

## RESEARCH ARTICLE

## FeAture Explorer (FAE): A tool for developing and comparing radiomics models

Yang Song<sup>1</sup> , Jing Zhang<sup>1</sup> , Yu-dong Zhang<sup>2</sup>, Ying Hou<sup>2</sup>, Xu Yan<sup>3</sup>, Yida Wang<sup>1</sup>, Minxiong Zhou<sup>4</sup>, Ye-feng Yao<sup>1\*</sup> , Guang Yang<sup>1\*</sup> 

**1** Shanghai Key Laboratory of Magnetic Resonance, East China Normal University, Shanghai, China, **2** Department of Radiology, The First Affiliated Hospital with Nanjing Medical University, Nanjing, China, **3** MR Scientific Marketing, Siemens Healthcare, Shanghai, China, **4** Shanghai University of Medicine & Health Sciences, Shanghai, China

 These authors contributed equally to this work.

\* [gyang@phy.ecnu.edu.cn](mailto:gyang@phy.ecnu.edu.cn) (GY); [yfyao@phy.ecnu.edu.cn](mailto:yfyao@phy.ecnu.edu.cn) (YY)

 OPEN ACCESS

**Citation:** Song Y, Zhang J, Zhang Y-d, Hou Y, Yan X, Wang Y, et al. (2020) FeAture Explorer (FAE): A tool for developing and comparing radiomics models. PLoS ONE 15(8): e0237587. <https://doi.org/10.1371/journal.pone.0237587>

**Editor:** Stavros I. Dimitriadis, Cardiff University, UNITED KINGDOM

**Received:** February 7, 2020

**Accepted:** July 29, 2020

**Published:** August 17, 2020

**Copyright:** © 2020 Song et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available from Figshare: <https://doi.org/10.6084/m9.figshare.12751391.v1>.

**Funding:** This study was supported in part by the National Key Research and Development Program of China (2018YFC1602800), and the Key Project of the National Natural Science Foundation of China (61731009). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors received no specific funding for this work. Siemens Healthcare provided support in the form

## Abstract

In radiomics studies, researchers usually need to develop a supervised machine learning model to map image features onto the clinical conclusion. A classical machine learning pipeline consists of several steps, including normalization, feature selection, and classification. It is often tedious to find an optimal pipeline with appropriate combinations. We designed an open-source software package named FeAture Explorer (FAE). It was programmed with Python and used NumPy, pandas, and scikit-learning modules. FAE can be used to extract image features, preprocess the feature matrix, develop different models automatically, and evaluate them with common clinical statistics. FAE features a user-friendly graphical user interface that can be used by radiologists and researchers to build many different pipelines, and to compare their results visually. To prove the effectiveness of FAE, we developed a candidate model to classify the clinical-significant prostate cancer (CS PCa) and non-CS PCa using the PROSTATEx dataset. We used FAE to try out different combinations of feature selectors and classifiers, compare the area under the receiver operating characteristic curve of different models on the validation dataset, and evaluate the model using independent test data. The final model with the analysis of variance as the feature selector and linear discriminate analysis as the classifier was selected and evaluated conveniently by FAE. The area under the receiver operating characteristic curve on the training, validation, and test dataset achieved results of 0.838, 0.814, and 0.824, respectively. FAE allows researchers to build radiomics models and evaluate them using an independent testing dataset. It also provides easy model comparison and result visualization. We believe FAE can be a convenient tool for radiomics studies and other medical studies involving supervised machine learning.

## Introduction

Radiomics extracts a large number of quantitative features from the regions of interest (ROIs) in medical images and maps them to clinical conclusions with machine learning (ML) [1, 2]. Radiomics utilizes not only shape, intensity (first-order), and texture (second-order) features

of salaries for authors [X. Y.], but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

**Competing interests:** The Author (Xu Yan) from a commercial company, Siemens Healthcare, was a MR collaboration scientist doing technical support in this study under Siemens collaboration regulation without any payment and personal concern regarding to this study. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

from the original image, but also higher and more abstract features such as those in the wavelet transform domain [3]. While features based on shape, intensity, and texture can be easily compared to observations of radiologists, abstract features often reveal hidden information in the image and may be used to explain the intuition of experts. Radiomics models can be integrated into computer-aided diagnosis (CAD) systems [4] to analyze medical images, and even resolve uncertain cases.

