# Web-based Radiomics Module for Image Feature Extraction for Tumor Characterization

Theresa Huebner [a,b,c], Aaron Berger [a,b], Daniel Wild [a,b], Jianning Li [a,b,c], Antonio Pepe [a,b,c], Christina Gsaxner [a,b], Yuan Jin [a,b,d], Gijs Luijten [a,b,c], Jens Kleesiek [c] and Jan Egger [a,b,c]*

[a] Graz University of Technology, Institute for Computer Graphics and Vision, Graz, Austria
[b] Computer Algorithms for Medicine Laboratory, Graz, Austria
[c] Institute for AI in Medicine (IKIM), University of Medicine Essen (AöR), Essen, Germany
[d] Research Center for Connected Healthcare Big Data, Zhejiang Lab, Hangzhou, Zhejiang, China
*Corresponding author: Jan Egger, E-mail: egger@tugraz.at

## INTRODUCTION

Nowadays, traditional interpretation of medical images by trained radiologists based on single image features is increasingly supported by radiomics approaches. This means that medical images are analyzed based on hundreds to thousands of radiomic features that are extracted in an automated fashion. Radiomics refers not only to the extraction of large amounts of high-dimensional quantitative features from multimodal medical images, but also to clinical data to detect quantitative correlations. Based on the analysis of these interrelationships, radiomics can provide insights into microanatomy, pathophysiology and yet unknown correlations for personalized medicine[1]. For understanding what radiomics does in particular, it is important to look at the radiomics pipeline as it consists of the following components that build up on each other[2]. The steps should be carried out in a standardized manner to minimize sources of error and to ensure comparability and reproducible results:

**Preprocessing, segmentation, feature extraction and modeling.**

## METHODS

The Radiomics module in StudierFenster (http://www.studierfenster.at)[3] allows the extraction of commonly used radiomic features in a user-friendly environment, similar to the cranial implant design and centerline modules. It provides users who have no coding background with the opportunity to upload already segmented data to the module for the feature extraction. The features can then be used for various down-stream image analysis and model building tasks, such as diagnosis.

The client side of the StudierFenster framework, defined in the frontend repository, is implemented with the three leading technologies of frontend web development Hypertext Markup Language (HTML), JavaScript and Cascading Style Sheets. A Python web server with a Flask framework provides the backend. The communication between the server and the client takes place by a Web Server Gateway Interface. By accepting requests from the client, the Flask framework manages the execution of new sub-processes within Python and C++ programs.

