
Fostering Open Science in Lung Cancer Lesion Sizing with ITK module LSTK

Release 1.00

Xiaoxiao Liu¹, Brian Helba¹, Karthik Krishnan¹, Patrick Reynolds¹, Matthew McCormick¹,
Wes Turner¹, Luis Ibáñez¹, David F. Yankelevitz², Rick Avila¹

June 26, 2012

Kitware Inc., 28 Corporate Dr., Clifton Park, NY
Radiology, Mount Sinai Hospital, New York, NY

Abstract

This document describes the latest efforts in integrating the Lesion Sizing Toolkit (LSTK) into ITK v4 as an external/remote module providing an Open Science dashboard website with a large open image archive of lung cancer CT images for LSTK development and testing.

Contents

1. Brief History of LSTK.....	1
2. Significance and Motivation.....	2
3. ITKv4 Integration.....	2
4. Open Science Dashboard.....	3
5. How to use LSTK module.....	5
6. Summary.....	6

1. Brief History of LSTK

The Lesion Sizing Toolkit (LSTK) is a generic and extensible software library for the segmentation of lesions in medical images [Krishnan10, Krishnan09]. LSTK consists of both a software architecture for developing lesion sizing algorithms as well as a reference algorithm for 3D segmentation of solid lesions

in high resolution CT (≤ 2 mm slice thickness). Kitware, with support from the Optical Society of America (OSA) and the Air Force Research Labs (AFRL), first developed and disseminated LSTK in 2008. LSTK was built using the ITK framework and image processing components. LSTK has participated as the only open source lesion sizing reference algorithm in the NIST “Biochange 2008” lesion sizing benchmark initiative and the “Volcano ‘09 Challenge” benchmark [Volcano09]. With recent funding support from National Library of Medicine (NLM), LSTK has been integrated to Insight Toolkit (ITK) version 4 as an externally distributed ITK module.

2. Significance and Motivation

Volumetric lesion analysis has two important and distinct roles, one for diagnostic purposes and the other to monitor therapeutic response to treatment. In the diagnostic domain, with increasing numbers of small nodules being identified as a result of improvements in CT (and other) imaging technology, measuring change in nodule volume so as to estimate growth rates has gained in importance. In measuring tumor response to treatment, the role of imaging continues to gain importance, as there has been general recognition of the need to find additional clinical trial endpoints beyond survival that can be measured more rapidly, particularly for Phase II trials. In both instances improvements in volumetric measurements will be important. The development of a database that has a wide variety of documented cases that were obtained with a variety of scan parameters will allow for assessment of the degrees of variation in measurement, a component that has yet to be fully evaluated in currently available software. Understanding the trade offs in terms of scan technique such as slice thickness, field of view, and dose will allow for better understanding of the accuracy of algorithms and improve their robustness [Reeves06].

ITK provides a large and rich collection of basic algorithmic components for image processing, segmentation and registration. Most of these components have been designed and implemented to be very generic in nature, and have not been fine tuned to any specific medical application. The LSTK provides a higher-level layer of image analysis components that are built by combining basic ITK components, and by defining pre-sets for many of their numerical parameters. By narrowing the focus of LSTK to lesion sizing applications it becomes possible to select, pre-configure and test, combinations of ITK components. The result is that LSTK provides pre-fabricated modules that are ready to use for lesion sizing clinical applications. LSTK represents a second generation of toolkit components that tailors ITK to specific clinical applications. It lowers the learning curve for developers and facilitates adoption in specific clinical fields. The algorithms, databases, and methods developed in this toolkit may have many additional uses for lesion sizing beyond the lung and will benefit the measurement of other diseases, such as chronic obstructive pulmonary disease (COPD).

3. ITKv4 Integration

One of the major benefits of integrating LSTK to ITK as an externally distributed module is that all additions and improvements made to the lesion sizing capabilities in ITK become immediately available to a wider community of algorithm developers and users. Following the modular architecture of new version 4 of ITK (ITKv4), LSTK is now made available as a remote/external module that can be downloaded as an option when configuring the ITK build tree.

ITK_LesionSizingToolkit (the complete module name of LSTK in ITK) is the first externally distributed ITK module after modularization of the ITKv4. The modules itself share the same modular structure as any internal ITK modules, and the only difference is that the source code of this module is stored separately from the main ITK repository. The old LSTK SVN repository is now replaced by the Git repository:

git://public.kitware.com/LesionSizingToolkit.git

Users can either download (use “[git clone](#)”) the LSTK repository manually into the “External” directory of the ITK source tree to build it, or just to fetch LSTK as a remote module by “fetching” it through CMake during ITK configuration time (Details of how to “fetch” LSTK source code through ITK CMake configuration can be found at Section 5). Details about how to create an ITK externally distributed module can be found [here](#).