Establishing radiomics models involves supervised learning, which is a subclass of machine learning. Most works of radiomics belong to a binary classification problem [5], which is usually a primer project about radiomics. A typical radiomics study includes the following steps: 1) The dataset is split into two parts: a training dataset used to develop the model, and an independent test dataset used to evaluate the model. 2) The training dataset is further split into a real training dataset and a validation dataset. The real training dataset is used to develop a model and update the parameters (e.g., the slope and intercept of a linear model), and the validation dataset is used to determine the hyperparameters (e.g., a linear model or quadratic model). When the size of the training dataset is limited, cross-validation is often used instead of an independent validation dataset. 3) Each feature is normalized to avoid the effect of different scales [6]. 4) Feature selection and dimension reduction are used to reduce the number of features and keep only those most relevant to the diagnosis to prevent overfitting [7]. 5) A classification or regression model is trained with selected features and the label. 6) The performance of the model on the validation dataset is evaluated using typical clinical statistics such as the area under the receiver operating characteristic (ROC) curve (AUC) [8]. 7) Steps 3 to 6 are repeated with different hyperparameters (e.g., different classifiers). 8) The model with the best performance on the validation dataset is determined, and its performance is evaluated on the independent test dataset.

In the above process, there are many candidate algorithms in each step. To find the optimal model, researchers often need to explore different combinations of algorithms, which can be tedious and may require expertise in machine learning. To address this problem, we designed an open-source software package called FeAture Explorer, which can be used in radiomic studies. Researchers can use FAE to develop models with different combinations of feature normalization methods, feature selection algorithms, and classifiers. FAE can automatically try out all specified combinations of different algorithms and compare the performances of the established models on the validation dataset. Further, in radiomics studies, researchers usually focus on the interpretation of the features in the final built model. FAE also provides common clinical statistics and visualization of models, which helps to facilitate the comparison of different models.

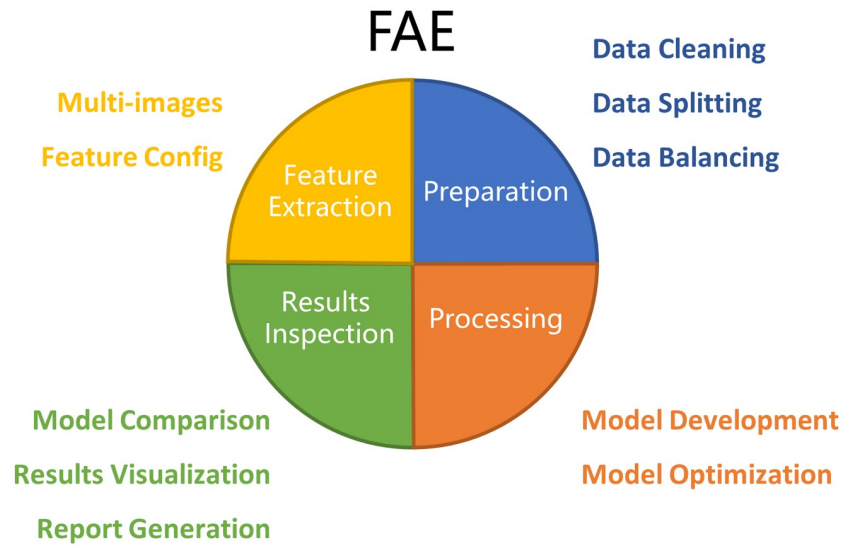
## Materials and methods

### Software design

FAE has four major functions: Feature Extraction, Data Preparation, Model Development, and Results Inspection (Fig 1).

**Feature extraction.** Pyradiomics, an open-source module for radiomics feature extraction, was used in FAE [9]. Users need to save the images and the corresponding ROIs in NIFTY format and store the files for each case in a separate subfolder in a root folder (Fig 2), before FAE can be used to extract radiomics features for all the cases in a batch. Features from multiple images, including transformed images, such as wavelet, log images, can also be extracted with the same ROIs.

