An Open-Source Toolkit for the Volumetric Measurement of CT Lung Lesions[◊]

Karthik Krishnan,[1,*] Luis Ibanez,[1] Wesley D. Turner,[1]
Julien Jomier,[1] and Ricardo S. Avila[1]

[1]Kitware Inc., 28 Corporate Drive, Clifton Park, NY 12065, USA
[*]karthik.krishnan@kitware.com

Abstract: An open source lesion sizing toolkit has been developed with a general architecture for implementing lesion segmentation algorithms and a reference algorithm for segmenting solid and part-solid lesions in lung CT scans. The CT lung lesion segmentation algorithm detects four three-dimensional features corresponding to the lung wall, vasculature, lesion boundary edges, and low density background lung parenchyma. These features form boundaries and propagation zones that guide the evolution of a subsequent shape detection level set. User input is used to determine an initial seed point for the level set and users may also define a region of interest around the lesion. The methods are validated against 18 nodules using CT scans of an anthropomorphic thorax phantom simulating lung anatomy. The scans were acquired under differing scanner parameters to characterize algorithm behavior under varying acquisition scenarios. We also validated repeatability using six coffee-break (zero volume change) clinical cases. The source code, datasets, and a running application are all provided under an unrestrictive license to encourage reproducibility and foster scientific exchange.

[◊]Datasets associated with this article are available at http://midas.osa.org/midaspre/midas/item/view/886?key=a3dGZzNXdnR3eFZqaw==.

References and links

1. WHO handbook for reporting results of cancer treatment. (1979).
2. P. Therasse, et al, "New guidelines to evaluate the response to treatment in solid tumors," J. National Cancer Institute 92(3), 205–16 (Feb 2000).
3. E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, J. Verweij, "New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)," Volume 45, Issue 2, Pages 228-247 (January 2009).
4. Lesion sizing toolkit, http://public.kitware.com/LesionSizingKit
5. K. Krishnan, L. Ibanez, W. D. Turner, R. S. Avila, "Algorithms, Architecture, Validation of an open source toolkit for segmenting CT Lung Lesions," Proc. MICCAI Workshop on Pulmonary Image Analysis, 365-75 (Sep 2009).
6. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: ITK Software Guide. Kitware Inc.
7. Schroeder, W., Martin, K., Lorensen, W.: Visualization Toolkit. Kitware Inc.
8. M.Descoteaux, M.Audette, K.Chinzei, el al.: "Bone enhancement filtering: Application to sinus bone segmentation and simulation of pituitary surgery". In: MICCAI. (2005) 9–16.
9. Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, R. Kikinis, "3D Multi-Scale line filter for segmentation and visualization of curvilinear structures in medical images," CVRMed MRCAS 213–22 (1997).
10. A.F.Frangi, W.J.Niessen, K.L.Vincken, et al.: "Multiscale vessel enhancement filtering". Lecture Notes in Comput Science 1496 (1998) 130–137.
11. J. F. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Analysis and Machine Intelligence **8**(6) 679–698 (1986).
12. J. A. Sethian, "Level Set Methods and Fast Marching Methods," Cambridge Press (1999).

13. R. Malladi, J. A. Sethian and B. C. Vermuri.:"Shape Modeling with Front Propagation: A Level Set Approach", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol 17, No. 2, pp 158-174, February 1995
14. Caselles, V. ,Kimmel, R.,Sapiro G.:"Geodesic Active Contours", International Journal on Computer Vision, Vol 22, No. 1, pp 61-97, 1997.
15. W. E. Lorensen, H. E. Cline, "Marching cubes: A high resolution 3d surface con-struction algorithm," SIGGRAPH. 21 163–169 (July 1987).
16. A. M. Alyassin, J. L. Lancaster, J. H. Downs, "Evaluation of new algorithms for interactive measurement of surface area and volume," Medical Physics 21(6) (1994).
17. M. A. Gavrielides, R. Zeng, L. M. Kinnard, K. J. Myers, N. A. Petrick, "A matched filter approach for the analysis of lung nodules in a volumetric CT phantom study," Proc. SPIE Med. Imaging (Feb 2009).
18. M. A. Gavrielides, L. Kinnard, S. Park, I. Kyprianou, B. Gallas, A. Badano, N. Petrick, K. J. Myers, "Quantitative in silico imaging and biomarker assessment using physical and computational phantoms: a review of new tools and methods available from the NIBIB/CDRH joint Laboratory for the Assessment of Medical Imaging Systems," Radiology (2008).
19. The Optical Society of America, The National Library of Medicine, Kitware Inc,. Interactive Science Publishing: http://www.opticsinfobase.org/isp.cfm
20. Kitware Inc. Volview, http://www.kitware.com/VolView

