JULY 19–22
2026 VANCOUVER, BC
JOINT AAPM | COMP MEETING

A Cancer Genomics Cloud project for Radiological and Radiotherapy Image Analyses

Aditya P. Apte, Ph.D; Aditi Iyer, M.S.; Eve LoCastro, M.S.; Joseph O. Deasy, PhD
Memorial Sloan Kettering Cancer Center, New York, NY

ABSTRACT

Purpose/Objective
A software resource in the form of Apps distributed via our public project on the Cancer Genomics Cloud (CGC) platform for reproducible, out-of-the-box application of analysis workflows for pre-trained radiological and radiotherapy AI models.

Development and Validation Methods
The framework comprises of two main components: the Python-native Computational Environment for Radiological Research (pyCERR) for radiological image processing, and the Cancer Genomics Cloud (CGC), a National Cancer Institute–funded resource that supports cloud-hosted collaborative projects, with tools for managing collaborators and compute-hardware from providers such as Amazon Web Services and Google Cloud Platform. pyCERR is open-source (GNU GPL) and distributed at **https://github.com/cerr/pycerr**; it provides user-friendly data structures, simplified access to metadata across radiological imaging modalities, and tools for data transformation, AI-based image segmentation, radiomics, dynamic contrast-enhanced MRI analysis, radiotherapy dose–volume histogram (DVH) feature extraction, and normal-tissue complication and tumor-control models. Radiomics features comply with the Image Biomarker Standardization Initiative (IBSI) standards and are validated against published reference values. Unit tests covering data input/output, radiomics, DVH features, and visualization run on every code update. Workflows are containerized for reproducible, identical execution on the CGC and on local hardware. The CGC enables AI inference from radiological images to be combined with other data types such as genomics and proteomics. Our public CGC project (https://cgc.sbgenomics.com/u/sevenbridges/pycerr-radiological-radiotherapy-imageanalysis), accessible with free registration, currently includes apps for organ-at-risk (OAR) and lung-nodule segmentation based on published models trained in-house and clinically validated. These apps can be run interactively from a web browser or programmatically via a REST API with client support for multiple programming languages. Notebooks demonstrating the CGC API and pyCERR capabilities are provided at https://github.com/cerr/pycerr-notebooks.

Potential Applications
The framework facilitates end-to-end radiological image analysis and reproducible research, including data access from public and private repositories, training and inference with AI models, and utilities for data management, visualization, and simplified access to image metadata.

CONTACT

Aditya P. Apte, Ph.D.
Memorial Sloan Kettering Cancer Center
New York, NY
aptea@mskcc.org

INTRODUCTION

Simplified application of analysis workflows

- The framework allows for deployment of radiological image analyses by pre-configuring software and hardware resources in the form of Apps on CGC and can be readily run as tasks via the REST API with support for various programming languages.
- Users can readily combine imaging and genomics data available on CGC.
- Pre-trained AI models can be run readily allowing for analysis on different datasets. Clinically validated AI segmentation of 70+ organs is available.

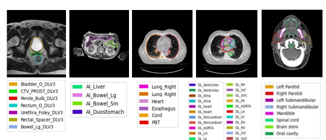
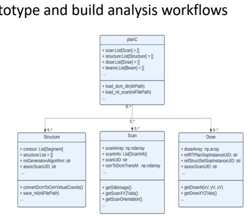


Table below lists AI models for Image Segmentation that are clinically used at our institution after prospective validation and usable with pyCERR. Jupyter Notebooks are distributed at https://github.com/cerr/pyCERR-Notebooks/tree/main/03_autosegmentation

Image Modality	Site	Network architecture	Outputs	Notebook
Planning CT Axial	Lung	Multiple resolution residual network [1]	Segmentations: Right lung, left lung, heart, esophagus, cord, PBT	autosegment_CT_Lung_OARs.ipynb
Planning CT Axial	Lung	DeepLabV3+ [2]	Segmentations: aorta, SVC, IVC, PA, LA, LV, RA, RV, atria, ventricles, and pericardium	autosegment_CT_Heart_OARs.ipynb
Planning CT Axial	Lung	Self-distilled Masked Image Transformer [7]	Segmentation: Lung tumor	autosegment_CT_Lung_GTV_SMIT.ipynb
T2w Axial MRI	Prostate	DeepLabV3+ [10]	Segmentations: CTV, bladder, penile bulb, rectum, urethra foley, rectal spacer, large bowel	autosegment_MR_Prostate_OARs.ipynb
T2w Axial MRI (longitudinal)	Pancreas	ProRSeg[11]	Segmentations: Liver, large bowel , small bowel, duo-stomach and Deformable vector field (first to last scan)	auto_register_segment_MR_Pancreas_OARs.ipynb
Planning CT Axial	Head & Neck	Self-Attention[13], DeepLabV3+[14]	Segmentations: Left parotid, right parotid, left submandibular gland, right submandibular gland, mandible, spinal cord, brain stem, oral cavity	autosegment_CT_HeadAndNeck_OARs.ipynb