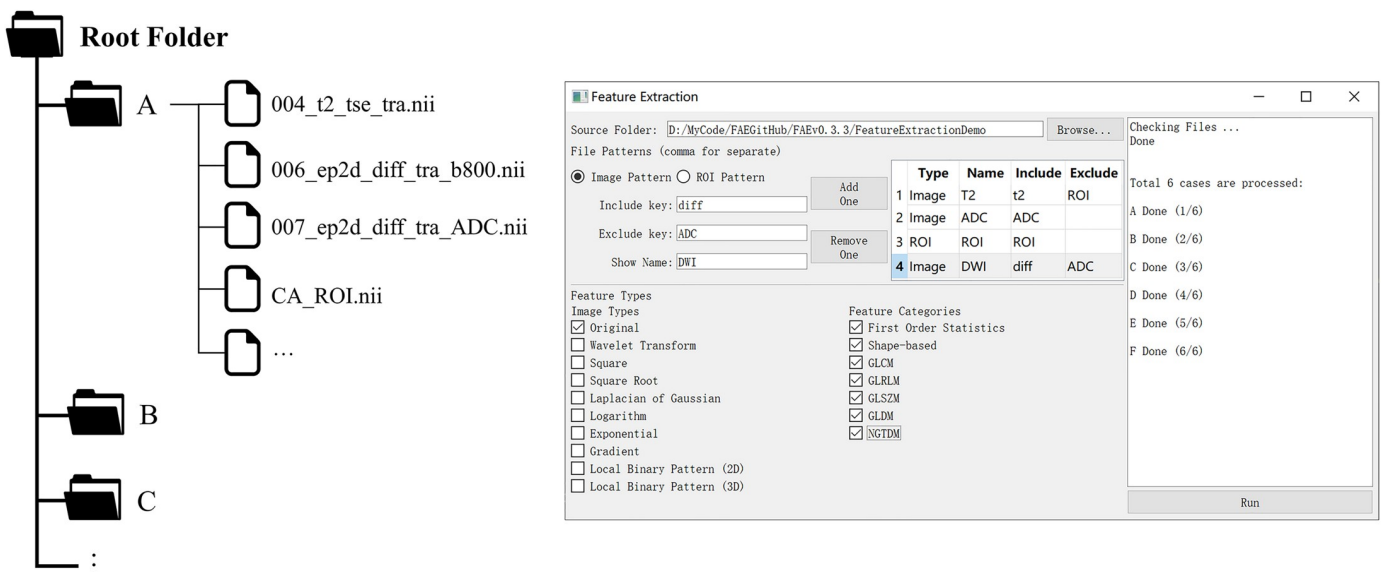
**Data preparation.** To use FAE, one must prepare a comma-separated values (CSV) file containing a feature matrix in the following format (Fig 3): 1) Each row contains one case, and



**Fig 1. Architecture of FAE.** Major functions of FAE include *Feature Extraction, Data Preparation, Model Development, and Result Inspection.*

<https://doi.org/10.1371/journal.pone.0237587.g001>

each column denotes one feature. 2) The first row contains the names of the features, and the first column contains index of the cases. 3) There should be one column with the header of “label” to denote the clinical conclusions. After the feature matrix is loaded, FAE checks the validity of the feature values. FAE can export a log file to help users see which cases and features have invalid values, such as null or text. The features or cases including invalid values can also be removed by one-click in FAE. It should be noted that if a feature is categorical instead of numeric (e.g., gender or tumor stage), the user needs to encode it numerically. FAE can also



**Fig 2. Folder structures and the user interface for feature extraction.** For each case, images and the corresponding ROIs were stored in a sub-folder. FAE can be used to extract different kinds of features from different kinds of images conveniently.

