2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A new machine learning based user-friendly software platform for automatic radiomics modeling and analysis

Zhiyong Zhou, Xusheng Qian, Jisu Hu, Jianbing Zhu, Chen Geng, Yakang Dai*

***Abstract*—Supervised machine learning methods are usually used to build a custom model for disease diagnosis and auxiliary prognosis in radiomics studies. A classical machine learning pipeline involves a series of steps and multiple algorithms, which leads to a great challenge to find an appropriate combination of algorithms and an optimal hyper-parameter set for radiomics model building. We developed a freely available software package for radiomics model building. It can be used to lesion labeling, feature extraction, feature selection, classifier training and statistic result visualization. This software provides a user-friendly graphic interface and flexible IOs for radiologists and researchers to automatically develop radiomics models. Moreover, this software can extract features from corresponding lesion regions in multi-modality images, which is labeled by semi-automatic or full-automatic segmentation algorithms. It is designed in a loosely coupled architecture, programmed with Qt, VTK, and Python. In order to evaluate the availability and effectiveness of the software, we utilized it to build a CT-based radiomics model containing peritumoral features for malignancy grading of cell renal cell carcinoma. The final model got a good performance of grading study with AUC=0.848 on independent validation dataset.**

***Clinical Relevance*—the developed provides convenient and powerful toolboxes to build radiomics models for radiologists and researchers on clinical studies.**

I. INTRODUCTION

Radiomics methods translate medical images into high-dimension features, which are used as noninvasive biomarkers for disease diagnosis, auxiliary prognosis [1-3] and prediction of pathological response [4, 5]. The features utilized in Radiomics can be categorized into shape, intensity (first order histogram) and texture (high order) feature. Moreover, higher order features and more expressive features are created by wavelet, 3D local binary pattern or other image filters. All these features are compliance with definitions of Imaging Biomarker Standardization Initiative (IBSI) [6]. Comparison to the observations of shape and intensity features by radiologists, radiomics extracts high order texture features, which reveals the hidden information in medical images and improves the ability of radiologists and computer-aided diagnosis systems. Radiomics model building is actually a supervised machine learning. The workflow of radiomics modeling includes image preparing, lesion labeling, feature extraction, feature selection, classifier training and evaluation.

Unfortunately, it is a great challenge to build an optimal radiomics model due to there are many candidate algorithms and their combinations in the workflow of radiomics studies. It requires expertise in machine learning and is a tedious work for radiologist. We currently developed a freely available software package, for radiomics studies. It builds a custom radiomics model by a pipeline with image preparing, lesion labeling, feature extraction, feature selection, and classifier training for clinical predict tasks. Further, it provides various data visualization ways and result statistics to compare the performances of the established models on the validation cohort, which helps the researchers focus on the interpretation of the radiomics signature in the final built model. Researchers can get this software by contacting daiyk@sibet.ac.cn

The key features of the software are as follows: 1) It is a user-friendly software, providing graphical user interfaces (GUI) for all processing functions without any coding or configuration file edition. 2) It supplies DICOM, NIFTY, CSV and .mat file IOs for users to cooperate with other software. 3) It is an all-in-one software package, including data preparing, lesion labeling, feature extraction, feature selection, classification training, evaluation, visualization. 4) It can automatically extract features from multi-modality images.

II. MATERIALS AND METHODS

This software was developed by Python 3.6 and scikit-learn 0.19 for machine learning algorithm, VTK 7.1 for medical image visualization. The software provides user-friendly GUI developed by Qt 5.8, which is shown as Fig. 1. It contains seven major modules following the processing pipeline of machine learning study. It automatically establish a radiomics modeling pipeline by graphic user interface (GUI) setting. The architecture and workflow of the software are shown in Fig 2.

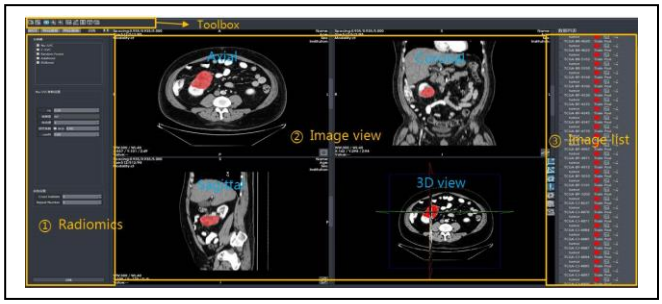

Fig. 1 Graphical user interfaces of the software

*Research supported by ABC Foundation.