---

1. Introduction

Computed Tomography (CT) is the preferred modality for the detection of lung cancer lesions. Quantifying tumor size and size change over time is necessary to evaluate the effectiveness of a drug. Current practice uses a low-dimensional surrogate for tumor volume such as the bi-dimensional WHO criteria [1] or the uni-dimensional RECIST measurement approach [2] [3], with RECIST being the currently preferred approach. These methods do not fully evaluate the volume of lesions and instead use a manually measured lesion diameter or diameters. The advent of high resolution CT lung scanning offers the opportunity to perform volumetric segmentation of lung nodules, thereby allowing for a more accurate quantification of changes in full tumor burden over time in response to a drug regimen.

In this paper we describe the software architecture available in the Lesion Sizing Toolkit [4] as well as the reference algorithm for CT lung lesion sizing provided with the toolkit. Performance of the CT lung lesion sizing algorithm was evaluated on synthetic phantom data as well as on clinical data. The paper is organized into five sections. Section 2 describes the general architecture of the toolkit. Sections 3 and 4 summarize the segmentation methods employed for CT lung lesion segmentation. Section 5 describes the validation of the methods on synthetic phantoms having known volumetric ground truth as well as results from a small set of "zero change" clinical data. Finally, Section 6 describes the implementation of the toolkit in ISP, a data visualization application developed to support interactive science publishing.

2. Software Architecture

The Lesion Sizing Toolkit [4][5] provides a cross platform, C++ software architecture based on the National Library of Medicine's Insight Segmentation and Registration Toolkit (ITK) [6] and the Visualization Toolkit (VTK) [7]. The main goal of the software architecture is to support a wide range of lesion sizing algorithms and methods with a set of modular software elements that can easily be modified or replaced. The toolkit framework consists of four software elements that plug into a main "Segmenter" class:

- *Feature Generator* : This generator produces one or more "features" to guide the segmentation thereby preventing the segmentation from bleeding into non-lesion regions. Feature generators provided with the toolkit include lung wall feature generators, vessel / airway feature generators, and edge detection feature generators.
- *Feature Aggregator* : Conceptually, each feature image represents the likelihood of a particular voxel being part of the lesion from a single perspective. The feature aggregator

defines how multiple features are combined to form a single aggregate feature that can be passed onto a segmentation module.

- *Segmentation Module* : The segmentation module runs a segmentation algorithm on the input image, with guidance from the feature image. A variety of segmentation methods are provided including level sets, geodesic active contours, and region growing.
- *Spatial Initialization* : Spatial initializations help initialize or constrain a segmentation algorithm. The toolkit supports the specification of one or more starting seed points or regions to be used to initialize the segmentation. Additionally, a bounding box or region of interest can be defined to contain the resulting segmentation.

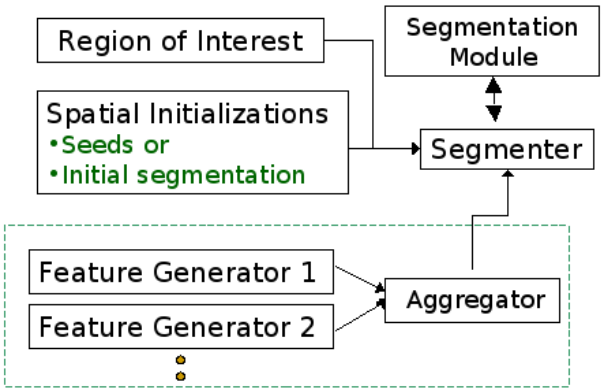

Figure 1. Toolkit architecture.

Figure 1 illustrates the general architecture of the toolkit. A segmentation module, such as a level set algorithm, is typically initialized by a spatial seed point and guided by at least one image based feature. This modular architecture, similar to the registration framework of ITK [6], supports rapid evaluation of different algorithmic components of a segmentation algorithm.