<https://doi.org/10.1371/journal.pone.0237587.g002>

	A	B	C	D	E	F	G
Row: Cases	1	ID	label	ADC_ge	ADC_ge	ADC_orig	ADC_orig
Column: Features	2	N10	0	1	338	1255	1762.009
	3	N100	0	1	71	1113	1327.984
	4	N101	1	1	253	734.96	1246.223
	5	N102	0	1	80	1168.8	1485.713
Case ID	6	N103	1	1	506	1162	1361.492
Label/Target	7	N106	0	1	75	857.83	1134.199
Feature Name	8	N108	0	1	50	1628.4	1959.877
Feature Matrix	9	N109	0	1	84	1183.2	1467.481
	10	N11	1	1	74	746.22	1047.104
	11	N110	1	1	116	1275	1535.002
	12	N111	0	1	37	1249.4	1646.031
	13	N114	1	1	586	766	1103.9

**Fig 3. Format of feature matrix used for FAE.** Case index is saved in first column and label in second column. Feature names are stored in first row. File should be saved in comma-separated values (CSV) format.

<https://doi.org/10.1371/journal.pone.0237587.g003>

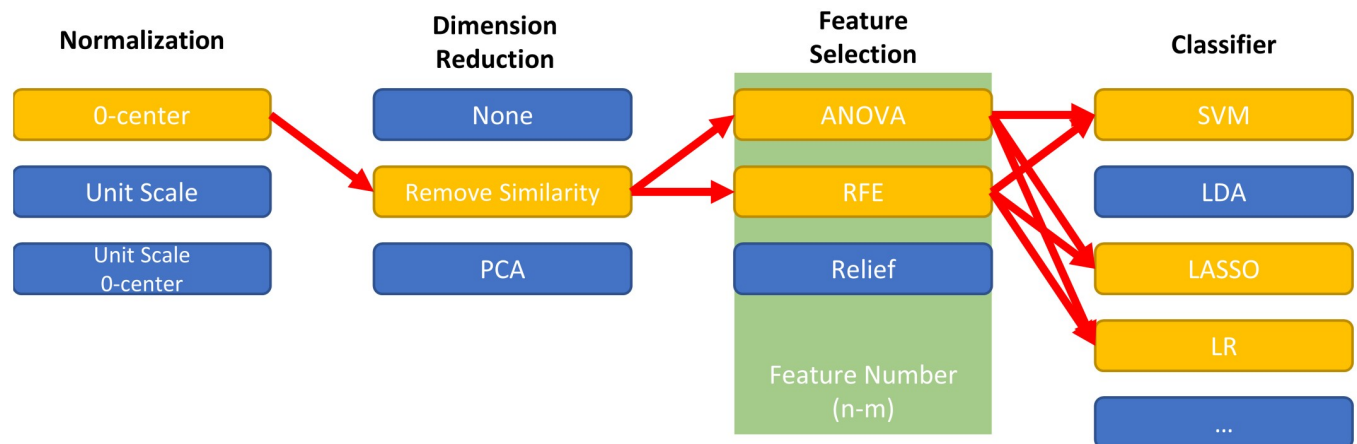
split the dataset into training and test datasets with a specified ratio, and it will maintain the ratio of positive cases to negative cases constant in the two datasets. If the positive and negative cases are not balanced in number, then FAE provides downsampling, upsampling, synthetic minority oversampling technique (SMOTE) [10] or SMOTETomek [11] to balance the training dataset.

**Model development.** FAE can help develop a radiomics pipeline that includes the following steps: data normalization, dimension reduction, feature selection, and classifier. Available options for all the components are listed in Table 1. FAE uses all possible combinations of these components to construct radiomics pipelines, train models using the training data, and evaluate the respective results on the validation dataset. For example, if one researcher selects

**Table 1. Available options for each step in the radiomics pipeline: Data normalization, dimension reduction, feature selection, and classifier.** All methods implemented based on scikit-learn [12].

Steps	Candidate
Data	Min-max Normalization
Normalization	Z-score Normalization
	Mean Normalization
Dimension	Pearson Correlation Coefficient (PCC)
Reduction	Principle Component Analysis (PCA)
Feature Selection	Analysis of Variance (ANOVA)
	Recursive Feature Elimination (RFE)
	Relief
Classifier	Linear Regression
	Least Absolute Shrinkage and Selection Operator (LASSO)
	Support Vector Machine (SVM)
	Linear Discriminant Analysis (LDA)
	Decision Tree
	Random Forest
	Adaboost
	Gaussian Process
Naïve Bayes	
Multilayer Perceptron	

<https://doi.org/10.1371/journal.pone.0237587.t001>



**Fig 4. Demo of combination of one radiomics pipeline.** For each step of radiomics pipeline, FAE provides some options from which researchers can choose. FAE will try out all possible combinations of selected options using the training dataset.