Zhiyong Zhou is with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu, China (email:zhouzy@sibet.ac.cn)

Xusheng Qian is with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu, China (email:qianxs@sibet.ac.cn)

Jisu Hu is with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu, China (email:hujs@sibet.ac.cn)

Jianbing Zhu is with the Suzhou Science & Technology Town Hospital, Suzhou, Jiangsu, China (email: zeno1839@126.com)

Yakang Dai is with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu, China (email:daiyk@sibet.ac.cn)

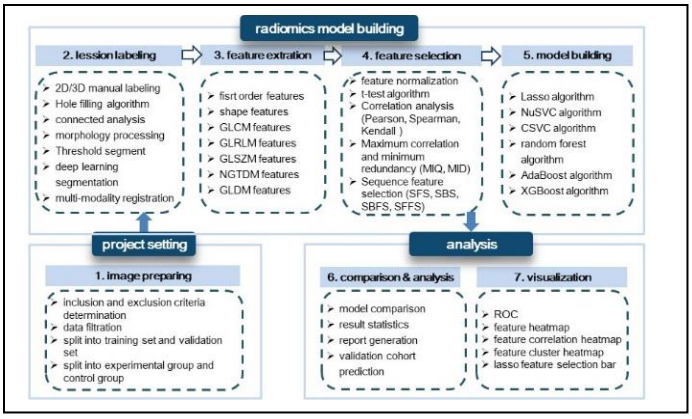

Fig 2. Architecture and workflow of the software

A. *Image preparation*

In order to use the software, users should prepare all medical images following the document architecture, which is shown in Fig 2. All medical images are considered to categorize into training cohort and independent validation cohort, where the training cohort is utilized to develop an optimal radiomics model, and the independent validation one used to evaluate the performance of the final model. Both of training and evaluation cohort subsequently are categorized into experimental group and control group according to the study task.

In the processing step of classifier training, software further splits training set into a real training subset and an internal validation subset for multi-fold cross-validation at running time, where the real training subset is used to develop and update a radiomics model, while internal validation subset used to determine optimal hyper-parameters. It is worthy to point out that the software is designed for automatically multi-modality medical image feature extraction and image processing. The users just need to prepare the multi-modality DICOM or NIFTY files in the corresponding modality folders in the same patient folder and assign modality names in GUI of the software.

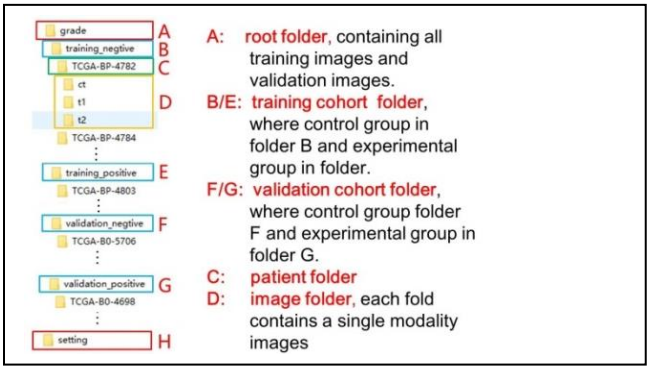

Fig 2. Document architecture.

B. *Lesion labeling*

It supplies human-computer interaction ways to label lesions or tissues. Users can delineate the outline of entire lesion in all contiguous slices, or design a semi-automatic segmentation pipeline by using segmentation algorithms integrated in this software. Comparing to time-costing manual delineating by radiologists, the main advantage of this semi-automatic process is to reduce manual intervention and guarantee the efficiency, consistency and reproducibility of labeling work. The software would automatically save the image, mask and their corresponding ROIs as NIFTY files for each sample in its respective folder.