The following is a list of feature generators, feature aggregators, and segmentation modules currently available with the toolkit:

Feature Generators:

- Descoteaux Sheetness feature [8]
- Sato Vesselness feature [9]
- Frangi tubularness feature [10]
- Gradient Magnitude Edge feature
- Canny Edge feature [11]
- Lung wall feature
- Intensity feature

Feature Aggregators

- Minimum feature aggregator
- Weighted sum feature aggregator

Segmentation Modules:

- Connected Threshold region growing
- Confidence connected region growing
- Isolated connected region growing
- Iterative connected region growing
- Fast Marching level set [12]
- Shape detection level set [13]

- Geodesic active contour level set [14]

3. CT Lung Lesion Segmentation Algorithm

Lung nodules exhibit a wide range of morphologies and shapes. They may be solid, part-solid (solid nodules surrounded by a diffuse region) or non-solid. Solid nodules consist of a CT density that is similar to soft tissue while non-solid nodules have varying densities between air and soft tissue. Nodules are often attached to vessels and can be attached to the lung wall or the mediastinum. The ability to define and combine an unlimited number of feature generators within the toolkit was designed to support the wide diversity of potential lesion presentations observed in the lung and other organs.

The main algorithm used for segmentation is a region evolution algorithm and specifically, a shape detection level set. The level set is allowed to evolve from a user-provided seed under the aggregate guidance of the defined features. Our current implementation uses four features to define areas that region growing should avoid. These are (i) the lung wall, (ii) lung vasculature, (iii) the boundary between the lesion and the background lung parenchyma, and finally (iv) the low CT density of air. Together, these features allow propagation of the segmentation front within the intensity thresholds specified by the intensity feature while deterring propagation into the lung wall, vessels, or across sharp edges. With a modular architecture, the addition of new features to specifically account for more complex anatomical areas, such as the hilum, is easily supported. The CT dataset is preprocessed by resampling to isotropic resolution matching the smallest voxel spacing.

3.1 Feature Generation

Next we describe the four features that are computed by the algorithm: Lung wall, vesselness, gradient, and intensity.

3.1.1 Lung wall feature

The lung wall is segmented by intensity thresholding followed by curvature constrained front propagation. The dataset is binarized so as to retain voxels above -400HU. We then run hole filling on the resulting binary image. The hole filling operation uses a quorum (voting) algorithm to decide whether a new pixel will be filled at every iteration. Changing the voting threshold is essentially equivalent to changing the maximum curvature allowed in the final binary segmentation. We use a 7×7×7 Manhattan kernel with a voting threshold of 1 (lowest possible curvature). The filter is run iteratively until convergence. The algorithm is summarized in Algorithm 1. The evolution of the front is shown in Figure 2. Admittedly, the final contour is slightly concave rather than convex, since the quorum voting does not distinguish between positive and negative curvature, but the wall detection recovers most of the nodule without allowing bleeding into the lung wall.

Algorithm 1. Voting based front propagation

1: Let Kernel Radius $r = 3$
2: Let BirthThreshold, $T = r^3 + 1$. For a 7×7×7 city-block kernel, $T = 171$
3: Let A_i be indices of voxels on the front at iteration i
4: Threshold CT image at -400HU. Voxels above are set to 1, below to 0.
5: Iterate over the image and add all indices within r voxels of the boundary to A_0
6: **repeat**
7: Let *PixelsChangingStatus* = 0
8: **for** each voxel in A_i **do**
9: Check for quorum, Q.
(Q is true if number of ON voxels within neighborhood (7×7×7) centered at the current voxel > T)
10: **if** Q is true **then**
11: Schedule this pixel for inclusion in foreground.
12: ++*PixelsChangingStatus*
13: Add background voxels in neighborhood to front for next iteration, A_{i+1}.