<https://doi.org/10.1371/journal.pone.0237587.g004>

two feature selection methods and three different classifiers, then FAE constructs six pipelines to develop six models (Fig 4). We fix the random seed for each component to make sure the result is reproducible and users can specify any random seed value they prefer for any component in the config file. FAE can be used to grid-search the best hyper-parameters if specified in a config file. For radiomics studies where the dataset size is limited, FAE can use five-fold or ten-fold cross-validation. All models are saved with common clinical statistics and all interim results, which can be loaded later for inspection and allows researchers to use other software to conduct further analysis.

**Result inspection.** FAE lists common clinical statistics and displays typical figures for model evaluation (Fig 5). The statistics presented include the AUC, 95% confidence intervals of AUC, accuracy, sensitivity, specificity, and others for the training dataset and validation dataset. Statistics for all models are ordered by the AUC on the validation dataset, and thus different models can be compared easily to find the best model. To help the researcher to select right number of features in the final model and to gain a better understanding of the model, FAE shows a plot of AUC against the number of selected features together with the weights of each feature in the model.

FAE was developed in Python 3.6 (<https://python.org>) and a Qt framework. Most algorithms in FAE were implemented with scikit-learn 0.19 (<https://scikit-learn.org>) [12]. The source code is openly available on GitHub (<https://github.com/salan668/FAE.git>).

## Retrospective experiments

To demonstrate the usability of FAE, we used it to differentiate clinically significant prostate cancer (CS PCa) and non-CS PCa (NCS PCa) using radiomics features of lesions in multiparametric magnetic resonance imaging (mp-MRI). We selected 185 cases from the PROSTATEx dataset [13], a public prostate cancer dataset containing 68 CS and 184 NCS PCa. We used the T2W image (TSE,  $0.5 \times 0.5 \times 3.6 \text{ mm}^3$ ), diffusion weighted images (DWI, SSEP,  $2 \times 2 \times 3.6 \text{ mm}^3$ ) with b-value =  $800 \text{ s/mm}^2$ , and the corresponding apparent diffusion coefficient (ADC) maps for feature analysis, which conform to the recommendation of prostate imaging-reporting and data system (PI-RADS). One radiologist with 11 years of clinical experience (Y. Z.) drew the boundary of lesions on T2W transverse images using biopsy results together with



**Fig 5. Graphic user interface of Result Inspection part of FAE.** Common clinical statistics for each model are shown on training dataset and validation dataset. Models are compared using ROC curve, AUC of each model, and feature contributions. FAE can also create high-resolution plots and generate brief description of algorithms used in selected model.

<https://doi.org/10.1371/journal.pone.0237587.g005>

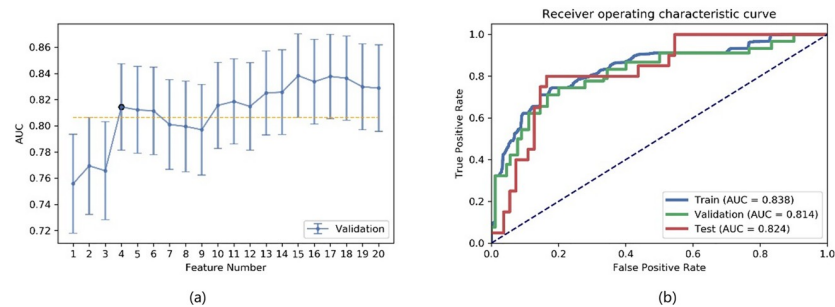
ADC and DWI images as reference. We used Elastix (<http://elastix.isi.uu.nl/index.php>) to align the DWI and ADC images to the T2W images [14].

## Results

We prepared the MR images and used FAE to extract quantitative features from these three series. Types of features extracted included the shape, histogram (intensity-based or first-order), gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), and gray level size zone matrix (GLSZM). From each PCa ROI, 264 features were extracted.