Following the new software practice in ITK community, we have also setup a group of dashboard testing machines to test the LSTK module and submit the results to [ITK Dashboards](#) nightly. If you search for "LSTK" in the "Build Name" of the dashboard web you will find the nightly dashboard machines in three differently platforms: Windows, Linux and Mac. In terms of development workflow, we adopted a [Git work flow](#) to commit new code and a [Gerrit review](#) procedure for code reviewing.

4. Open Science Dashboard

In addition to the aforementioned code quality and regression tests and dashboard, algorithmic performance was also quantitatively tested, reported, and analyzed in a publicly accessible and reproducible process.

1120 high-resolution CT datasets containing lung lesions that were obtained during the course of CT guided percutaneous lung biopsies were provided by Dr. David Yankelevitz of Mount Sinai Hospital. The CT datasets contain pre- and intra-procedure scans of 84 patients. As all scans of a given patient were obtained during the procedure, sometimes with varying scan parameters, allowing for a genuine “zero-change” dataset, where the objective lesion size is considered to be constant excluding the possible motion artifact. These zero-change datasets provide important insight into the algorithm’s precision and resiliency.

Those datasets which all contain series with slice thicknesses of 1.25mm, some with additional series with slice thickness of 2.5mm, were manually examined and the physical coordinates of a seed point for each lesion was recorded to run the lesion segmentation algorithm using LSTK.

The original CT datasets, lesion seed coordinates, and LSTK output are all openly available to view or download (Figure 1) on the [Open Science Dashboard](#). The impact of any future changes or refinements to the LSTK algorithm will be empirically tested, by running LSTK with the same input and examining the resultant longitudinal performance. The online platform allows for additional tumor coordinates, new datasets, and even entirely different algorithm classes to be added and shared in the future.

LSTK Results for 2012-05-02 02:07:09

Regression Test Results

Source Code

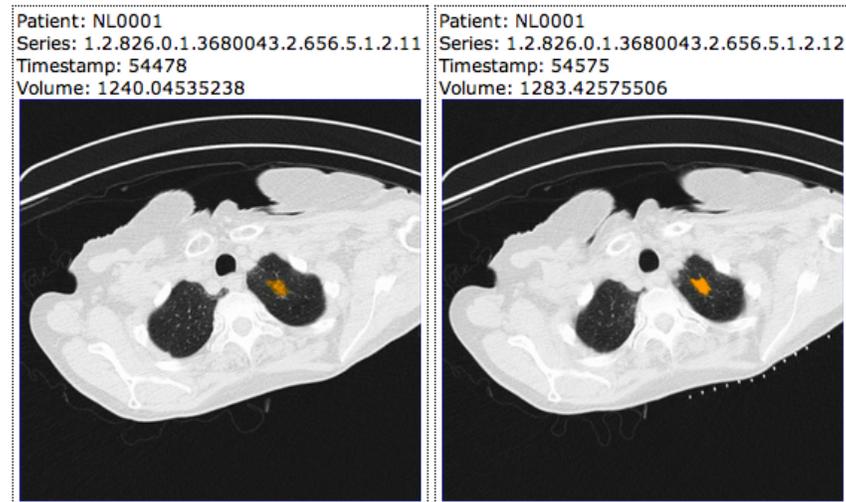


Figure 1. A screenshot of the MiDAS web interface when exploring the LSTK results.

Preliminary multiple-scan tumor segmentation results

We have applied the lesion segmentation algorithm to 29 patients (from the Mt. Sinai database) that have at least two scans with 1.25mm slice thickness. Most datasets have two scans available, several have 3 ~ 6 scans available. We obtained valid segmentation results on 27 of them. The algorithm failed to segment two cases, one of which has a non-solid nodule and the other scan does not capture the entire tumor region due to the limited field of view. Given that the LSTK CT lung reference algorithm was only designed to support segmentation of solid lesions in thin slice CT scans, the inability to handle a non-solid lesion was not considered an issue.

Since no manual segmentation ground truth was available, we used the mean volume value of the patient results as an approximation of the ground truth, and calculated the error coefficient percentage as: $\text{error coefficient} = (\text{volume} - \text{mean volume}) / \text{mean volume} * 100\%$. For all the 27 patients, LSTK segmentation gives an average error coefficient of 24.8 % with a standard deviation of 24.4%. The minimum percentage is 3.5% and the maximum percentage is 91.74%.

Among the 27 patients, 12 of them have inconsistent segmentation results with obvious (mostly more than 50%) volume differences with either under segmentations or over segmentations. Over segmentations tend to happen when the tumor is too close or completely attached to the wall, or when the tumor is connected with a noticeable vascular structure. Under segmentations tend to happen when the tumor region intensities are quite heterogeneous. The other 15 patients had relatively consistent segmentation result among multiple scans. The average error coefficient of the 15 patients is 8.09% with a standard deviation of 6.41%. The minimum percentage is 0.35% and the maximum percentage is 23.86%.

