oncology. As an emerging field in precision medicine, radiomics utilizes quantitative imaging features to assess the characteristics of tumor phenotype and has potential applicability in treatment planning and monitoring. For example, the changes of radiomic features extracted from post-treatment CT images can serve as early indicators of progression to local recurrence within six months after SABR in early-stage lung cancer (30). In another study, 440 imaging features extracted from CT data

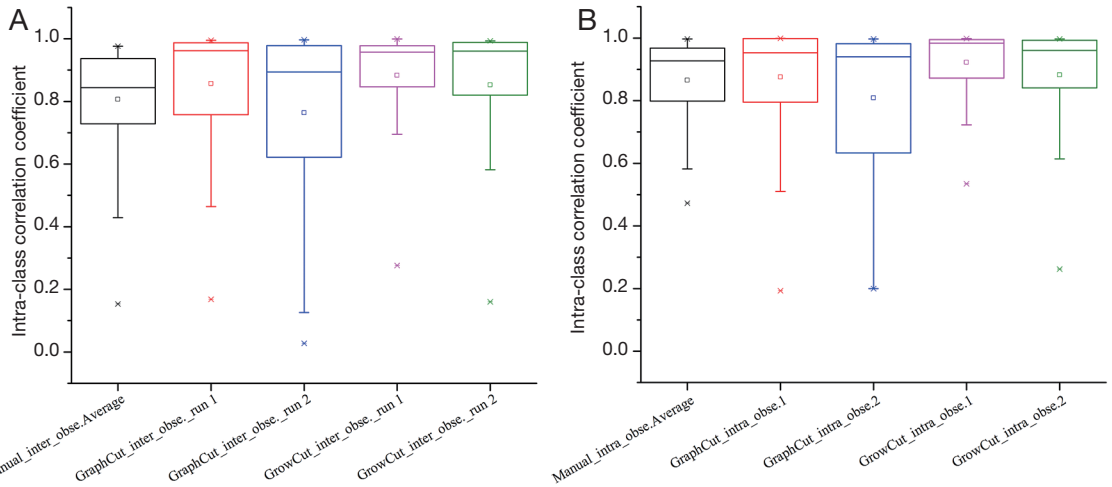


Figure 2 Box-plot comparing (A) inter- and (B) intra-observer reproducibility of radiomic features. Run1 and run2 are different segmentation sets defined by different observers.

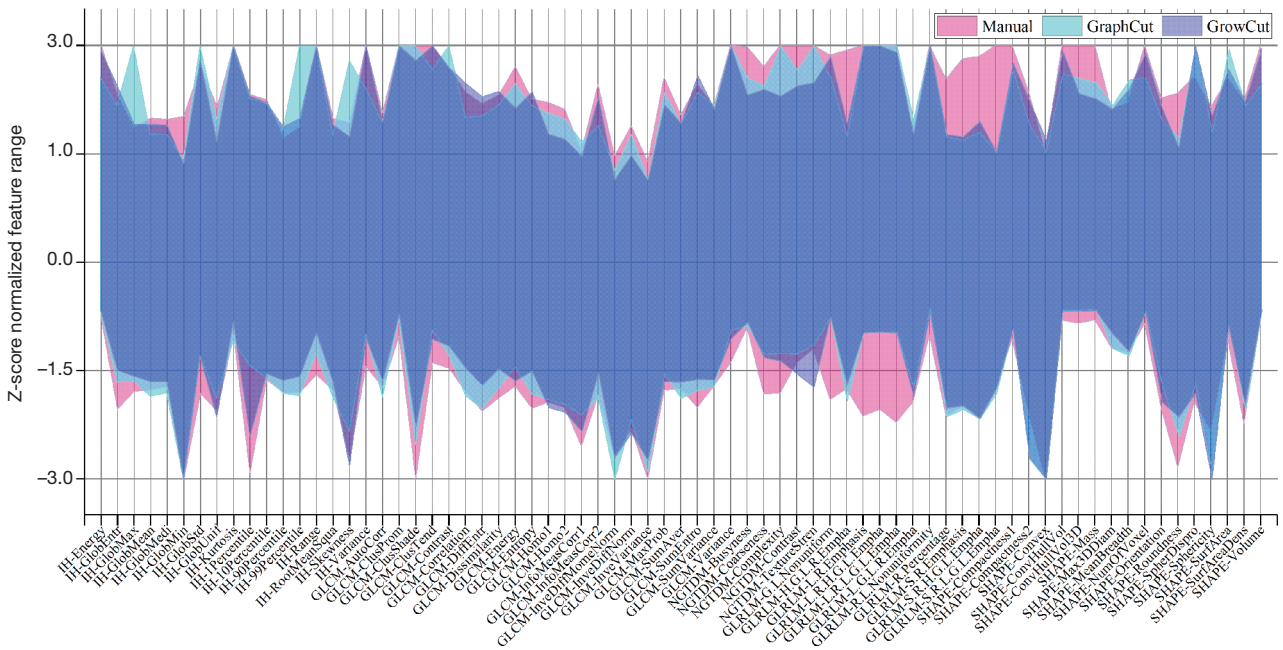


Figure 3 Comparison of normalized feature range between manual and two 3D Slicer segmentations.

of 1,019 patients with lung or head-and-neck cancer can capture intratumor heterogeneity and are associated with gene expression patterns, TNM staging and prognosis of patients (6).

Tumor segmentation is an essential step in the workflow of radiomics. Many semi-automatic and automatic segmentation algorithms have been developed for tumor delineation. Therefore, the GraphCut and GrowCut semiautomatic methods were used in liver tumor segmentation in this study. The detailed workflow of these two semiautomatic segmentation tools can be found in *Figure S3*. These methods yield more stable segmentation and need less time compared with manual delineation, as manual delineation is time-consuming and prone to higher

inter-observer variability. Parmar *et al*. concluded that the quantitative imaging feature extracted from semiautomatic tumor segmentations showed significantly higher reproducibility than manual delineations (18). However, there were few reports in the literature about the effect on the stability of radiomic features derived from tumors segmented using different semiautomatic algorithms.