14: **else**
15: Schedule this pixel for inclusion in background.
16: Add this location to the front for the next iteration, A_{i+1}
17: **end if**
18: **end for**
19: Update the status of scheduled voxels.
20: $++i$
21: **until** *PixelsChangingStatus* = 0

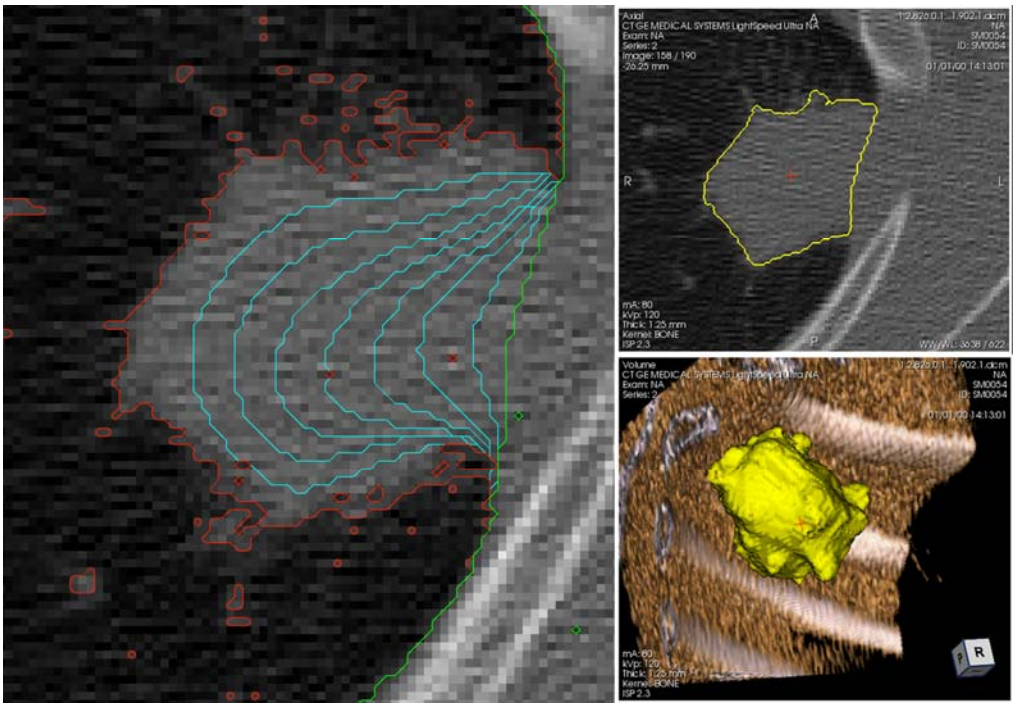

Figure 2: *Left:* Segmentation of the lung wall by voting based front propagation on a lesion attached to the lung wall. The front starts with the red contour (obtained by thresholding the dataset at -400HU. In cyan are contours of the front at iterations of 10, 20, 30, 40, 50, 60, 80. The algorithm converges in 151 iterations at the green front. *Right:* The lesion segmented and volume rendered in ISP (View 1).

The resulting image is transformed using a sigmoid with inversion, resulting in a grayscale representation of the inner lung, with a gradual rolloff towards the lung wall.

3.1.2 Vesselness feature

Vesselness features are computed using Sato [9] tubularness with $\alpha = 0.1$ and $\beta = 2.0$. The resulting image is transformed using a sigmoid filter with inversion resulting in a grayscale image with a rolloff towards tubular structures.

3.1.3 Gradient feature

The gradient feature attempts to localize the edges of the lesion where the lesion is adjacent to a low density structure, such as the background lung parenchyma. This is modeled using a Canny edge detector [11], which achieves optimal edge localization (with voxel accuracy), and results in a voxelized edge map. The resulting image is transformed using a sigmoid with inversion resulting in a grayscale image with a roll-off towards the identified boundary edges.

