NSCLC-Radiomics

Reproduced from https://wiki.cancerimagingarchive.net

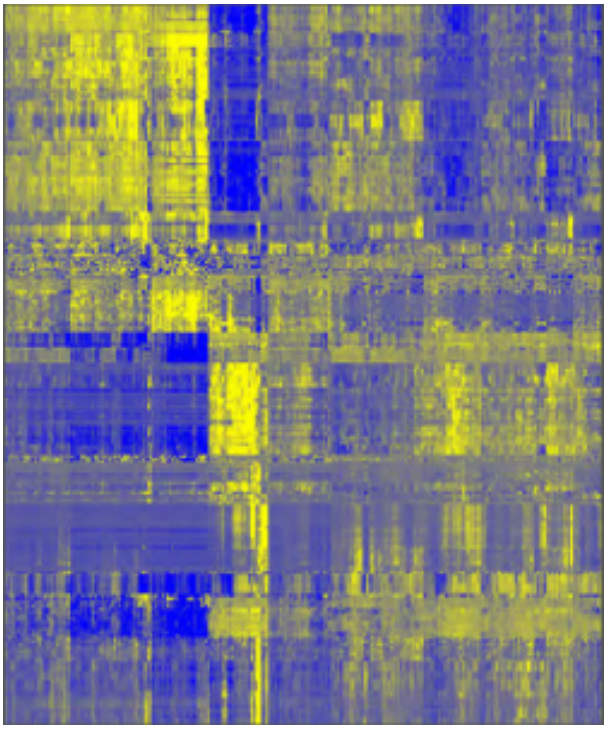
Redirection Notice

This page will redirect to https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI in about 10 seconds.

Reproduced from https://wiki.cancerimagingarchive.net

Summary

This collection contains images from 422 non-small cell lung cancer (NSCLC) patients. For these patients pretreatment CT scans, manual delineation by a radiation oncologist of the 3D volume of the gross tumor volume and clinical outcome data are available. This dataset refers to the Lung1 dataset of the study published in Nature Communications.

