Standardized representation of the TCIA LIDC-IDRI annotations using DICOM

Data Citation

Andrey Fedorov, Matthew Hancock, David Clunie, Mathias Brockhausen, Jonathan Bona, Justin Kirby, John Freymann, Hugo Aerts, Ron Kikinis, Fred Prior. Standardized representation of the TCIA LIDC-IDRI annotations using DICOM. (2018) The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.2018.h7umfurq

Description

This dataset contains standardized DICOM representation of the annotations and characterizations collected by the LIDC/IDRI initiative, originally stored in XML and available in the TCIA LIDC-IDRI collection. Only the nodules that were deemed to be greater or equal to 3 mm in the largest planar dimensions have been annotated and characterized by the expert radiologists performing the annotations. Only those nodules are included in the present dataset.

Conversion was enabled by the *pylidc* library (https://pylidc.github.io/) (parsing of XML, volumetric reconstruction of the nodule annotations, clustering of the annotations belonging to the same nodule, calculation of the volume, surface area and largest diameter of the nodules) and the *dcmqi* library (https://github.com/qiicr/dcmqi) (storing of the annotations into DICOM Segmentation objects, and storing of the characterizations and measurements into DICOM Structured Reporting objects). The script used for the conversion is available at https://github.com/qiicr/lidc2dicom. The details on the process of the conversion and the usage of the resulting objects are available in the preprint:

Preprint Citation

Fedorov A, Hancock M, Clunie D, Brochhausen M, Bona J, Kirby J, Freymann J, Pieper S, Aerts H, Kikinis R, Prior F. 2018. Standardized representation of the LIDC annotations using DICOM. PeerJ Preprints 6:e27378v1 https://doi.org/10.7287/peerj.preprints.27378

Please also cite the following original datasets and manuscript when citing this dataset:

Data Citation

Armato III, Samuel G., McLennan, Geoffrey, Bidaut, Luc, McNitt-Gray, Michael F., Meyer, Charles R., Reeves, Anthony P., … Clarke, Laurence P. (2015). Data From LIDC-IDRI. The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX

Publication Citation

Armato SG III, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, Zhao B, Aberle DR, Henschke CI, Hoffman EA, Kazerooni EA, MacMahon H, van Beek EJR, Yankelevitz D, et al.: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Medical Physics, 38: 915--931, 2011. (paper)

Download

- Version 2: http://bit.ly/lidc-dicom-v2
  - DICOM SEG objects no longer encode empty slices to reduce object size
  - the coded terms used to describe the nodule annotations now use fewer non-standard (99QIICR) codes
  - SegmentLabel attribute is populated in the DICOM SEG objects to list nodule annotation name instead of "Nodule", to help with readability
    for the user
- Version 1: http://bit.ly/lidc-dicom
  - Note: Version 1 of this dataset is currently located in a shared Google Drive folder while undergoing verification. When testing is complete the Google Drive folder will be replaced by a different link to the final dataset. If you identify any issues with the data please report them to the TCIA Helpdesk.