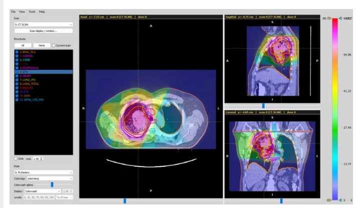
pyCERR DATA STRUCTURE

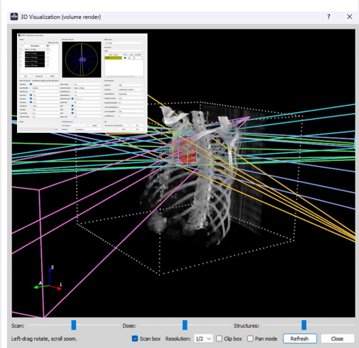
- The plan container class planC houses various medical imaging data types.
- Metadata from DICOM and research-friendly modalities such as Nifti are well-organized in planC by retaining temporal and spatial associations.
- This data organization enables users to efficiently prototype and build analysis workflows

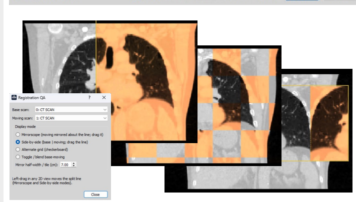


pyCERR GUI

Python-native Qt-based viewer supporting -
→ Loading and browsing multiple image modalities.
→ 3D visualization
→ Registration QA
→ Contouring
→ IMRTP beams definition and QIB dose calculation
→ Programmatic control of GUI for batch runs







RESOURCES

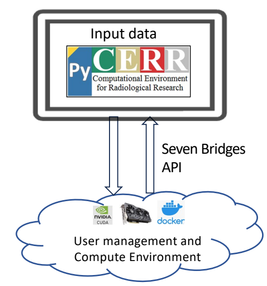
Resource	Location
Codebase	https://github.com/cerr/pyCERR (GNU GPL copyright)
Documentation	https://pycerr.readthedocs.io
Example Notebooks	https://github.com/cerr/pyCERR-Notebook
CGC Apps	https://cgc.sbgenomics.com/u/sevenbridges/pycerr-radiological-radiotherapy-imageanalysis
Model Weights	https://github.com/<modelName>/model.txt where <modelName> is one of the following: CT_cardiac_structures_deeplab CT_LungOAR_incrMRRN MR_Prostate_Deeplab CT_Lung_SMIT MRI_Pancreas_Fullshot_AnatomicCtxShape CT_HeadAndNeck_OARs CT_HN_SMIT CT_HeartSubstruct_SMIT CT_WHOLEBODY_SMITplus MR_Rectum_GTV_SMIT MR_HN_Nodule_SMIT
Sample Datasets	https://github.com/cerr/pyCERR/tree/main/cerr/datasets
Unit tests	https://github.com/cerr/pyCERR/tree/main/tests
Python Package Index	https://pypi.org/project/pycerr

CGC APPS

Apps from our public CGC project can be cloned by users to run out of the box analysis.

Cancer Genomics Cloud (CGC) provides integrated bioinformatics tools, automatic logging, versioning, and access-controlled sharing, facilitating reproducibility and collaboration.

Data can be persistently hosted in linked cloud buckets from providers such as AWS and GCP or accessed from external cloud repositories, including public sources such as TCIA and IDC as well as proprietary databases such as XNAT, BOX.



(a) Apps and workflows can be run from the command-line on client computers by using the API. The API supports various programming languages including Python, Java, C# etc.

(b) App execution via the web interface with user-configurable inputs, outputs, and runtime settings.

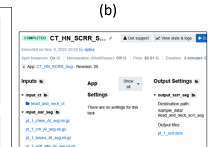
(a)

```python
import sevenbridges as sbg
from getpass import getpass

url = 'https://cgc-api.sbgenomics.com/v2'
proj = '/aptea/pycerr-analyses'
# Authenticate via authentication token
authtoken = getpass()
api = sbg.Api(url=url, token=authtoken)
```

```python
task = api.tasks.create(name=taskName,
                        project=proj,
                        app=app,
                        inputs=appInputs,
                        run=True)
```

(b)



CONCLUSIONS

CGC and pyCERR together facilitate end-to-end radiological image analysis and reproducible research, including data access from cloud sources, training or inferring from pre-configured AI models, utilities for data management, visualization, and simplified access to image metadata.

ACKNOWLEDGEMENTS

Funding sources:
U54 CA274291 – ROBIN Network
P30 CA008748 – Cancer Center Core
Computation credits from the CGC