In this study, 71 commonly used quantitative imaging features were selected and organized into three categories (17 tumor intensity histogram based features, 38 textural features and 16 shape based features). We analyzed the robustness of these features when they were extracted from tumor regions segmented using three methods. In all 71 radiomic features, GrowCut segmentations showed significantly higher ICC values than GraphCut segmentations and manual delineations ($P<0.001$). While GraphCut is not as significant as GrowCut, GraphCut had slightly better robustness than manual delineations ($P=0.036$), indicating that 3D Slicer tumor segmentation tools can extract more reproducible quantitative imaging features. These results can be explained by the fact that semiautomatic tumor segmentation algorithms require no more manual intervention since algorithm initialization was performed, then the tumor was segmented by an efficient algorithm. There is too much underlying uncertainty in the manual tumor delineation because of the inter-observer variability may be accumulated through slice-by-slice manual delineation.

We observed that GLCMs-based features were more robust to semiautomatic tumor segmentations, GrowCut and GraphCut has significantly higher reproducibility compared to manual delineations ($P=0.001$, $P<0.001$, respectively). Additionally, the tumor intensity histogram features were more reproducible when they were extracted from GrowCut segmentations. However, there was no significant difference in reproducibility of other feature categories. These results indicate that CT textural features derived from semiautomatic segmentations are highly reproducible for HCC patients.

To evaluate the performance of three segmentation strategies, we analyzed inter- and intra-observer reproducibility. We found that features derived from GrowCut had higher ICC values in both inter- and intra-observer, indicating that it was able to extract more reproducible features against the different algorithm initializations. We also observed that reproducible features extracted from GraphCut-based segmentations were unstable for inter-observer. The feature range that we observed was significantly smaller in GrowCut compared with other two segmentation methods.

Our findings demonstrate that stable and reproducible radiomic features can be extracted from semiautomatic tumor segmentations and that textural features extracted using these segmentation tools are more suitable. However, inter-observer initialization differences result in various segmentations due to different principles of semiautomatic methods. We can apply these methods that were evaluated in radiomics studies to yield reproducible results. In this study, tumor segmentation based on GrowCut presented great performance on the feature extraction both inter- and intra-observer.

A limitation of this study is that although clinical data for these patients are available, the small patient cohort prevented prediction/prognosis models from being devised. The conclusions drawn from this study should be applicable to predict outcomes in HCC patients. This work will be done when we have collected large prospective patient cohorts. Another limitation is that tumor size could also be an important factor related to feature reproducibility and prognostic value. These volume-dependent features will be involved in future research. Furthermore, many software packages are available for use in radiomics research (31). The increased usage of these computational resources will bridge the gaps between radiomics and clinical oncology.

Conclusions

Our study reveals that variations exist in the reproducibility of quantitative imaging features extracted from tumor region segmented using different methods. For HCC radiomics studies, tumor intensity histogram-based features and textural features were more reproducible when they were extracted from GrowCut semiautomatic segmentations. Therefore, 3D Slicer can serve as a better alternative to the manual delineation method, and care must be taken when selecting segmentation tools to draw tumor regions.

Acknowledgments

The authors would like to thank Dr. Lifei Zhang (MD Anderson Cancer Center) for her guidance.
Funding: This research was supported by the National Natural Science Foundation of China (No. 81472811 and No. 81272699), the Science and Technology Planning Project of Shandong Province (2014GSF118011) and the Shandong Natural Science Foundation (ZR2017PH071).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2017.09.47). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by Institutional Review Board (IRB) and ethics committee of Shandong Cancer Hospital Affiliated to Shandong University and the ID/number of ethics approval was 201606021. Informed consent was waived.

References

1. Singh S, Singh PP, Roberts LR, et al. Chemopreventive strategies in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol 2014;11:45-54.
2. Dutta R, Mahato RI. Recent advances in hepatocellular carcinoma therapy. Pharmacol Ther 2017;173:106-17.
3. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441-6.
4. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.
5. Yip SS, Aerts HJ. Applications and limitations of radiomics. Phys Med Biol 2016 ;61:R150-66.
6. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
7. Bashir U, Siddique MM, Mclean E, et al. Imaging Heterogeneity in Lung Cancer: Techniques, Applications, and Challenges. AJR Am J Roentgenol 2016;207:534-43.
8. Rao SX, Lambregts DM, Schnerr RS, et al. CT texture analysis in colorectal liver metastases: A better way than size and volume measurements to assess response to chemotherapy? United European Gastroenterol J 2016;4:257-63.
9. Zhou Y, He L, Huang Y, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. Abdom Radiol (NY) 2017;42:1695-704.
10. O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14:169-86.
11. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234-48.
12. Mackin D, Fave X, Zhang L, et al. Measuring Computed Tomography Scanner Variability of Radiomics Features. Invest Radiol 2015;50:757-65.
13. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and Prognosis of Quantitative Features Extracted from CT Images. Transl Oncol 2014;7:72-87.
14. Balagurunathan Y, Kumar V, Gu Y, et al. Test-retest reproducibility analysis of lung CT image features. J Digit Imaging 2014;27:805-23.
15. Hunter LA, Krafft S, Stingo F, et al. High quality machine-robust image features: identification in nonsmall cell lung cancer computed tomography images. Med Phys 2013;40:121916.
16. Leijenaar RT, Carvalho S, Velazquez ER, et al. Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol 2013;52:1391-7.
17. Echegaray S, Gevaert O, Shah R, et al. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. J Med Imaging (Bellingham) 2015;2:041011.
18. Parmar C, Rios Velazquez E, Leijenaar R, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. PLoS One 2014;9:e102107.
19. Fave X, Zhang L, Yang J, et al. Impact of image preprocessing on the volume dependence and prognostic potential of radiomics features in non-small cell lung cancer. Transl Cancer Res 2016;5:349-63.
20. Hu P, Wang J, Zhong H, et al. Reproducibility with repeat CT in radiomics study for rectal cancer. Oncotarget 2016;7:71440-6.