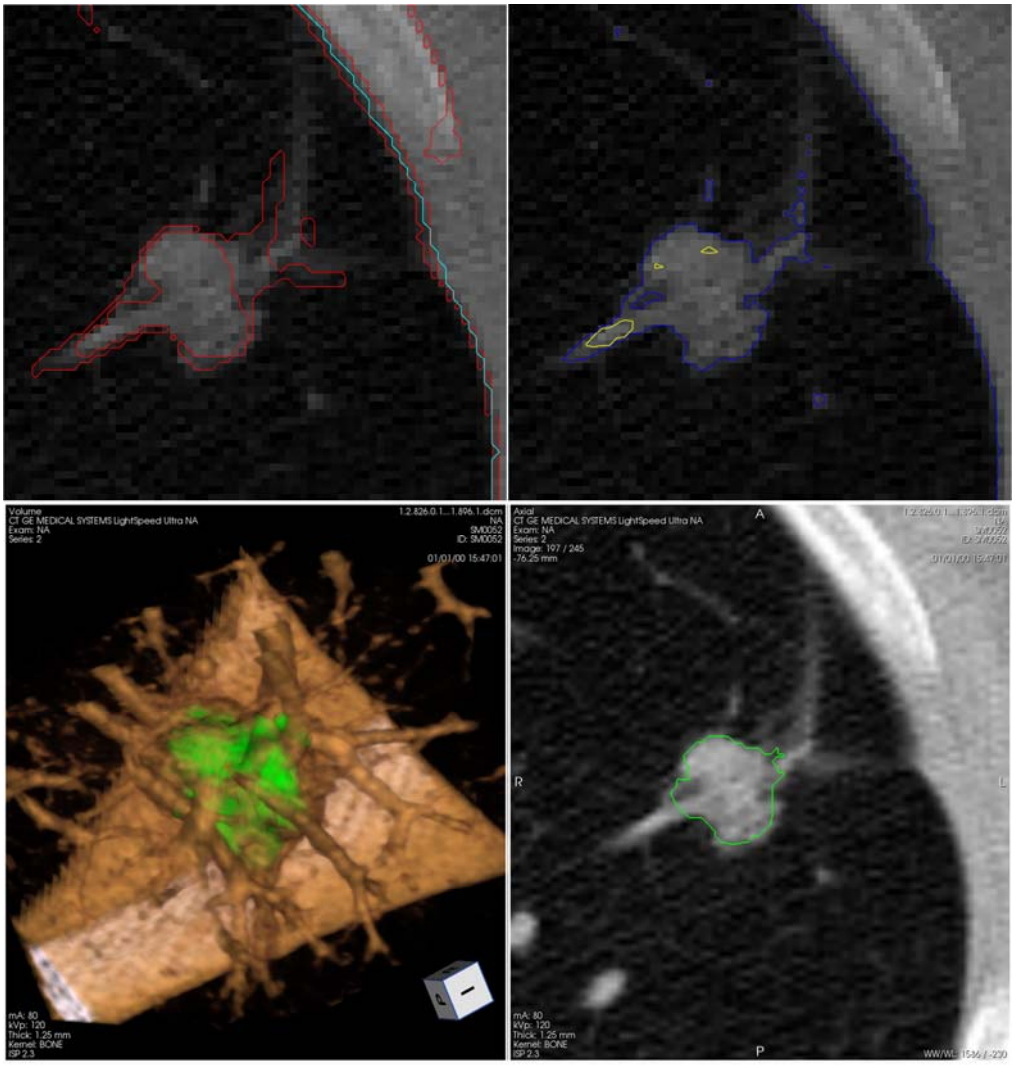

Figure 3: *Top left*: Lung wall features in cyan, Canny Edges features in red; a contour around the single voxel thick edges is shown. Note that the voxels may have a thickness of 1 along the Z dimension and hence several voxels thick in-plane as shown. *Top right*: Intensity features for solid lesion parameters in blue, Sato Vesselness features in yellow. *Bottom*: Lesion segmented and volume rendered using the ISP lesion sizing tool (View 2).

3.1.4 Intensity feature

The intensity feature is based on a simple intensity threshold applied to the dataset, followed by a transformation with a sigmoid filter to generate gradual roll-offs. The user specifies the type of the lesion (solid / part-solid), which determines the threshold level: -200HU for solid lesions and -500HU for part-solid lesions.

3.1.5 Aggregation of features

Figures 3 and 5 show the application of feature generators on a solid and a part-solid lesion. An aggregation of the features is obtained by normalizing each individual feature to lie within the range [0,1] and then choosing the minimum feature value at each voxel. The resulting image is then passed onto the segmentation module.

4. Segmentation Module

We use the fast marching algorithm [12], followed by refinement with a shape detection level set [13]. In our user interface, the user identifies the lesion with a single seed point. The seed is used to initialize the fast marching front, which starts at the seed and proceeds outwards, solving the eikonal equation, its propagation guided solely by the aggregate feature image and an optional, user-defined, region of interest (ROI). The output is a time crossing map, with each voxel representing the time taken for the front to reach that voxel. We use a stopping time of 5s implying that the front stops propagating when all voxels on the front have an arrival time greater than 5. The fast marching front is conservative enough to ensure that the segmentation results lie within the lesion. This resulting time crossing image is renormalized to the range [−4, 4], so as to be symmetric about zero and passed onto the shape detection module.