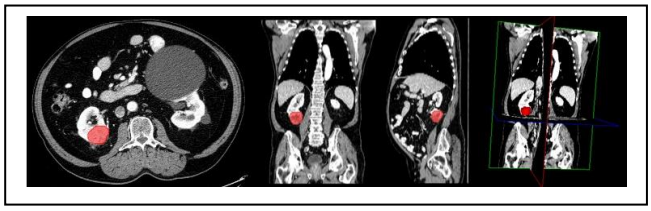

Fig 3. Visualization of lesion mask in software, subfigures from left to right are axial, coronal, sagittal, and 3D view.

C. *Feature extraction*

The software extracts quantitative image features by using PyRadiomics [7], which is an open-source python package for feature extraction. The software not only directly extract features from ROI in original images, but also from transformed images, including Laplacian of Gaussian, square, square root, logarithm, exponential, gradient, 3D local binary pattern, wavelet (LHL, LHH, HLL, LLH, HLH, HHH, HHL, and LLL). Finally, it generates and saves two CSV files in the root folder (training cohort and validation cohort respectively), which contain feature names and feature value matrix, as shown in Fig 4.

StudyID	ClassifyValue	ct_tumor_original_firstorder_90Percentile	ct_tumor_original_firstorder_Energy	ct_tumor_original_firstorder_Entropy
TCGA-CJ-6033	1	160	917140980257.00	2.768650301
TCGA-CW-5590	1	132	406007039219.00	2.684929376
TCGA-CZ-4863	1	105	209309509967.00	1.965770139
TCGA-B0-5088	0	119	58025209753.00	2.424008676
TCGA-BP-4173	0	162	473008515316.00	2.681813179
TCGA-BP-4177	0	184	942000467 2.00	2.291374124
TCGA-BP-4349	0	155	723169890 6.00	2.495235124
TCGA-BP-4353	0	129	4394388076.00	2.887873471

Fig 4. Data matrix in CSV file generated by the software. The first column indicates the image sample index, the second one shows the experimental group (here encoded as 1) or control group (encoded as 0). The top row, colored as yellow, describes the feature names, and the feature value matrix locates in the blue region.

D. *Feature selection*

Feature selection is to reduce the amount of features and keep the most relevant features for preventing over-fitting. It plays an important role in building a radiomics model of good performance and consistency.

The software firstly normalizes each feature column to avoid the effect of different scales. It subsequently supplies four types methods for selecting normalized features: t-test method, correlation analysis method, minimum redundancy maximum relevance (mRMR) [8], and sequential feature select method [9]. Correlation analysis was performed to identify the distinctiveness of features, remove the redundant and low reproducible features. Three classical correlation analysis algorithms: Pearson, Spearman, and Kendall methods are integrated in it. mRMR is a multivariate ranking method, which is applied to identify the most important features on the basis of a heuristic scoring criterion. The common algorithms: MIQ and MID can be used in it. Sequential feature select method is a widely used method, applied to select the discriminative features to enhance the performance of classifiers since accuracy of classification is achieved by a linear discriminant analysis [10] using leave-one-out cross validation, which is also employed in sequential feature select methods for measuring the significance of features. This software also contains four typical algorithms: sequential forward selection (SFS), sequential backward selection (SBS),

sequential forward floating selection (SFFS), and sequential backward floating selection (SBS). Users can select features with a single selection algorithm, or design a pipeline combined multiple algorithms. The final selected features are also automatically saved in a CSV file with the same format as its corresponding file of feature extraction.

E. Classifier training

This software provides seven machine learning algorithms for building radiomics models. The classifiers or regression models are trained by the selected features and corresponding labels. The users could set or choose the hyper-parameters in GUI of it to guarantee reproducibility of statistic results. It also supplies a convenient way in GUI to grid-search to optimize initial hyper-parameters. In order to evaluate the consistency and reproducibility of statistic results, researchers could use cross-validation with multiple repeat times, the software finally export a statistic inspection report.