21. Velazquez ER, Aerts HJ, Gu Y, et al. A semiautomatic CT-based ensemble segmentation of lung tumors: Comparison with oncologists' delineations and with the surgical specimen. Radiother Oncol 2012;105:167-73.
22. Velazquez ER, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. Sci Rep 2013;3:3529.
23. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323-41.
24. Vezhnevets V, Konouchine V. GrowCut: Interactive multi-label N-D image segmentation by cellular automata. Proc Graphicon 2005;1:150-6.
25. Freedman D, Zhang T. Interactive graph cut based segmentation with shape priors. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE 2005;1:755-62.
26. Zhang L, Fried DV, Fave XJ, et al. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. Med Phys 2015;42:1341-53.
27. Fave X, Mackin D, Yang J, et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? Med Phys 2015;42:6784-97.
28. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.
29. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284-290.
30. Mattonen SA, Palma DA, Johnson C, et al. Detection of Local Cancer Recurrence After Stereotactic Ablative Radiation Therapy for Lung Cancer: Physician Performance Versus Radiomic Assessment. Int J Radiat Oncol Biol Phys 2016;94:1121-8.
31. Court LE, Fave X, Mackin D, et al. Computational resources for radiomics. Transl Cancer Res 2016;5:340-8.

Cite this article as: Qiu Q, Duan J, Gong G, Lu Y, Li D, Lu J, Yin Y. Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma. Transl Cancer Res 2017;6(5):940-948. doi: 10.21037/tcr.2017.09.47

Table S1 The measurements of GLCM-Entropy from 18 segmentations (10 Manual + 4 GraphCut + 4 GrowCut)

Patient	Manual										GraphCut				GrowCut			
	Obser.1		Obser.2		Obser.3		Obser.4		Obser.5		Obser.1		Obser.2		Obser.1		Obser.2	
	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2	Run1	Run2
1	4.052	3.755	3.863	3.985	4.221	4.879	4.427	4.816	4.721	4.771	4.608	4.292	4.357	4.979	3.957	3.769	3.889	3.761
2	3.457	3.180	3.641	3.790	3.541	3.963	4.084	4.418	4.011	4.074	4.225	3.785	3.780	4.114	3.531	3.635	3.705	3.818
3	2.810	2.779	3.036	2.768	2.668	3.064	3.135	2.845	2.966	2.870	2.953	2.913	3.058	3.083	3.004	2.850	2.861	2.620
4	3.232	3.818	3.075	3.199	3.251	3.317	3.242	3.329	3.307	3.252	3.525	3.351	3.348	3.581	3.284	3.255	3.254	3.245
5	2.902	3.030	2.887	2.846	3.005	3.655	6.313	5.773	3.555	2.248	2.744	2.902	2.817	3.051	3.150	2.858	3.049	3.332
6	3.288	3.348	3.335	3.320	2.817	3.790	3.066	3.360	3.396	3.466	3.425	3.373	3.483	3.544	3.285	3.295	3.287	3.313
7	3.673	3.785	3.702	3.435	3.453	3.870	3.958	3.627	4.114	4.866	4.021	5.540	3.917	3.987	3.506	3.646	3.694	3.599
8	3.088	2.895	3.199	3.058	3.373	5.251	3.183	3.165	2.946	2.772	3.185	3.301	3.384	3.189	2.847	2.947	2.969	2.759
9	3.525	3.871	3.517	3.458	3.491	3.438	3.616	3.607	3.533	4.469	3.461	3.458	3.466	3.562	3.467	3.556	3.694	3.577
10	2.688	2.163	2.741	2.562	2.745	3.321	3.501	3.301	3.210	4.933	2.885	2.298	2.389	3.264	2.473	2.537	2.658	4.152
11	3.077	3.049	3.139	3.164	3.310	3.520	3.501	3.460	3.466	3.363	3.147	3.158	3.128	3.081	3.084	3.157	3.215	3.064
12	3.601	3.204	3.608	3.575	3.691	3.667	3.754	3.639	3.704	3.722	3.667	3.692	3.683	3.560	3.613	3.657	3.645	3.560
13	4.255	4.510	4.418	4.238	4.644	4.515	4.444	4.744	4.300	4.897	4.382	4.515	3.462	3.320	4.382	4.490	4.491	4.441
14	4.845	4.784	4.342	4.409	4.675	5.084	4.986	4.726	5.117	5.082	5.129	5.128	5.124	4.441	4.637	4.999	5.105	4.829
15	2.337	2.576	2.190	2.402	2.481	3.407	2.857	2.904	2.544	2.548	2.463	2.557	2.560	2.400	2.528	2.404	2.439	2.995