The shape detection level set is guided by the propagation and curvature components.

$$\varphi_t + F\,(1 - wK)\,|\nabla\varphi| = 0$$

The propagation F is the aggregate feature image. w is the curvature weight for the curvature, K. We use a propagation to curvature weight ratio of 500:1, to account for the wide variety of shapes that lesions can assume. The final lesion boundary is then taken as the -0.5 isosurface of the shape detection level set. Note that this isosurface value is driven by the characteristics of our aggregate feature map. In the nominal case of the lesion having a boundary within the parenchyma, we find that the canny edge is the dominant feature for lesion segmentation. Our canny edge feature only localizes the true edge boundary to a one voxel precision (Figure 3). By taking the level set at -0.5 we force the segmentation to the interior of the canny detected boundary voxel and reduce our error bound from 1 to 0.5 voxels.

A surface of the resulting level set is generated using the Marching Cubes algorithm [15] at an isovalue of -0.5. The volume is computed using a discretized version of the divergence theorem [9], to compute the volume in a closed triangle mesh.

5. Validation of the segmentation algorithm

5.1. Evaluation on an anthropomorphic thorax phantom

We evaluated the results on CT scans of an anthropomorphic thorax phantom [17][18] (courtesy US FDA), containing nodules of various densities, sizes, shapes and at various locations in the chest. This phantom provides ground truth for simple nodules in a realistic anatomy, as shown in Figure 4, that includes attachments to vasculature, airways, and the lung wall. Exposure varied from 20mAs to 200mAs. Ground truth volume information was available only for the spherical nodules. The database consists of 6 CT scans, each containing a 5, 8, and 10mm diameter spherical nodule with a density of 100HU. Three of the scans have a slice thickness of 0.8mm and a z spacing of 0.4mm, while the other three have a slice thickness of 3mm and a z spacing of 1.5mm.

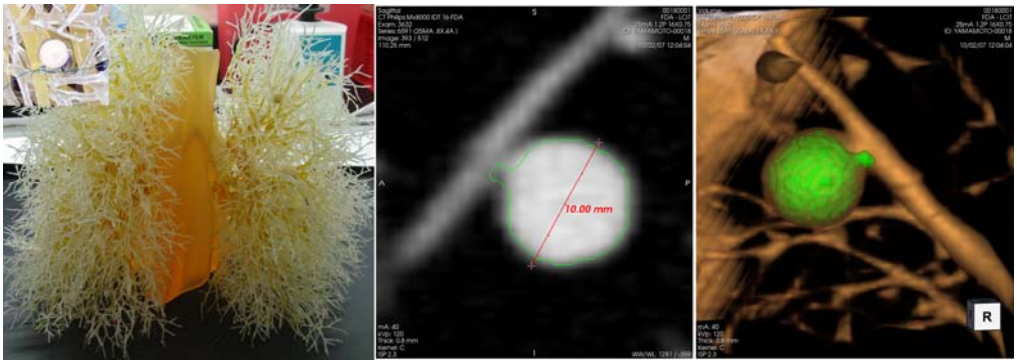

Figure 4: *Left*: A photograph of the anthropomorphic phantom and an inlaid nodule (courtesy US FDA). *Middle, Right*: Lesion segmented using the ISP wizard and volume rendered in ISP (View 3).

Defining the error as,

$$\xi = 100 \times |ComputedVolume - TrueVolume| / TrueVolume,$$

we found that for the 18 lesions in the FDA phantom, our average ξ was 34.88% (S.D 19.77%). One source of error is a poor estimate of the features due to the resolution across slices. This was confirmed by noting that on the 9 lesions, with 0.8mm slice thickness, our average volumetric error, ξ was 21.81% (S.D 12.22%). Furthermore, the half voxel error (as mentioned above) causes an amplification of the errors for the small 5mm lesions. This is confirmed by noting that on the 6 lesions with a diameter ≥ 8mm and a slice thickness of 0.8mm, our average volumetric error ξ is 14.14% (S.D. 5.2%).