F. Evaluation

The area under the receiver operating characteristic curve (ROC) is used to evaluate the statistic performance of the radiomics model on both of cross-validation of training cohort and independent validation cohort, as shown in Fig 5. The software displays all ROCs on the same figure for conveniently comparing and finding best performance of final models. It also lists common statistic indicators, including AUC, 95% confidence intervals of AUC, accuracy, sensitivity, specificity, precision, and others for quantitatively and comprehensively assessing performance. The best performed model is considered as the final radiomics model used to predict the prospective data.

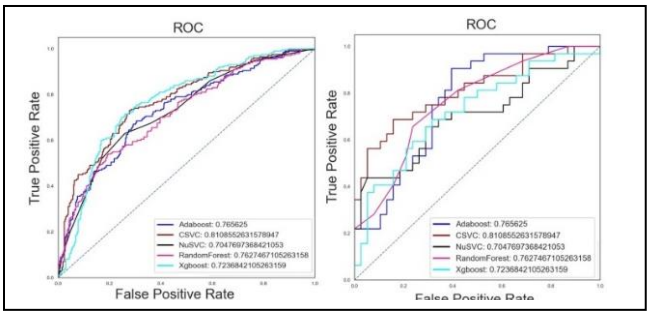

Fig 5. ROC for evaluation. The left figure shows performance of radiomics models on training cohort by five-fold cross-validation with repeating five times, and the right one shows performance on validation cohort.

G. Visualization

In order to intuitively demonstrate the role and relationship of selected features, the software automatically creates and saves feature heat-maps (shown in Fig 6) after model building. Users could switch to statistical chart for viewing these figures.

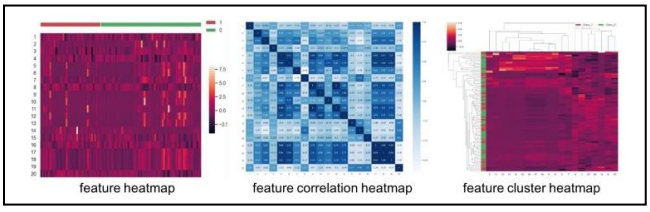

Fig 6. Feature visualization

III. EXPERIMENT AND RESULTS

We prepared CT images and used the software to build a radiomics model for grading clear cell renal cell carcinoma (ccRCC) in 2-tiered Fuhrman grade system, which is strategized into low grade and high grade. All CT images were downloaded from the Cancer Genome Atlas-Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) [11, 12]. The images in TCGA-KIRC were required from seven centers in US including multiple manufacture of GE, SIEMENS, and Philips. In this retrospective study, we prepared 203 CT images of ccRCC confirmed by pathology reports, where 122 samples (low/high=76/46) in training cohort and 81 samples (low/high=47/34) in validation cohort.

Radiologists labeled the ROIs by delineating the outline of entire tumor in all contiguous slices. The software subsequently created a corresponding peritumoral region by isotropically expanding and shrinking the borders of the tumor masses by 5mm in three dimensions. The main advantage of automatic morphology process is that it provides a convenient and effective tool to create peritumoral regions in medical images for cancer research. The morphology process and other tools in software offers a flexible and non-interactive pipeline for medical image preprocessing, leading to consistency and reproducibility of radiomics models.

The software respectively extracted 1760 quantitative 3D radiomics features from each lesion region and its corresponding peritumoral region, including seven feature types: first-order gray-level statistic features, 3D shape-based features, gray level co-occurrence matrix (GLCM) features, gray level run length matrix (GLRLM) features, gray level size zone matrix (GLSZM) features, neighboring gray tone difference matrix (NGTDM) features, and gray level dependence matrix (GLDM) features.

As a preprocessing procedure of feature selection, a standardization method was applied to normalize each radiomics feature to a zero mean and unit variance to prevent features in greater numeric ranges from dominating those in smaller numeric ranges. Subsequently, we combined Pearson correlation analysis [13] and minimum redundancy maximum relevance (mRMR) in the feature selection pipeline.

The Pearson correlation analysis was performed to identify the distinctiveness of features and remove the redundant and low reproducible features. In this radiomics study, we randomly removed one in pair-wise features if their Pearson correlation coefficient was larger than 0.9. mRMR was applied to identify the most important features on the basis of a heuristic scoring criterion after removing the redundant and low reproducible features. The software finally selected the top 20 ranked features as radiomics signature for ccRCC malignancy grading.