One factor that causes the inconsistency among multiple scans of the same patient is the image quality difference of the scans. For the same patient, the first large volume scan (more than 200 slices) and the

follow-up smaller (6~12 slices) needle placement scans might have different contrast and noise level due to differences in configuration parameters of the imaging devices. The tumor tissue appearance therefore may be different. When the tumor texture is homogeneous (with tight intensity distribution), the segmentation results are expected to be more accurate. The respiratory motion also contributes to the inconsistency. We have observed that in some cases the shape of the large tumor are either squeezed or extended noticeably in inferior-superior direction between multiple scans. More detailed error analysis and validations will be carried out on more datasets in the future.

5. How to use LSTK module

Here is a brief list for potential users of LSTK module.

1) Software prerequisites

You need to have the following software ready in your system:

- ITK (4.0 +) source
- VTK (5.4+), installed
- CMake (2.8+), installed
- Git, installed
- Common essential C++ development environment

2) How to build LSTK

There are two ways of obtaining the source code of LSTK into ITK source tree:

- Fetch LSTK source code as a remote module in ITK (preferred, easier)

When running CMake configuration to build ITK tree, enter the advanced mode in CMake GUI and turn on “Fetch_LesionSizingToolkit” CMake variable. After configuration, you should be able to find the source code of LSTK has been downloaded to your ITK source tree.

- Manually download LSTK source code

Go to the source tree of ITK, and inside the External directory, run “git clone [git://public.kitware.com/LesionSizingToolkit.git](https://public.kitware.com/LesionSizingToolkit.git)” to download the module.

To build LSTK as an ITK Module, turn on *Module_LesionSizingToolkit* CMake variable when configuring ITK build tree with CMake. All LSTK depending modules will be automatically turned on. To use the visualization functionality in LSTK, *LSTK_USE_VTK* variable needs to be turned on as well (*VTK_DIR* needs to be specified, if not automatically detected).

3) Running Lesion Segmentation on CT images

LSTK provides an executable program called *LesionSegmentation* to run the lesion segmentation algorithms on CT images. An example command line of running lesion segmentation on an CT image is given as follows with a seed point’s physical coordinates provided by the user:

```
./bin/LesionSegmentation -InputImage <./inputImage.mha> -OutputImage  
<./outputTumorImage.mha> -Seeds 3 326.2 231.9 172.6 -MaximumRadius 20 -OutputROI  
<./outputROI.mhd> -OutputMesh <./output.stl> -Visualize 1 -Screenshot <./screenShotTumor>
```



Figure 2. Lesion Segmentation result screen shots. From left to right, orthogonal views of sagittal, coronal and axial views of the segmented tumor volume overlapped with the corresponding input CT image slices are shown. The reported volume size of the detected lesion is 1264.67 mm³.

6. Summary

In this paper we briefly introduced the latest efforts supported by NLM in integrating the Lesion Sizing Toolkit (LSTK) into ITK v4 and providing an Open Science Dashboard with a large open image archive of lung cancer CT images for LSTK development and testing. A preliminary study on multiple-scan lung tumor LSTK segmentation results were presented and brief introductions on how to use LSTK were given. To keep up with latest or future developments in LSTK and the Open Science Dashboard, please visit the [LSTK wiki page](#).

References

[Krishnan10] Karthik Krishnan, Luis Ibanez, Wesley D. Turner, Julien Jomier, and Ricardo S. Avila, *An open-source toolkit for the volumetric measurement of CT lung lesions*, Optics Express, Vol. 18, Issue 14, pp. 15256-15266, 2010

[Krishnan09] Krishnan K., Ibanez L., Turner W.D., Avila R.S., *Algorithms, architecture, validation of an open source toolkit for segmenting CT lung lesions*, Second International Workshop on Pulmonary Image Analysis, London, 2009.

[Volcano09] A. P. Reeves, A. C. Jirapatnakul, A. M. Biancardi, T. V. Apanasovich, C. Schaefer, J. J. Bowden, M. Kietzmann, R. Korn, M. Dillmann, Q. Li, J. Wang, J. H. Moltz, J. Kuhnigk, T. Hayashi, X. Zhou, H. Fujita, T. Duindam, B. van Ginneken, R. Avila, J. P. Ko, K. Melamud, H. Rusinek, R. Wiemker,

G. Soza, C. Tietjen, M. Thorn, M. F. McNitt-Gray, Y. Valenciaga, M. Khatonabadi, Y. Kawata, and N. Niki, *The VOLCANO'09 challenge: Preliminary results*, Second International Workshop of Pulmonary Image Analysis, pp. 353-364, Sept. 2009.

[Reeves06] Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ., *On measuring the change in size of pulmonary nodules*, IEEE Trans Med Imaging. 25 (4):435-50, 2006