

Standardized Environment for Radiomics Analysis (SERA)

Saeed Ashrafinia, Arman Rahmim

Introduction

The Standardized Environment for Radiomics Analysis (SERA) is a Matlab®-based framework developed at Johns Hopkins University that calculates radiomic features based on guidelines from [the Image Biomarker Standardization Initiative](#) (IBSI). SERA is capable of processing images from various clinical imaging modalities such as CT, MRI, PET and SPECT. Radiomic features calculated with SERA are standardized and in compliance with IBSI, which ensures their reproducibility.

Radiomic Features

SERA calculates 487 IBSI-standardized features, including: 79 first-order features (morphology, statistical, histogram and intensity-histogram features), 272 higher-order 2D features, and 136 3D features. In addition, it also calculates 10 moment invariant features, that are not included in IBSI. Different subsets of features can be selected, such as the default of 215 features (first-order + higher-order 3D).

Question, Feedback, Collaborations

If you have any questions regarding SERA, want to provide feedback, or to propose a collaboration, please contact s.ashrafinia AT gmail.com.

Reference

➤ Please cite the following reference if you publish results with help from SERA:

Ashrafinia, Saeed. "Quantitative Nuclear Medicine Imaging using Advanced Image Reconstruction and Radiomics", Ph.D. Dissertation, Johns Hopkins University, 2019

Table: Radiomic features SERA calculates in each feature family. Different 2D, 2.5D and 3D configurations are explained in detail in the IBSI guideline available at <https://arxiv.org/pdf/1612.07003.pdf> *Moment Invariant features are not included in IBSI, but are provided additionally. Users can set to return only a selected subset of these feature, as explained below.

| Feature Family | Subtypes | Number of Features |
|--|-----------------------|--------------------|
| Morphology | - | 29 |
| Local Intensity | - | 2 |
| Intensity-based Statistics | - | 18 |
| Intensity Histogram | - | 23 |
| Intensity-Volume Histogram | - | 7 |
| Gray Level Co-occurrence Matrix (GLCM) | 2D Averaged | 25 |
| | 2D Slice-Merged | 25 |
| | 2.5D Direction Merged | 25 |
| | 2.5 D All Merged | 25 |
| | 3D Averaged | 25 |
| | 3D Merged | 25 |
| Gray Level Run Length Matrix (GLRLM) | 2D Averaged | 16 |
| | 2D Slice-Merged | 16 |
| | 2.5D Direction Merged | 16 |
| | 2.5 D All Merged | 16 |
| | 3D Averaged | 16 |
| | 3D Merged | 16 |
| Gray Level Size Zone Matrix (GLZSM) | 2D | 16 |
| | 2.5 D | 16 |
| | 3D | 16 |
| Gray Level Distance Zone Matrix (GLDZM) | 2D | 16 |
| | 2.5 D | 16 |
| | 3D | 16 |
| Neighborhood Grey Tone Difference Matrix (NGTDM) | 2D | 5 |
| | 2.5 D | 5 |
| | 3D | 5 |
| Neighboring Grey Level Dependence Matrix (NGLDM) | 2D | 17 |
| | 2.5 D | 17 |
| | 3D | 17 |
| Moment Invariants* | - | 10 |
| Total | | 497 |

INSTRUCTIONS:

1) Feature Evaluation Settings

SERA has options to set and modify all parameters that have been defined or used in the IBSI guideline. The image preparation setting includes (for detailed information please refer to [IBSI documentation](#)):

- Resampling and interpolation:
 - resample to 2D and 3D isotropic voxel sizes; interpolation algorithm used in resampling image and ROI (nearest/linear/cubic); partial volume threshold (mostly used for CT HU).
- Discretization:
 - bin size, discretization type (fixed bin size/fixed bin numbers), discretization algorithm (uniform/Lloyd)
- Other:
 - grey-level rounding, image re-segmentation (range re-segmentation, outliers re-segmentation)

2) Loading datasets into SERA:

Images and their corresponding ROIs

SERA can handle images with multiple ROIs. First, make sure that each patient has its image and ROI saved in a separate folder. Inside each folder, there should be two “.mat” files: one containing a **3D double matrix** variable that has the image, and one containing a variable that has the ROI. Each single ROI should be prepared or converted to a **3D matrix of 0s and 1s** to be able to work in SERA.

While the patient folder names should be different, it is required that all image “.mat” files and ROI “.mat” files inside those folders have an identical name; e.g. PT001\img.mat, PT001\ROI.mat, PT002\img.mat, PT002\ROI.mat, and so on. Note that “img.mat” and “ROI.mat” are the same inside each folder. Also, the variable names that are saved inside these two files should be identical. For example, all “img.mat” files in the above example can contain a 3D matrix variable named “image_volume”, and all “ROI.mat”s can contain a 3D matrix – a single binary 3D variable – named “contours”.