We simultaneously tried five classifiers (NuSVC, CSVC, random forest, AdaBoost and XGBoost) to find the best model. In order to determine an optimal hyper-parameter configuration of classifiers, we compared AUCs of classifiers on training cohort by using five-fold cross-validation with 5 repeat times. The finally performance was evaluated on the independent validation dataset.

In this study, the software costs 4272 seconds to complete the workflow of radiomics analysis. Actually, radiomics feature extraction is the most time-consuming processing, especially the final feature set from image types including Log, logarithm, square, square root, exponential, gradient, LBP, and wavelet image types.

In order to promote the time performance of this software, it automatically crops the images and only extract the features in the ROIs in the flow of feature extraction, which leads the time-consuming to be independent of image size and linearly increase corresponding to the total quantity of input images.

In this study, we focused on the role of peritumoral microenvironments in malignancy grading. We therefore built four types of radiomics models following the aforementioned model-building pipeline. Four types of radiomics models are: 1) features in tumor mass (PTV), 2) features in peritumoral region (TMV), 3) features in merging region of peritumoral region and tumor mass (rTM-PTV), 4) features combined with peritumoral region feature and tumor mass features (sTM-PTV). The performance of four type models is shown in TABLE 1.

TABLE I. PERFORMANCE OF RADIOMICS MODELS

Model Type	Training		Validation	
	AUC	*95% CI*	*AUC*	*95% CI*
PTV	0.807	0.800-0.834	0.848	0.760-0.936
TMV	0.773	0.744-0.802	0.810	0.706-0.914
rTM-PTV	0.797	0.768-0.825	0.811	0.712-0.970
sTM-PTV	0.802	0.775-0.829	0.836	0.744-0.928

IV. DISCUSSION

We developed a freely available software to speed up radiomics researches, which provides various functions, including image preparing, lesion labeling, feature extraction, feature selection, classifier training, evaluation and visualization. Users could build their custom models in pipeline since the software has a loosely coupled architecture and algorithms are independent of each other. It has a graphical user interface, providing a friendly interaction for users, avoiding configuration file edition or coding for radiologists.

This software also supplies multiple image IOs to read and export medical images. It could read multi-modality and multi-series DICOM files and NIFTY files. The imported image will be display its axial, coronal and sagittal slices in the software. It could automatically export and save labeled images as NIFTY files when lesion labeling finished. Further, the software integrates some image process algorithms and their real-time interactions developed by ITK 4.11 and VTK 7.1, which are open-source C++ packages. Users could designs a flexible semi-automatic segmentation pipeline composed of these algorithms to reduce intervention of manual segmentation for consistent and reproducibly labeling. An elaborate semi-automatic segmentation pipeline also offers possibility and convenience to efficiently label various lesions. In addition, this software also have a labeled image IO to read binary DICOM or NIFTY image files as masks. It offers a convenient way for users to label images with other freely available software packages, such as itk-SNAP (http://www.itksnap.org/) and 3D Slicer (https://www.slicer.org/). Beside of the image export IO, the developed software will automatically save all intermediate results in the pipelines, e.g. feature value matrix saved as a CSV file and statistic results matrix saved as .mat files. Researchers can process these intermediate results with other software, such as FeAture Explorer [14] and Matlab for further analysis or data visualization.

Comparison to the existing software (package), the main advantages of our software are (as shown in): 1) the developed software provides a complete pipeline of radiomics workflow in the convenient GUI; 2) provides semi-automatic and full-automatic segmentation for lesion labeling; 3) provides registration for lesion labeling in multi-modality or multi time-phase images.

TABLE II. COMPRISON OF THE EXISTING SOFTWARE

Competing systems	Open access	GUI	Custom modeling	Segment	Registraion
FeAture Explorer	√	√	√	×	×
GE A.K.	×	√	√	√	×
Radiomics	√	√	×	√	×
3D slicer	√	√	√	√	×
Ours	√	√	√	√	√

We use Pyradiomics in the software to extract image features. Researchers could choose specialized feature types and image filters. The software will extract all desired features from all images in bulk and combine them into a feature value matrix. It additionally provides an option for senior users to edit or modify feature extraction configuration file, which leads it to extract more features. It can select features with a single algorithm or a pipeline combined multiple algorithms. The final select features are used to train and optimize classifiers to predict prospective data.