Reproducible radiomic features of three segmentation groups

Manual Group	GraphCut Group	GrowCut Group
IH-Energy	IH-Energy	IH-Energy
IH-GlobEntr	IH-GlobEntr	IH-GlobEntr
IH-GlobMean	IH-GlobMean	IH-GlobMax
IH-GlobMedi	IH-GlobMedi	IH-GlobMean
IH-GlobStd	IH-GlobUnif	IH-GlobMedi
IH-GlobUnif	IH-10Percentile	IH-GlobStd
IH-10Percentile	IH-90Percentile	IH-GlobUnif
IH-90Percentile	IH-99Percentile	IH-Kurtosis
IH-99Percentile	IH-Range	IH-1Percentile
IH-RootMeanSqua	IH-RootMeanSqua	IH-10Percentile
IH-Variance	IH-Variance	IH-90Percentile
GLCM-AutoCorr	GLCM-AutoCorr	IH-99Percentile
GLCM-DiffEntr	GLCM-ClusTend	IH-RootMeanSqua
GLCM-Dissimilarity	GLCM-Contrast	IH-Variance
GLCM-Energy	GLCM-Correlation	GLCM-AutoCorr
GLCM-Entropy	GLCM-DiffEntr	GLCM-ClusTend
GLCM-Homo1	GLCM-Dissimilarity	GLCM-Contrast
GLCM-Homo2	GLCM-Energy	GLCM-Correlation
GLCM-InveDiffNorm	GLCM-Entropy	GLCM-DiffEntr
GLCM-InveVariance	GLCM-Homo1	GLCM-Dissimilarity
GLCM-MaxProb	GLCM-Homo2	GLCM-Energy
GLCM-SumAver	GLCM-InfoMeasCorr1	GLCM-Entropy
GLCM-SumEntro	GLCM-InfoMeasCorr2	GLCM-Homo1
GLCM-SumVariance	GLCM-InveDiffMomeNorm	GLCM-Homo2
GLCM-Variance	GLCM-InveDiffNorm	GLCM-InfoMeasCorr1
GLRLM-G.L.Nonuniform	GLCM-MaxProb	GLCM-InfoMeasCorr2
GLRLM-H.G.L.R.Empha	GLCM-SumAver	GLCM-InveDiffMomeNorm
GLRLM-L.R.Emphasis	GLCM-SumEntro	GLCM-InveDiffNorm
GLRLM-L.R.H.G.L.Empha	GLCM-SumVariance	GLCM-InveVariance
GLRLM-L.R.L.G.L.Empha	GLCM-Variance	GLCM-MaxProb
GLRLM-L.G.L.R.Empha	GLRLM-G.L.Nonuniform	GLCM-SumAver
GLRLM-R.L.Nonuniformity	GLRLM-H.G.L.R.Empha	GLCM-SumEntro
GLRLM-R.Percentage	GLRLM-L.R.Emphasis	GLCM-SumVariance
GLRLM-S.R.Emphasis	GLRLM-L.R.H.G.L.Empha	GLCM-Variance
GLRLM-S.R.H.G.L.Empha	GLRLM-L.R.L.G.L.Empha	GLRLM-G.L.Nonuniform
GLRLM-S.R.L.G.L.Empha	GLRLM-L.G.L.R.Empha	GLRLM-H.G.L.R.Empha
NGTDM-Busyness	GLRLM-R.L.Nonuniformity	GLRLM-L.R.Emphasis
NGTDM-Complexity	GLRLM-R.Percentage	GLRLM-L.R.H.G.L.Empha
NGTDM-Contrast	GLRLM-S.R.Emphasis	GLRLM-L.R.L.G.L.Empha
SHAPE-Compactness1	GLRLM-S.R.H.G.L.Empha	GLRLM-L.G.L.R.Empha
SHAPE-Compactness2	GLRLM-S.R.L.G.L.Empha	GLRLM-R.L.Nonuniformity
SHAPE-ConvHullVol	NGTDM-Busyness	GLRLM-R.Percentage
SHAPE-ConvHullVol3D	NGTDM-Coarseness	GLRLM-S.R.Emphasis
SHAPE-Mass	NGTDM-Contrast	GLRLM-S.R.H.G.L.Empha
SHAPE-Max3DDiam	SHAPE-Compactness1	GLRLM-S.R.L.G.L.Empha
SHAPE-MeanBreadth	SHAPE-Compactness2	NGTDM-Busyness
SHAPE-NumOfVoxel	SHAPE-ConvHullVol	NGTDM-Coarseness
SHAPE-Orientation	SHAPE-ConvHullVol3D	NGTDM-Contrast
SHAPE-Sphericity	SHAPE-Mass	SHAPE-Compactness1
SHAPE-SurfArea	SHAPE-Max3DDiam	SHAPE-ConvHullVol
SHAPE-SurfAreaDens	SHAPE-MeanBreadth	SHAPE-ConvHullVol3D
SHAPE-Volume	SHAPE-NumOfVoxel	SHAPE-Mass
	SHAPE-SurfArea	SHAPE-Max3DDiam
	SHAPE-SurfAreaDens	SHAPE-MeanBreadth
	SHAPE-Volume	SHAPE-NumOfVoxel
		SHAPE-SurfArea
		SHAPE-SurfAreaDens
		SHAPE-Volume

Figure S1 Reproducible radiomic features of three segmentation groups.

A Intra-class correlation coefficient (ICC) value of radiomic features

Category		Feature	Manual Group	GraphCut Group	GrowCut Group
Intensity Histogram (IH)		Energy	0.968	0.992	0.976
		GlobalEntropy	0.911	0.934	0.946
		GlobalMax	0.647	0.120	0.855
		GlobalMean	0.779	0.982	0.988
		GlobalMedian	0.979	0.988	0.987
		GlobalMin	0.664	0.563	0.675
		GlobalStd	0.853	0.139	0.887
		GlobalUniformity	0.914	0.907	0.958
		Kurtosis	0.360	0.605	0.938
		1Percentile	0.745	0.679	0.784
		10Percentile	0.985	0.969	0.995
		90Percentile	0.961	0.980	0.996
		99Percentile	0.911	0.973	0.977
		Range	0.104	0.782	0.726
		RootMeanSquare	0.878	0.976	0.988
		Skewness	0.338	0.542	0.523
		Variance	0.844	0.904	0.924
Textural Features	Gray-level Co-occurrence Matrix (GLCM)	AutoCorrelation	0.966	0.983	0.990
		ClusterProminence	0.649	0.618	0.608
		ClusterShade	0.595	0.589	0.510
		ClusterTendendcy	0.229	0.894	0.908
		Contrast	0.189	0.895	0.936
		Correlation	0.592	0.850	0.863
		DifferenceEntropy	0.911	0.933	0.936
		Dissimilarity	0.919	0.954	0.969
		Energy	0.922	0.943	0.976
		Entropy	0.928	0.948	0.961
		Homogeneity	0.961	0.974	0.981
		Homogeneity2	0.958	0.973	0.980
		InformationMeasureCorr1	0.731	0.881	0.903
		InformationMeasureCorr2	0.721	0.844	0.865
		InverseDiffMomentNorm	0.507	0.896	0.936
		InverseDiffNorm	0.928	0.956	0.970
		InverseVariance	0.939	0.748	0.984
		MaxProbability	0.951	0.954	0.982
		SumAverage	0.893	0.983	0.990
		SumEntropy	0.878	0.929	0.945
		SumVariance	0.929	0.985	0.992
		Variance	0.571	0.894	0.908
	Neighbor Gray-Tone Difference Matrix (NGTDM)	Busyness	0.959	0.941	0.944
		Coarseness	0.451	0.901	0.969
		Complexity	0.828	0.672	0.490
		Contrast	0.946	0.890	0.898
		TextureStrength	0.730	0.402	0.419
	Gray-Level Run Length Matrix (GLRLM)	GL.Nonuniformity	0.927	0.987	0.985
		HGLR.Empha	0.896	0.971	0.985
		LR.Emphasis	0.818	0.993	0.996
		LRHGL.Empha	0.895	0.993	0.996
		LRLGL.Empha	0.791	0.993	0.996
		LGLR.Empha	0.985	0.976	0.964
		RL.Nonuniformity	0.960	0.992	0.960
		R.Percentage	0.996	0.990	0.996
		SR.Emphasis	0.995	0.987	0.984
		SRHGL.Empha	0.993	0.780	0.986
		SRLGLE.mpha	0.892	0.989	0.982
Shape		Compactness1	0.957	0.973	0.971
		Compactness2	0.761	0.801	0.529
		Convex	0.425	0.245	0.188
		ConvexHullVolume	0.964	0.988	0.984
		ConvexHullVolume3D	0.903	0.987	0.985
		Mass	0.961	0.994	0.982
		Max3DDiameter	0.935	0.978	0.991
		MeanBreadth	0.910	0.861	0.880
		NumberOfVoxel	0.751	0.992	0.976
		Orientation	0.832	0.480	0.528
		Roundness	0.620	0.641	0.689
		SphericalDisproportion	0.746	0.074	0.286
		Sphericity	0.756	0.118	0.436
		SurfaceArea	0.954	0.924	0.986
		SurfaceAreaDensity	0.916	0.795	0.847
		Volume	0.961	0.982	0.994