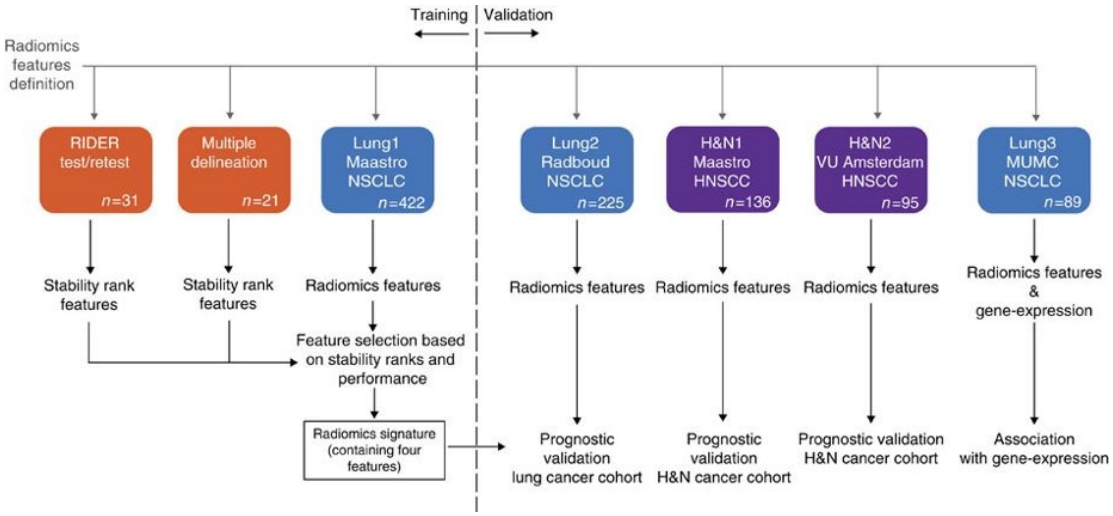


In short, this publication applies a radiomic approach to computed tomography data of 1,019 patients with lung or head-and-neck cancer. Radiomics refers to the comprehensive quantification of tumour phenotypes by applying a large number of quantitative image features. In present analysis 440 features quantifying tumour image intensity, shape and texture, were extracted. We found that a large number of radiomic features have prognostic power in independent data sets, many of which were not identified as significant before. Radiogenomics analysis revealed that a prognostic radiomic signature, capturing intra-tumour heterogeneity, was associated with underlying gene-expression patterns. These data suggest that radiomics identifies a general prognostic phenotype existing in both lung and head-and-neck cancer. This may have a clinical impact as imaging is routinely used in clinical practice, providing an unprecedented opportunity to improve decision-support in cancer treatment at low cost. The DICOM Radiotherapy Structure Sets (RTSTRUCT) and DICOM Segmentation (SEG) files in this data contain a manual delineation by a radiation oncologist of the 3D volume of the primary gross tumor volume ("GTV-1") and selected anatomical structures (i.e., lung, heart and esophagus). Of note, DICOM SEG objects contain a subset of annotations available in RTSTRUCT.

The dataset described here (Lung1) was used to build a prognostic radiomic signature. The Lung3 dataset used to investigate the association of radiomic imaging features with gene-expression profiles consisting of 89 NSCLC CT scans with outcome data can be found here: NSCLC-Radiomics-Genomics.

Other data sets in the Cancer Imaging Archive that were used in the same study published in Nature Communications: Head-Neck-Radiomics-HN1, NSCLC-Radiomics-Interobserver1, RIDER Lung CT Segmentation Labels from: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach (RIDER-LungCT-Seg).

Reproduced from https://wiki.cancerimagingarchive.net



For scientific or other inquiries about this dataset, please contact the TCIA Helpdesk.

Acknowledgements

We would like to acknowledge the individuals and institutions that have provided data for this collection:

- Leonard Wee, MAASTRO (Dept of Radiotherapy), Maastricht University Medical Centre+, Maastricht, Limburg, The Netherlands.
- Dirk de Ruysscher, MAASTRO (Dept of Radiotherapy), Maastricht University Medical Centre+, Maastricht, Limburg, The Netherlands.
- Andre Dekker, MAASTRO (Dept of Radiotherapy), Maastricht University Medical Centre+, Maastricht, Limburg, The Netherlands.
- Hugo Aerts, Computational Imaging and Bioinformatic Laboratory, Dana-Farber Cancer Institute & Harvard Medical School, Boston, Massachusetts, USA.
- Harmonization of the components of this dataset, including into standard DICOM representation, was supported in part by the NCI Imaging Data Commons consortium. NCI Imaging Data Commons consortium is supported by the contract number 19X037Q from Leidos Biomedical Research under Task Order HHSN26100071 from NCI.

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Images, Segmentations, and Radiation Therapy Structures (DICOM, 33 GB)	*Download* *Search* (Download requires the NBIA Data Retriever)	CC BY-NC 3.0
Lung1 clinical (CSV, 23 kB)	*Download*	CC BY-NC 3.0

Reproduced from https://wiki.cancerimagingarchive.net

Click the Versions tab for more info about data releases.

Additional Resources for this Dataset

The NCI Cancer Research Data Commons (CRDC) provides access to additional data and a cloud-based data science infrastructure that connects data sets with analytics tools to allow users to share, integrate, analyze, and visualize cancer research data.

- Imaging Data Commons (IDC) (Imaging Data)

Third Party Analyses of this Dataset

TCIA encourages the community to publish your analyses of our datasets. Below is a list of such third party analyses published using this Collection:

- Thoracic Volume and Pleural Effusion Segmentations in Diseased Lungs for Benchmarking Chest CT Processing Pipelines (PleThora)

Detailed Description

Detailed Description

Collection Statistics	
Modalities	CT, RTSTRUCT, SEG
Number of Participants	422
Number of Studies	422
Number of Series	1265
Number of Images	52073
Image Size (GB)	33

Radiation Oncologist Tumor Segmentations

The DICOM Radiotherapy Structure Sets (RTSTRUCT) and DICOM Segmentation (SEG) files in this data contain a manual delineation by a radiation oncologist of the 3D volume of the primary gross tumor volume ("GTV-1") and selected anatomical structures (i.e., lung, heart and esophagus). Of note, DICOM SEG objects contain a subset of annotations available in RTSTRUCT.

For viewing the annotations the authors recommend 3D Slicer that can be used to view both RTSTRUCT and SEG annotations (make sure you install the SlicerRT and QuantitativeReporting extensions first!). Visualization of the DICOM annotations is also supported by the OHIF Viewer.