We demonstrated the availability of the software with the clinical study of ccRCC malignancy grading on a multi-center CT image dataset. It extracted 1760 features for each ROI. We used 20 features was selected to build a radiomics model and evaluated the final performance of models on an independent validation cohort. Comparison the AUCs of different models (PTV, TMV, rTM-PTV, and sTM-PTV), we found the high-order texture features in peritumoral regions have more power ability to quantify heterogeneity of tumors in multiple perspectives and multi-scales, which revealed that peritumoral microenvironment plays an important role in malignancy grading of ccRCC.

We will sequentially develop the software in the following two ways: 1) providing artificial-intelligence-based segmentation algorithms for some tissues and lesions, 2) providing non-rigid registration algorithms for multi-modality medical image modeling, which maps a lesion mask in one modality to other modalities, reducing the manual interactions and guaranteeing consistency of segmentation between different modalities.

V. CONCLUSION

We developed a convenient and powerful software to build radiomics models for radiologists and researchers. It provides a user-friendly graphical interface, standardized radiomics modeling pipeline and practical image processing algorithms in a loosely coupled architecture. Users could utility it to rapidly build radiomics models, take a comprehensive comparison to get an optimal model, visualize the results and

feature heatmaps for their custom radiomics studies without any coding or configuration file edition.

VI. ACKNOWLEDGMENT

This study was supplied by National Key Research and Development Plan (2018YFA0703101), National Nature Science Foundation (61971413, 81971685), Jiangsu Key Technology Research Development Program (BE2020625), Suzhou key clinical diseases diagnosis and treatment technology project (LCZX201930, LCZX202001), Suzhou Science & Technology Projects for People's Livelihood (SS2019068, SS2019012), Youth Innovation Promotion Association CAS (2021324), Quancheng 5150 Project.

REFERENCES

[1] Ding, J., et al., CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. European Journal of Radiology, 2018. 103: p. 51-56.

[2] P, L., et al., Radiomics: Extracting more information from medical images using advanced feature analysis. European Journal of Cancer, 2012. 48(4): p. 441-446

[3] X, Z., et al., Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. European Radiology, 2018. 28(7): p. 2772-2778.

[4] Choi, W., et al. Radiomics Analysis of Pulmonary Nodules in Low Dose CT for Early Detection of Lung Cancer. in Aapm Meeting & Exhibition. 2017.

[5] Ferreira Junior, J.R., et al., Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. 2018. 159(1): p. 23-30.

[6] Zwanenburg A , A.M.A., Apte A , et al. , Image biomarker standardisation initiative. Radiotherapy & Oncology, 2018(127): p. S543-S544.

[7] Griethuysen J J M V , F.A., Parmar C , et al. , Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research, 2017. 77(21): p. e104-e107.

[8] Peng, H., et al., Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. 2005. 27(8): p. 1226-1238.

[9] Pudil, P., J. Novovičová, and J. Kittler, Floating search methods in feature selection. Pattern Recognition Letters, 1994. 15(11): p. 1119-1125.

[10] Wei-Chien, W., C. Liang-Bi, and C.J.A.S. Wan-Jung, Development and Experimental Evaluation of Machine Learning Techniques for an Intelligent Hairy Scalp Detection System. 2018. 8(6): p. 853.

[11] Clark, K., et al., The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. Journal of Digital Imaging, 2013. 26(6): p. 1045-1057.

[12] Akin, P.E., M. Heller, R. Jarosz, B.J. Erickson, S. Kirk, J. Filippini, Radiology Data from The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma [TCGA-KIRC] collection. Cancer Imaging Arch, 2016.

[13] Gillies, R.J., P.E. Kinahan, and H. Hricak, Radiomics: Images Are More than Pictures, They Are Data. Radiology, 2016. 278(2): p. 563-77.

[14] Song, Y., et al., FeAture Explorer (FAE): A tool for developing and comparing radiomics models. PLOS ONE, 2020