It is instructive to relate the error in volumetric estimate to the error in the uni-dimensional RECIST measurement. A 30% overestimate in a lesion with a diameter of 8mm is equivalent to a change in diameter of 0.73mm, Hence a ξ of 30% for a 8mm lesion would roughly translate to an error of a voxel.

5.2. Validation on "zero change" CT scans

Validation on clinical data was done on six zero change CT scans. These are CT scans of a patient obtained hours apart, presumably with insufficient time for a change in lesion volume. However, breathing artifacts, subtle differences in patient positioning, and image noise result in slightly different segmentations. In addition, the two timepoints often had different acquisition parameters. Table 1 provides the performance of the algorithm on all 6 zero change datasets. The average ξ for the "zero change" data was 7.3% (S.D. 9.1%).

Table 1. Zero change Volumes V_1, V_2, for 6 cases each acquired at two time points

Case	V_1 (mm^3)	V_2 (mm^3)	Resolution (t_1)	Resolution (t_2)	$200\times\|V_1-V_2\|/(V_1+V_2)$
ST0108 (View 4)	1925.18	1928.56	.703×.703×1.25	.703×.703×1.25	0.17%
ST0109 (View 5)	2351.59	2548.45	.703×.703×1.25	.703×.703×1.25	8.03%
ST0111 (View 6)	2250.42	1752.87	.703×.703×1.25	.703×.703×1.25	24.85%
ST0112 (View 7)	277.7	293.84	.703×.703×1.25	.703×.703×1.25	5.64%
ST0113 (View 8)	669.15	641.06	.703×.703×1.25	.78×.78×1.25	4.29%
ST0114 (View 9)	2121.84	2102.07	.56×.56×1.25	.56×.56×5	0.93%

6. Implementation

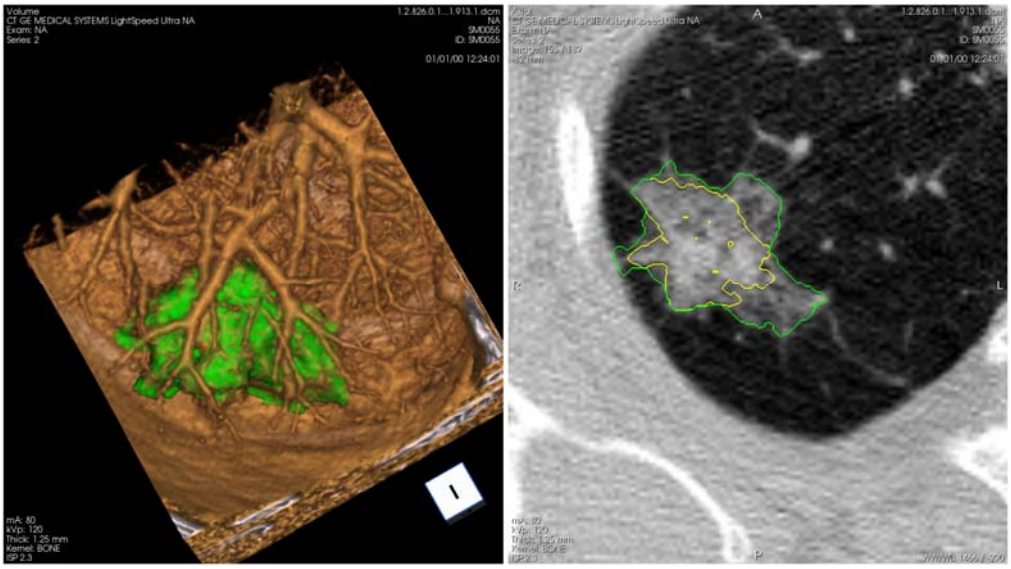

Figure 5: The segmentation of a part-solid lesion (View 10).

The lesion sizing algorithm was added to ISP 2.2[19], a 2D and 3D volume visualization application based on VolView [20]. ISP provides a guided wizard to perform one-click segmentations of solid and part-solid lesions. The user identifies the region containing the lesion with a box, places a seed within the lesion and selects the lesion type (solid or part-solid). The application then segments the lesion, reports volumetric statistics and displays the lesion in the context of the surrounding anatomy. A typical segmentation takes between 5 to 30 seconds, depending on the lesion size on an Intel 2.66GHz Core2Duo with 2GB of RAM.