Other tools include:

- Dicompyler is an open source, cross-platform DICOM RT viewer.
- The Radiotherapy DICOM toolkit, which may also be useful for working with this data.

Reproduced from https://wiki.cancerimagingarchive.net

- dcmqi and Plastimatch libraries can be used to support conversion of the DICOM SEG and RTSTRUCT representations, respectively, into the popular research formats, such as NIfTI or NRRD.

Clinical Data

Corresponding clinical data can be found here: Lung1.clinical.csv.

Please note that survival time is measured in days from start of treatment. DICOM patients names are identical in TCIA and clinical data file.

The deadstatus.event follows the standard epidemiological / biomedical definition for time to event survival modelling whereby :

1 == “death has occurred at the time interval value stated on the next column”, hence

0 == “death has NOT occurred up until the time interval value stated in the next column, and is therefore right-censored”.

Note:

Some CT (LUNG1-014 , LUNG1-021 , LUNG1-085) are missing some slices from the complete volume, unfortunately not detected until recently. The skipped CT slice did not cut through the Gross Tumour Volume which was the principal focus of this Collection. We understand the user may have a different principal focus for re-using this collection, but perhaps some other workaround like interpolation may be timely and feasible for those affected volumes.

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the TCIA Data Usage Policy and Restrictions. Attribution should include references to the following citations:

Data Citation

Aerts, H. J. W. L., Wee, L., Rios Velazquez, E., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., Lambin, P. (2014). **Data From NSCLC-Radiomics (version 4)** [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI

Publication Citation

Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., Lambin, P. (2014, June 3). **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.** Nature Communications. Nature Publishing Group. https://doi.org/10.1038/ncomms5006 (link)

TCIA Citation

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. https://doi.org/10.1007/s10278-013-9622-7

Reproduced from https://wiki.cancerimagingarchive.net

Questions may be directed to help@cancerimagingarchive.net.

Other Publications Using This Data

TCIA maintains a list of publications which leverage our data. If you have a publication you'd like to add, please contact the TCIA Helpdesk.

Versions

Version 4 (Current): Updated 2020/10/22

Data Type	Download all or Query/Filter
Images (DICOM, 33 GB)	*Download* *Search* (Download requires the NBIA Data Retriever)
Lung1 clinical (CSV)	*Download*

- RTSTRUCT and SEG study instance UID changed to match study instance uid with associated CT image.
- Added missing structures in SEG files to match associated RTSTRUCTs.
- Patient Id copied to Patient Name in CT images (for consistency).
- Added 1 missing image for LUNG1-246.

Version 3: Updated 2019/10/23

Data Type	Download all or Query/Filter
Images (DICOM, 29GB)	*Download* (Download requires the NBIA Data Retriever)
Lung1 clinical (CSV)	*Download*

- Re-checked and updated the RTSTRUCT files to amend issues in the previous submission due to missing RTSTRUCTS or regions of interest that were not vertically aligned with the patient image.
- In 4 cases (LUNG1-083,LUNG1-095,LUNG1-137,LUNG1-246) re-submitted the correct CT images.
- The regions of interest now include the primary lung tumor labelled as “GTV-1”, as well as organs at risk.
- For one case (LUNG1-128) the subject does not have GTV-1 because it was actually a post-operative case; we retained the CT scan here for completeness.
- Added DICOM SEGMENTATION objects to the collection, which makes it easier to search and retrieve the GTV-1 binary mask for re-use in quantitative imaging research.
- Clinical data updated as follow-up time has been extended.

Version 2: Updated 2016/05/31

Data Type	Download all or Query/Filter
Images (DICOM, 25GB)	not available, see version 3
Lung1 clinical (CSV)	*Download*

Reproduced from https://wiki.cancerimagingarchive.net

Added 318 RTSTRUCT files for existing subject imaging data. (NOTE: version 2 removed as RTSTRUCTs or regions of interest were not vertically aligned with patient images. See version 3 for updated files).

Version 1: Updated 2014/07/02

Data Type	Download all or Query/Filter
Images (DICOM, 25GB)	*Download* (Download requires the NBIA Data Retriever)
Lung1 clinical (CSV)	*Download*

Reproduced from https://wiki.cancerimagingarchive.net