B Inter-observer reproducibility of radiomic features (ICC)

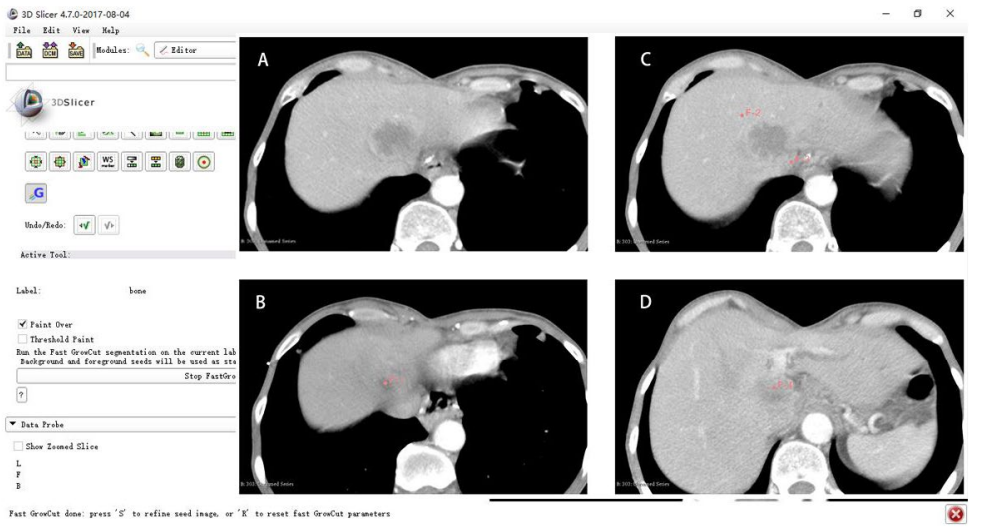
Feature	manual_ inter_obse. Average	GraphCut_ inter_obse._ run1	GraphCut_ inter_obse._ run2	GrowCut_ inter_obse._ run1	GrowCut_ inter_obse._ run2
Energy	0.912	0.991	0.957	0.980	0.993
GlobalEntropy	0.723	0.671	0.451	0.748	0.160
GlobalMax	0.697	0.341	0.646	0.832	0.398
GlobalMean	0.844	0.921	0.770	0.837	0.918
GlobalMedian	0.896	0.957	0.028	0.929	0.904
GlobalMin	0.651	0.552	0.126	0.872	0.900
GlobalStd	0.906	0.958	0.847	0.956	0.874
GlobalUniformity	0.880	0.976	0.810	0.965	0.953
Kurtosis	0.622	0.964	0.872	0.979	0.963
1Percentile	0.762	0.962	0.855	0.964	0.940
10Percentile	0.834	0.986	0.930	0.980	0.968
90Percentile	0.732	0.985	0.925	0.979	0.966
99Percentile	0.883	0.758	0.498	0.888	0.949
Range	0.223	0.655	0.334	0.889	0.884
RootMeanSquare	0.795	0.957	0.421	0.929	0.905
Skewness	0.894	0.977	0.841	0.966	0.954
Variance	0.776	0.983	0.898	0.976	0.983
AutoCorrelation	0.926	0.978	0.928	0.972	0.978
ClusterProminence	0.948	0.991	0.958	0.980	0.993
ClusterShade	0.738	0.936	0.776	0.951	0.924
ClusterTendendcy	0.820	0.994	0.980	0.987	0.990
Contrast	0.937	0.921	0.178	0.837	0.918
Correlation	0.817	0.989	0.991	0.944	0.987
DifferenceEntropy	0.925	0.936	0.887	0.958	0.914
Dissimilarity	0.779	0.780	0.080	0.920	0.794
Energy	0.976	0.979	0.979	0.977	0.991
Entropy	0.842	0.951	0.803	0.982	0.848
Homogeneity	0.429	0.168	0.672	0.880	0.712
Homogeneity2	0.906	0.884	0.386	0.917	0.813
InformationMeasureCorr1	0.748	0.950	0.894	0.969	0.938
InformationMeasureCorr2	0.556	0.672	0.597	0.776	0.204
InverseDiffMomentNorm	0.789	0.968	0.963	0.835	0.673
InverseDiffNorm	0.972	0.994	0.996	0.995	0.993
InverseVariance	0.943	0.962	0.966	0.959	0.989
MaxProbability	0.153	0.710	0.635	0.979	0.970
SumAverage	0.804	0.464	0.763	0.900	0.737
SumEntropy	0.849	0.978	0.966	0.977	0.991
SumVariance	0.721	0.744	0.137	0.777	0.369
Variance	0.729	0.895	0.192	0.871	0.880
Busyness	0.707	0.907	0.949	0.886	0.820
Coarseness	0.886	0.893	0.854	0.964	0.960
Complexity	0.791	0.517	0.483	0.818	0.313
Contrast	0.901	0.817	0.933	0.847	0.897
TextureStrength	0.743	0.391	0.500	0.325	0.599
GL.Nonuniformity	0.972	0.981	0.984	0.977	0.982
HGLR.Empha	0.938	0.976	0.958	0.975	0.985
LR.Emphasis	0.684	0.995	0.987	0.999	0.990
LRHGL.Empha	0.968	0.994	0.987	0.999	0.990
LRLGL.Empha	0.963	0.995	0.987	0.999	0.991
LGLR.Empha	0.551	0.976	0.979	0.978	0.980
RL.Nonuniformity	0.941	0.990	0.990	0.911	0.978
R.Percentage	0.961	0.993	0.982	0.998	0.992
SR.Emphasis	0.629	0.993	0.976	0.992	0.969
SRHGL.Empha	0.931	0.992	0.961	0.991	0.975
SRLGLE.mpha	0.650	0.990	0.982	0.991	0.963
Compactness1	0.970	0.979	0.949	0.947	0.974
Compactness2	0.838	0.634	0.623	0.484	0.483
Convex	0.522	0.398	0.466	0.366	0.298
ConvexHullVolume	0.945	0.990	0.986	0.965	0.991
ConvexHullVolume3D	0.955	0.987	0.987	0.965	0.992
Mass	0.930	0.992	0.993	0.958	0.990
Max3DDiameter	0.954	0.974	0.978	0.988	0.991
MeanBreadth	0.929	0.929	0.793	0.775	0.973
NumberOfVoxel	0.958	0.989	0.991	0.947	0.985
Orientation	0.761	0.285	0.381	0.530	0.582
Roundness	0.603	0.837	0.565	0.695	0.715
SphericalDisproportion	0.652	0.647	0.622	0.276	0.470
Sphericity	0.802	0.653	0.630	0.406	0.477
SurfaceArea	0.923	0.959	0.895	0.976	0.988
SurfaceAreaDensity	0.915	0.702	0.772	0.796	0.911
Volume	0.960	0.993	0.993	0.957	0.990