Figure 1 shows all the major folders, important files, and links of the Radiomics module. The red arrows show the connections between the files, and the numbers show the general order, starting with the upload and ending with the download. All files and folders are categorized into two main folders of the study window, namely, the frontend, and the backend. Starting from the frontend, the main HTML file *index.html* defines the user interface, such as the button for data uploading.
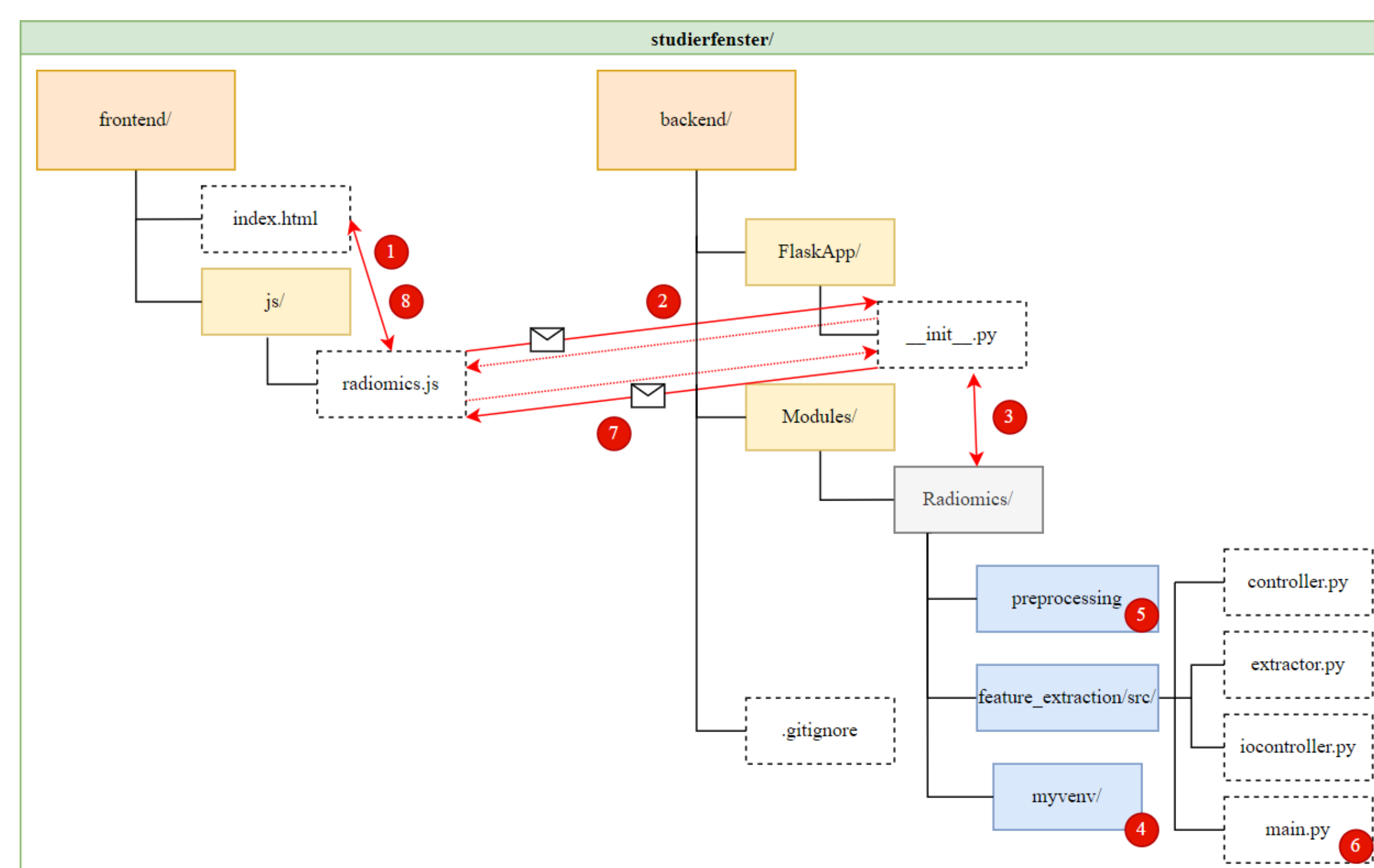


Fig. 1: Overview of relevant folders and files for the Radiomics module.

## RESULTS

The Radiomics module was tested with data of different sizes, case numbers and formats. The test cases came from different public data collections, namely KiTS19[4], RIDER[5] and the Skull-stripped MRI GBM Datasets[6]. The respective runtimes can be seen in Table 1. The time it takes to upload the data to the server was excluded, as it depends heavily on the user's internet connection.

| Runtime List | | | | | |
|---|---|---|---|---|---|
| Nr. | Size | Format | Nr. Cases | Dataset | Runtime |
| 1 | 46.262 KB | .nii.gz | 2 | KiTS19 | 2 min |
| 2 | 306.014 KB | .nii.gz | 7 | KiTS19 | 5 min |
| 3 | 508.063 KB | .nii.gz | 4 | KiTS19 | 7 min |
| 4 | 64.602 KB | .nrrd | 1 | RIDER | 2 min |
| 5 | 78.140 KB | .nrrd | 1 | RIDER | 1 min |
| 6 | 1.438 KB | .nrrd | 1 | Skull-stripped MRI GBM Datasets | <1 min |

Table 1: Runtimes of the test cases, which have been taken from KiTS19, RIDER and the Skull-stripped MRI GBM Datasets.

Furthermore, about 100 cases of the KiTS19 dataset were processed using the radiomics module. The 1410 features obtained for each case were stored in CSV files. Figure 2 shows a typical radiomics correlation diagram of these features.