We loaded the feature matrix in the *Data Preparation* page in FAE and split the dataset into a training dataset of 177 samples (CS-PCa/NCS-PCa = 48/129) and an independent test dataset of 75 samples (CS-PCa/NCS-PCa = 48/129). In the *Model Development* page of FAE, we designed the pipelines of radiomics models. We normalized the dataset by *Z-score Normalization*, which subtracts the mean value and divides the standard deviation for each feature. To reduce the dimensions of the row space of the feature matrix, we used a *Pearson Correlation Coefficient (PCC)* for each pair of two features [15]. If the PCC was larger than 0.9, we removed one of them randomly in FAE. We chose the Analysis of Variance (ANOVA), Relief [7], and Recursive Feature Elimination (RFE) for feature selection. The range of the feature number was set from 1 to 20. We also tried a Support Vector Machine (SVM) [16], Linear Discriminative Analysis (LDA) [5], Random Forest [17], and Logistical Regression (LR) [18] to find the best model. The performance of all pipelines was evaluated using five-fold cross-validation.

We compared the AUC of all pipelines on the validation dataset with FAE. The pipeline using ANOVA feature selection and an LDA classifier yielded the highest AUC using 15 features. When “one-standard error” rule was used, FAE also produced a simpler model with only four features [Fig 6(a)] [5] whose ROC curves are shown in Fig 6(b). The AUCs of the training/validation/test datasets achieved 0.838/0.814/0.824, respectively.



**Fig 6. Performance of model generated by FAE.** (a) FAE suggested a candidate four-feature model according to “one-standard error” rule, and (b) ROC curves of this model on different datasets.

<https://doi.org/10.1371/journal.pone.0237587.g006>

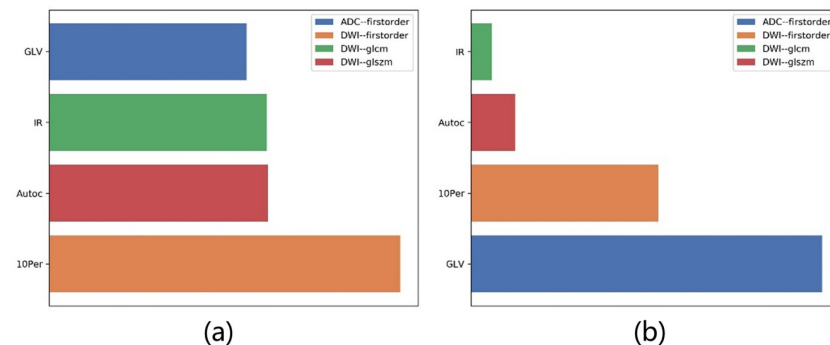
Features selected by FAE were the gray level variance of the gray level size zone matrix (GLSZM) from DWI ( $GLV_{DWI}$ ), interquartile range of the histogram from DWI ( $IR_{DWI}$ ), auto-correlation of the gray level cooccurrence matrix (GLCM) from DWI ( $AC_{DWI}$ ), and 10% percentage point of the histogram from ADC ( $10Per_{ADC}$ ).

In the feature selection step, we sorted the features according to their F-values by ANOVA. The F-values of four selected features were  $GLV_{DWI} = 72.91$ ,  $IR_{DWI} = 45.47$ ,  $AC_{DWI} = 45.22$ , and  $10Per_{ADC} = 41.03$ . We used these four features to develop an LDA model. The coefficients were  $10Per_{ADC} = 7.98$ ,  $IR_{DWI} = -1.89$ ,  $AC_{DWI} = 0.95$ , and  $GLV_{DWI} = 14.96$ . The F-values and their contributions are shown in Fig 7(a) and 7(b). We found that the order of the contributions of the features was different in ANOVA and LDA.

The histograms of each selected feature in the two classes are shown in Fig 8. While no single feature can be used to classify the CS and NCS PCa perfectly, each feature has different distributions in the two classes. This explains why we can combine these four features to establish a good model.