C Intra-observer reproducibility of radiomic features (ICC)

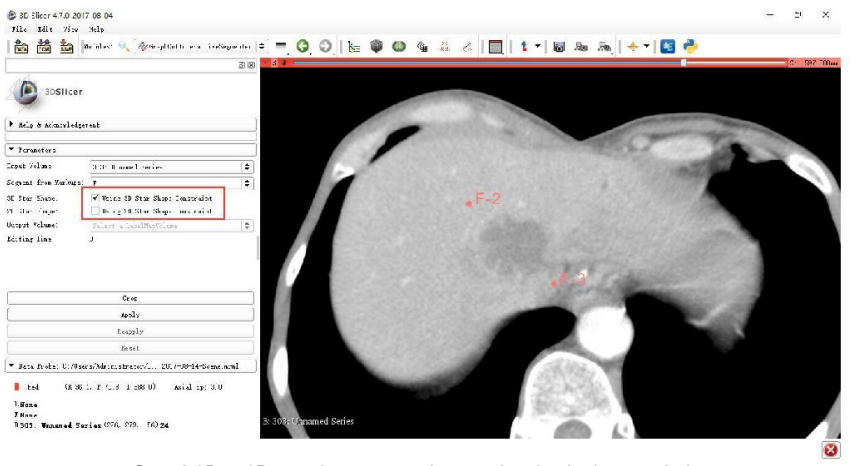
Feature	manual_ intra_obse. Average	GraphCut_ intra_obse.1	GraphCut_ intra_obse.2	GrowCut_ intra_obse.1	GrowCut_ intra_obse.2
Energy	0.967	0.999	0.976	0.997	0.994
GlobalEntropy	0.883	0.988	0.498	0.968	0.841
GlobalMax	0.642	0.940	0.498	0.886	0.343
GlobalMean	0.975	0.997	0.490	0.984	0.877
GlobalMedian	0.976	0.695	0.420	0.988	0.892
GlobalMin	0.582	0.659	0.543	0.951	0.748
GlobalStd	0.817	0.795	0.945	0.986	0.913
GlobalUniformity	0.889	0.697	0.885	0.994	0.962
Kurtosis	0.518	0.596	0.964	0.995	0.960
1Percentile	0.800	0.796	0.939	0.993	0.943
10Percentile	0.989	0.797	0.971	0.996	0.979
90Percentile	0.946	0.867	0.968	0.996	0.978
99Percentile	0.891	0.900	0.702	0.963	0.829
Range	0.510	0.952	0.727	0.957	0.759
RootMeanSquare	0.974	0.795	0.391	0.988	0.894
Skewness	0.734	0.797	0.913	0.994	0.964
Variance	0.816	0.829	0.940	0.994	0.985
AutoCorrelation	0.974	0.847	0.966	0.997	0.986
ClusterProminence	0.534	0.999	0.977	0.997	0.994
ClusterShade	0.651	0.865	0.705	0.990	0.909
ClusterTendendcy	0.798	0.759	0.788	0.798	0.895
Contrast	0.880	0.771	0.490	0.633	0.877
Correlation	0.686	0.999	0.994	0.747	0.994
DifferenceEntropy	0.895	0.994	0.968	0.661	0.922
Dissimilarity	0.928	0.942	0.864	0.540	0.758
Energy	0.902	0.999	0.983	0.997	0.993
Entropy	0.885	0.921	0.919	0.986	0.872
Homogeneity	0.952	0.969	0.633	0.797	0.633
Homogeneity2	0.949	0.995	0.747	0.979	0.898
InformationMeasureCorr1	0.750	0.991	0.961	0.991	0.936
InformationMeasureCorr2	0.618	0.921	0.402	0.907	0.725
InverseDiffMomentNorm	0.880	0.842	0.974	0.894	0.779
InverseDiffNorm	0.930	0.879	0.997	0.649	0.997
InverseVariance	0.927	0.828	0.984	0.535	0.984
MaxProbability	0.939	0.843	0.611	0.994	0.974
SumAverage	0.981	0.863	0.538	0.872	0.658
SumEntropy	0.863	0.999	0.971	0.996	0.993
SumVariance	0.988	0.936	0.199	0.834	0.751
Variance	0.814	0.988	0.606	0.983	0.842
Busyness	0.977	0.633	0.916	0.959	0.882
Coarseness	0.968	0.747	0.969	0.988	0.969
Complexity	0.994	0.961	0.835	0.940	0.775
Contrast	0.997	0.402	0.817	0.945	0.935
TextureStrength	0.996	0.974	0.803	0.954	0.843
GL.Nonuniformity	0.994	0.998	0.991	0.995	0.997
HGLR.Empha	0.940	0.999	0.967	0.996	0.989
LR.Emphasis	0.997	0.999	0.992	0.999	0.995
LRHGL.Empha	0.995	0.999	0.992	0.998	0.995
LRLGL.Empha	0.992	0.999	0.992	0.999	0.996
LGLR.Empha	0.988	0.999	0.987	0.996	0.986
RL.Nonuniformity	0.956	0.999	0.993	0.983	0.991
R.Percentage	0.905	0.998	0.993	0.999	0.995
SR.Emphasis	0.676	0.997	0.987	0.997	0.982
SRHGL.Empha	0.899	0.996	0.981	0.997	0.984
SRLGLE.mpha	0.472	0.998	0.987	0.997	0.980
Compactness1	0.961	0.998	0.972	0.879	0.988
Compactness2	0.797	0.742	0.552	0.858	0.760
Convex	0.527	0.193	0.640	0.806	0.262
ConvexHullVolume	0.968	0.999	0.982	0.993	0.995
ConvexHullVolume3D	0.961	0.999	0.978	0.863	0.994
Mass	0.966	0.999	0.994	0.992	0.995
Max3DDiameter	0.954	0.994	0.974	0.757	0.995
MeanBreadth	0.922	0.989	0.727	0.995	0.876
NumberOfVoxel	0.944	0.999	0.995	0.991	0.994
Orientation	0.870	0.766	0.510	0.723	0.614
Roundness	0.704	0.560	0.628	0.856	0.592
SphericalDisproportion	0.788	0.661	0.204	0.860	0.279
Sphericity	0.796	0.727	0.217	0.860	0.846
SurfaceArea	0.962	0.995	0.879	0.936	0.988
SurfaceAreaDensity	0.939	0.510	0.857	0.985	0.798
Volume	0.965	0.999	0.995	0.962	0.995