Voxel size information

In addition to the images and their corresponding ROI, SERA requires an Excel or CSV spreadsheet that contains voxel sizes (voxel size in x/y dimension, voxel size in z dimension) for every patient. In this spreadsheet, the first column refers to patient ID and has to be the same as the folder containing patient’s image and ROI. The second and the third columns include voxel size in x(=y) and z dimensions, respectively. Rows of this spreadsheet should be alphabetically sorted, and be the same order as the patient folder names in their parent directory. The path to this file should be provided at the beginning of the code.

3) Running SERA

To set up the parameters and run SERA, open “SERA_Main.m”. Detailed comments are provided in front of every variable that helps users to select the right configuration. Please refer to IBSI documentation for more detail elaboration of these variables at <https://arxiv.org/pdf/1612.07003.pdf>.

Parameters that need modifications

They are **four sections** in “SERA_Main.m” that **need to be configured**:

```
%% Path to dataset and voxel size info
```

The path to the patients’ dataset containing image and ROI folder as stated above is given in this section. You also can name your dataset. Also, the path to an Excel or CSV file containing voxel size information for every patient should be given here. This spreadsheet should contain 2 columns: x(=y) dimension voxel

thickness, and z dimension slice thickness. They should be stored in the same sequence as the patients' directory names sorted alphabetically.

%% Radiomics Framework Settings

Workflow parameters are configured under this section. Please refer to comments in front of each variable, as well as IBSI documentation for more information.

%% Verifying Image and ROI variable names

SERA automatically loads the image and ROI from the file with filenames specified at the end of "Radiomics Framework Settings" section. But the data inside these files might have been saved under a different variable name. Users should specify the variable name that they have stored image and ROI inside their corresponding files in this section. In the released version, these two variables are set as "vol_vals" for image and "total" for ROIs. Please modify these two names to your specific variable names. If the variable cannot be found, it returns an error.

%% Loading ROIs and Decrypting contours

SERA is capable of handling more than one ROI per image, but it is up to the user to write few lines of code to properly import multiple ROIs into the framework. Due to diversity of methods to save ROIs, this has left to the user. To run SERA on multiple ROIs per image, **it is required that all ROIs are stored under a 4D matrix of 0s and 1s named "ROIsCollection4D"**. The proper conversion should be applied to create ROIsCollection4D—for example, if ROIs are saved under different variable names inside the ROI ".mat" file, or are saved under a structured variable, etc. The released version of SERA assumes just one 3D matrix of 0s and 1s inside the ROI ".mat" file.

Output Features

Users can choose subset of 497 features SERA can calculate to return in the output. Currently there are five options to choose from, and they can be set using variable "Feats2out" in the main code:

- 1- All IBSI features (487 features)
- 2- 1st-order+all 3D features, (215 features)
- 3- 1st-order+all 2D/2.5D features, (351 features)
- 4- 1st-order + selected 2D + all 3D features, (351 features)
- 5- all features + moment invariant (497 features)

Selecting features for different imaging modalities

SERA can be used for any 2D or 3D medical imaging modalities such as CT, MRI, PET and SPECT. However, there are recommended settings for specific imaging modalities. For instance, as stated in the IBSI document, 2D feature calculation is more relevant in CT and MRI where slice thickness (voxel size in z dimension) is very different from the voxel size in x and y dimensions. As a result, users are encouraged to include 2D features in their calculation for CT images, or for PET images with anisotropic voxel sizes (z dimension voxel size \neq x,y dimension voxel size). This means choosing options 1, 3, 4 or 5 for "Feats2out" variable.

For raw images with isotropic voxel dimensions (x=y=z voxel dimension), 2D features may not capture any significant pattern, thus suggesting the option 2 to exclude 2D and 2.5D features.

List of output features

The detailed list of SERA output features can be found in “SERA Features Name and Tags.xlsx”. Each sheet contains the feature sets for each “Feats2out” setting.

Test Dataset

Preloaded data

The released version of SERA is preloaded with the IBSI benchmark dataset, including a CT image and an ROI. The pre-configured settings for SERA is based on “config D” of IBSI benchmark test. The IBSI benchmark configurations are available at “configuration.pdf”.

Verifying compliance of SERA with IBSI

The released version also includes a spreadsheet “SERA features compliance check with IBSI.xlsx” that contains all the benchmark values for IBSI configurations, the radiomic features calculated by SERA, and whether they match IBSI or not. If you made any changes to SERA inner functions, make sure you verify the results with this spreadsheet.

If you run SERA right from the batch, it is going to calculate features based on config D as mentioned above. Then if you copy the values of “FullFeatures” and paste them into column “your_results” under “config D” sheet of “SERA features compliance check with IBSI.xlsx”, you will get all matches, except one.