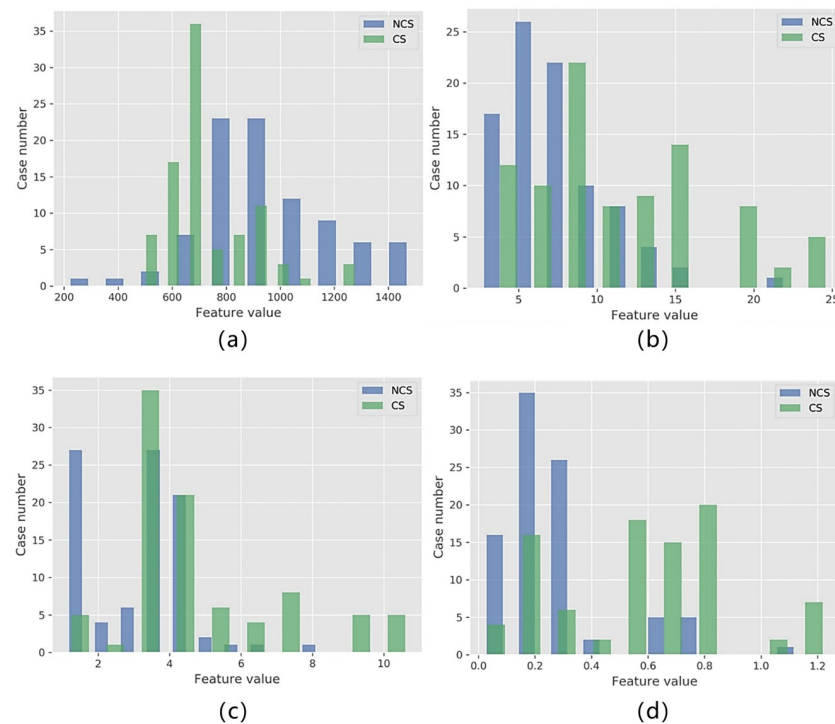
## Discussions

We developed an open-source software package named FeAture Explorer to facilitate model development and feature exploration in radiomics studies. FAE can be used to process the feature matrix, develop different models, and visualize the results conveniently. For each model, FAE provides a number of statistics together with typical diagrams in which radiologists might



**Fig 7. Contribution of features in final model.** F-values of selected features by (a) ANOVA, and (b) coefficients of these features in LDA model. Order of contributions of these features is different in ANOVA and LDA.

<https://doi.org/10.1371/journal.pone.0237587.g007>



**Fig 8. Histograms of selected features in CS PCA and NCS PCA.** Distribution of  $GLV_{DWI}$ ,  $IR_{DWI}$ ,  $AC_{DWI}$ , and  $10Per_{ADC}$  shown as (a) through (d), respectively.

<https://doi.org/10.1371/journal.pone.0237587.g008>

be interested. FAE can also produce a brief description of the model, including its predictive ability, components in the pipeline, and features selected.

We demonstrated the use of FAE using CS/NCS PCA differentiation on the PROSTATEx dataset. We tried 12 different pipelines (combining 3 feature selection methods and 4 classifiers) and compared their AUCs on the validation dataset to find the optimal model. Then, FAE was used to find a more robust model that used four features but achieved a comparable AUC. We used an independent testing dataset to evaluate the performance of the final model. The AUC of 0.824 on the test dataset is comparable to that in the work of L. Wang [19]. The AUC of the model on the test dataset was close to those on the training and validation datasets, implying that the model was neither overfitting nor underfitting.

FAE can be used not only for radiomics studies but also for supervised binary classification problems. Since all processors (algorithms) in the FAE pipeline are independent of each other, researchers can add their own algorithms to construct specific pipelines. In addition, since FAE saves all intermediate results in the pipelines, researchers can process these intermediate results with other software. For instance, researchers often compare the model built with only clinical features and the model with image features. FAE can build the radiomics model and export selected radiomics features. Then we can combine them with the clinical features for model development [20]. Moreover, models developed with FAE can be combined with the ensemble method [21].

We used pyradiomics in FAE to extract image features. However, we did not integrate a UI for ROI labeling. ROI labeling requires real time interaction and is often compute-intensive when semi-automatic segmentation algorithms are involved, where C and C++ programs are more suitable. Some free (e.g., 3D Slicer) or commercial (e.g., Frontier of Siemens Healthcare)



software packages can be used for ROI labeling. While some of them can also be used for feature extraction, FAE is more efficient because it can extract all desired features from all cases and combine them into a feature matrix in one click.