Figure S2 The intra-class correlation coefficient (ICC) of (A) radiomic features, (B) inter-observer and (C) intra-observer.

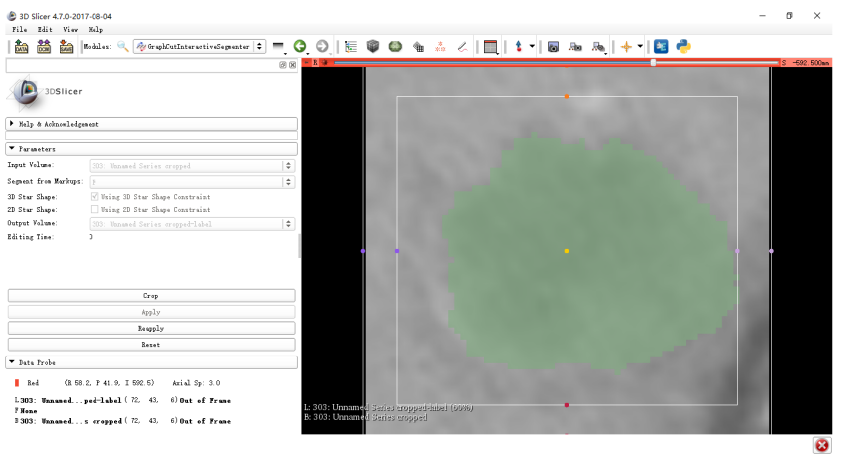
Part I Workflow of GraphCut segmentation tool in 3D Slicer

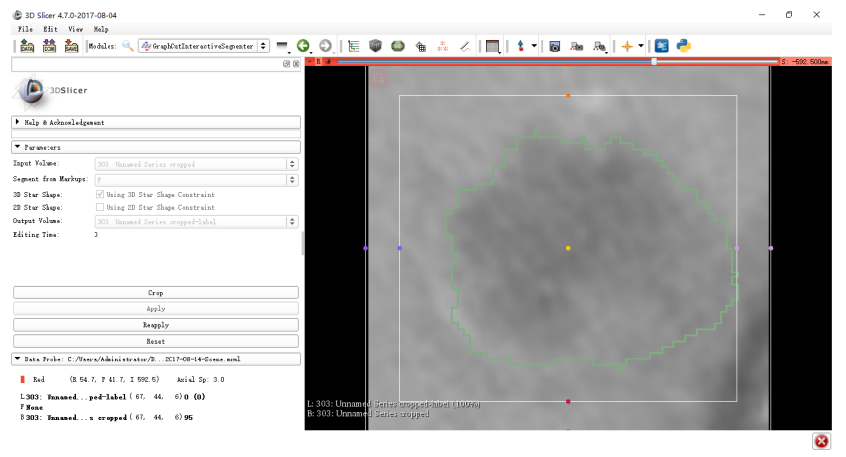


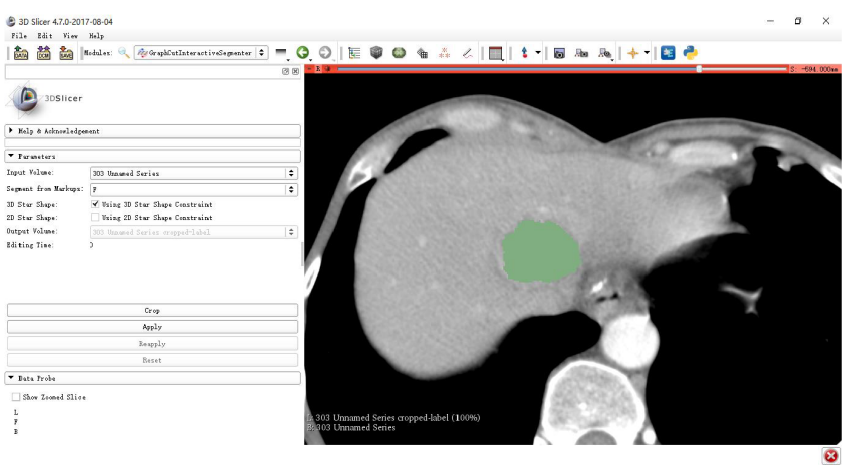
Step 1 (A) Load the CT images and add four fiducials around the tumor. Two fiducials on the first slice and last slice where the (B) tumor begins and (C) ends to show, at the center of the tumor. Two fiducials on the middle slice where the (D)tumor area is the largest, at the diagonal corners of a rectangle which can contain the tumor.