Figure 6 shows the lesion sizing algorithm applied to a large lung cancer lesion. The lesion was scanned with 2.5mm CT slice thickness at two time points separated by a 10 month time interval. This demonstrates the potential ability of the toolkit to quantitatively measure complex lesions over time, such as is commonly performed in a lung cancer clinical trial setting.

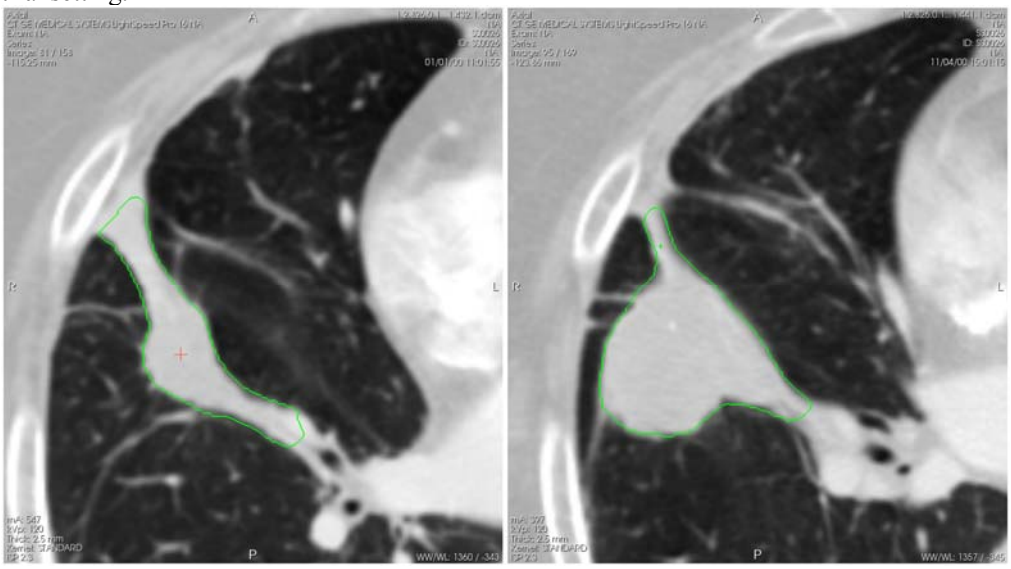

Figure 6: The segmentation of a lung cancer lesion at two time points, imaged 10 months apart. Quantitative analysis indicates that the lesion has more than doubled in volume (View 11).

6. Loading the views in this paper

To load any of the views in this paper, first click on the view link in the paper and wait for the full resolution dataset to be downloaded and displayed. Next navigate to the *Advanced Algorithms* tab. The seed points and the bounding box have already been defined in the view, so executing the segmentation algorithm only requires that the user set the lesion type (solid or part-solid (e.g. Figure 5)) before clicking through the prompts for seed points and region of interest selecting "next" or "finished" as appropriate. For views with more than one time point, such as the zero change datasets, select the first time point's image window (it will become highlighted in green) and walk through the wizard to perform the segmentation. After the segmentation completes, select the second time point's image window and repeat the process.

7. Conclusions

This paper presents an extensible architecture for developing lesion sizing algorithms and demonstrates a CT lung lesion segmentation algorithm developed based on it. Both the framework and the lesion sizing algorithm are available as open access resources. The algorithm has been evaluated on clinical data and an anthropomorphic lung phantom. An important component of this work is reproducibility through open source and sharing of data. Future enhancements will incorporate improved modeling of the acquisition system and additional models of lung anatomy. With continued improvements made by the open source imaging community we hope this toolkit will accelerate the development of algorithms capable of robustly quantifying response to therapy.

7. Acknowledgements

We thank the Optical Society of America and the Air Force Research Laboratory for funding this work. We also thank David Yankelevitz (Weill Cornell Medical Center) and Anthony Reeves (Cornell University) for the I-ELCAP clinical and zero change cases. We also thank Nick Petrick, Lisa Kinnard, Marios Gavrielides and Kyle Myers for the anthropomorphic lung phantom data provided by the U.S. Food and Drug Administration CDRH/OSEL/DIAM. The FDA collection was partially supported with funding through the intramural programs of NCI and NIBIB.