There are several limitations of FAE as follows: 1) In machine learning, the hyperparameters of the feature selection and classifiers have an important effect on the performance of the model [5]. While FAE provides an option to tune hyper-parameters with a config file, it lacks a user-friendly UI for hyper-parameter tuning. 2) FAE can be used conveniently for binary classification, but it has not provided an integrated UI for multi-label classification and regression problems for now. These problems should be addressed in future versions.

## Conclusion

FAE provides a convenient way to develop radiomics models. Radiologists can use FAE to clean the feature matrix, develop different models, compare the performance, and visualize the results for their own radiomics studies. We opened the source code of FAE, and researchers can use it freely and extend it to test new ideas and algorithms at will.

## Author Contributions

**Conceptualization:** Yang Song.

**Data curation:** Yu-dong Zhang, Ying Hou.

**Formal analysis:** Yang Song, Jing Zhang.

**Funding acquisition:** Ye-feng Yao, Guang Yang.

**Project administration:** Guang Yang.

**Software:** Yang Song, Jing Zhang.

**Validation:** Jing Zhang, Xu Yan, Yida Wang, Minxiong Zhou.

**Writing – original draft:** Yang Song.

**Writing – review & editing:** Guang Yang.

## References

1. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012; 30(9):1234–48. Epub 2012/08/18. <https://doi.org/10.1016/j.mri.2012.06.010> PMID: 22898692.
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2015; 278(2):563–77. <https://doi.org/10.1148/radiol.2015151169> PMID: 26579733
3. Manjunath BS, Ma WY. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1996; 18(8):837–42. <https://doi.org/10.1109/34.531803>
4. Wang S, Burt K, Turkbey B, Choyke P, Summers RM. Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research. *Biomed Res Int*. 2014; 2014:789561. Epub 2014/12/20. <https://doi.org/10.1155/2014/789561> PMID: 25525604.
5. James G. An introduction to statistical learning: with applications in R. New York: Springer; 2017.
6. Kotsiantis S, Kanellopoulos D, Pintelas P. Data preprocessing for supervised learning. *International Journal of Computer Science*. 2006; 1(2):111–7.
7. Kira K, Rendell LA. A practical approach to feature selection. *Machine Learning Proceedings 1992*: Elsevier; 1992. p. 249–56.
8. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982; 143(1):29–36. <https://doi.org/10.1148/radiology.143.1.7063747> PMID: 7063747

9. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*. 2017; 77(21):e104. <https://doi.org/10.1158/0008-5472.CAN-17-0339> PMID: 29092951
10. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002; 16:321–57.
11. Batista GEAPA, Prati RC, Monard MC. A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor Newsl*. 2004; 6(1):20–9. <https://doi.org/10.1145/1007730.1007735>
12. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*. 2011; 12(Oct):2825–30.
13. Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans Med Imaging*. 2014; 33(5):1083–92. Epub 2014/04/29. <https://doi.org/10.1109/TMI.2014.2303821> PMID: 24770913.
14. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. Elastix: a toolbox for intensity-based medical image registration. *IEEE transactions on medical imaging*. 2010; 29(1):196. <https://doi.org/10.1109/TMI.2009.2035616> PMID: 19923044
15. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. *Noise reduction in speech processing*: Springer; 2009. p. 1–4.
16. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*. 2011; 2(3):27.
17. Breiman L. Random forests. *Machine learning*. 2001; 45(1):5–32.
18. Hinton GE. Connectionist learning procedures. *Machine learning*: Elsevier; 1990. p. 555–610.
19. Min X, Li M, Dong D, Feng Z, Zhang P, Ke Z, et al. Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: Cross-validation of a machine learning method. *Eur J Radiol*. 2019; 115:16–21. Epub 2019/05/16. <https://doi.org/10.1016/j.ejrad.2019.03.010> PMID: 31084754.
20. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol*. 2016; 34(18):2157–64. Epub 2016/05/04. <https://doi.org/10.1200/JCO.2015.65.9128> PMID: 27138577.
21. Kuncheva LI, Whitaker CJ. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine learning*. 2003; 51(2):181–207.