Step 2 2D or 3D star shape constraints can be checked as needed.

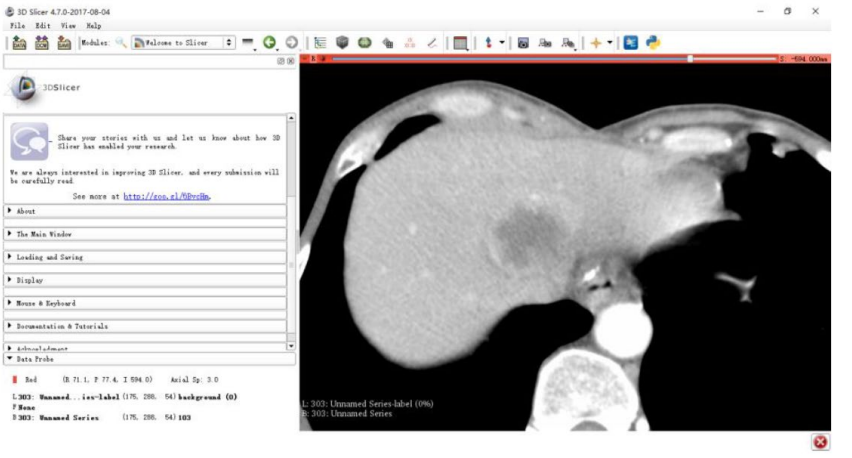




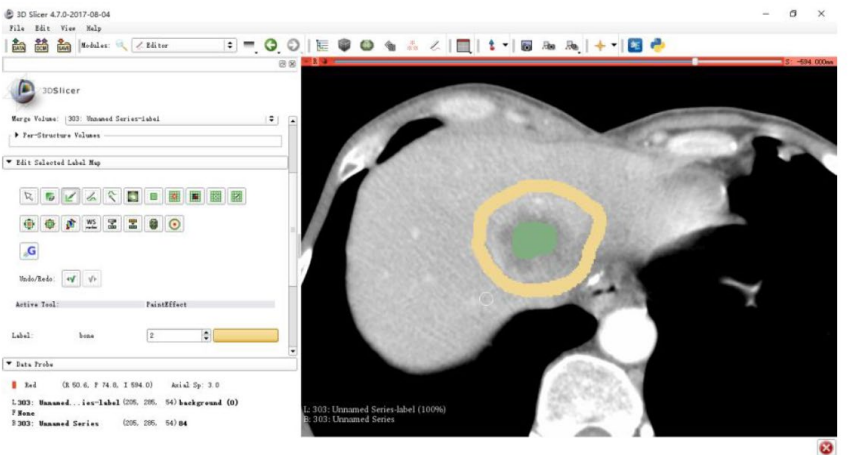


Step 3 The final segmentation result.

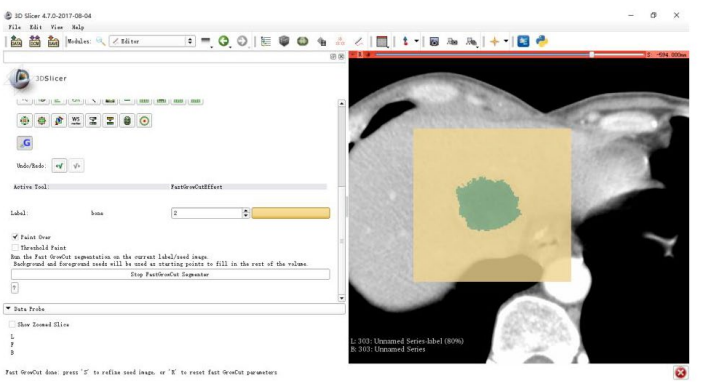
Part II Workflow of GrowCut segmentation tool in 3D Slicer



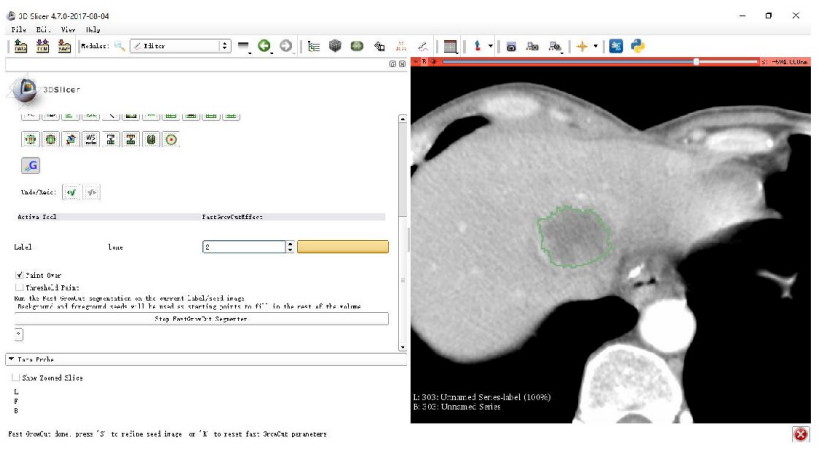
Step 1 Load the CT images.



Step 2 Define the tumor (green labels) and non-tumor (yellow labels) region in different label value. A set of initial labels need given by users to mark foreground and background.



Step 3 The GrowCut algorithm automatically computes a region of interest by using cellular automata with a weighted similarity score. If not satisfy with the result, the foreground tumor region can be edited manually both in GrowCut and GraphCut.



Step 4 The final segmentation result.

Figure S3 (Part I) Workflow of GraphCut segmentation tool in 3D Slicer and (Part II) workflow of GrowCut segmentation tool in 3D Slicer.