Noticeable are the white bands resulting from non-valid values of features. There are also many yellow and green clusters, which are taken as groups of features with very similar results. Features that are very similar or meaningless should be removed in the next step, the feature selection. The selection also reduces overfitting in subsequent model building. The fewer features and the more cases included in a model, the lower the risk of overfitting.
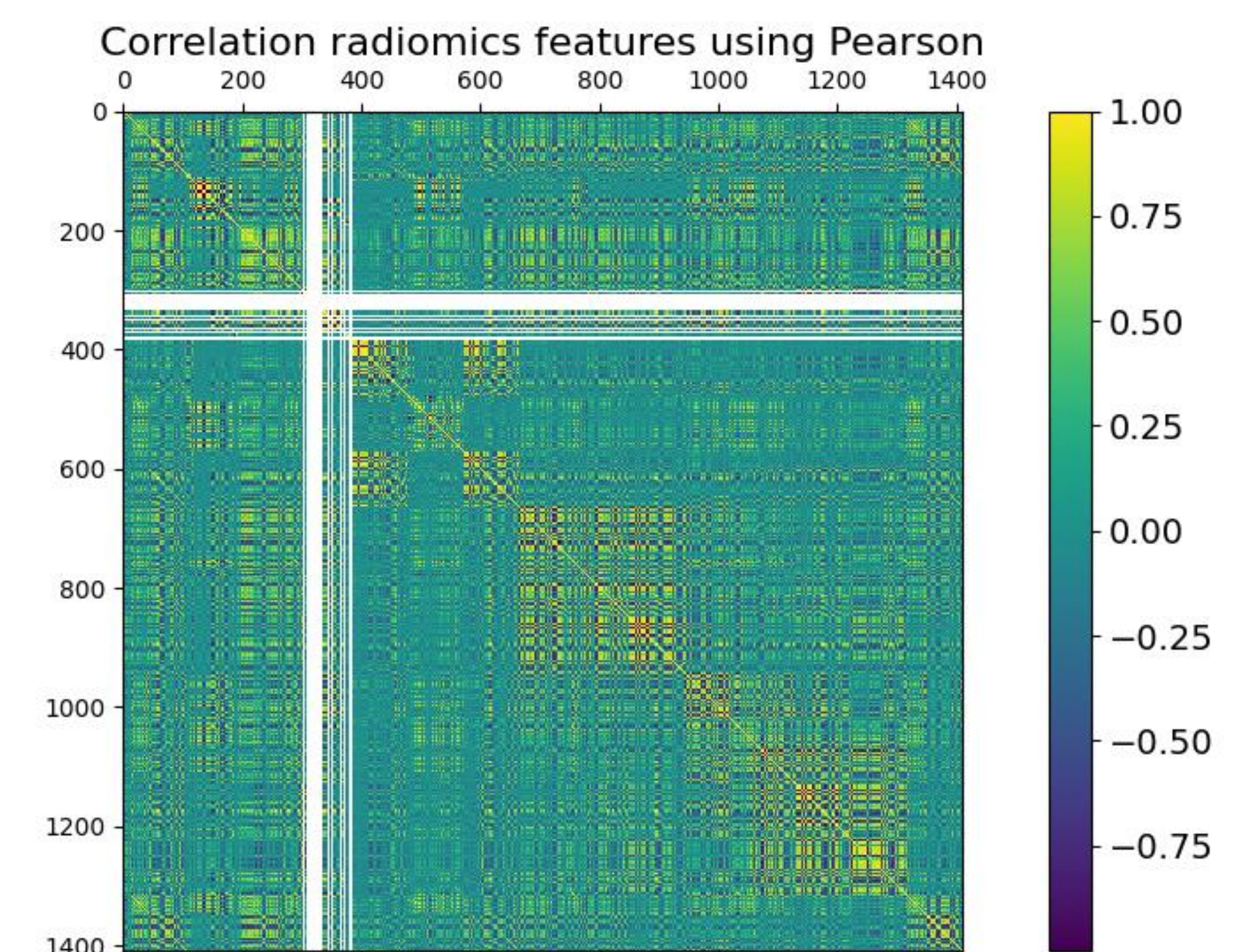


Fig. 2: Correlation matrix of 1410 features from 98 KiTS19 cases.

## CONCLUSIONS

In this contribution, we presented a web-based radiomics functionality for end users that enables the extraction of image feature for a tumor characterization. In doing so, we implemented and added a radiomics module to the existing StudierFenster framework.

## REFERENCES

1. Lee SH, Cho HH, Kwon J, Lee HY, Park H. Are radiomics features universally applicable to different organs?. Cancer Imaging. 2021 Dec;21(1):1-0.
2. Murray JM, Kaissis G, Braren R, Kleesiek J. Wie funktioniert Radiomics?. Der Radiologe. 2020 Jan;60(1):32-41.
3. Egger J, Wild D, Weber M, Bedoya CA, Karner F, Prutsch A, Schmied M, Dionysio C, Krobath D, Jin Y, Gsaxner C. Studierfenster: an Open Science Cloud-Based Medical Imaging Analysis Platform. Journal of Digital Imaging. 2022 Apr;35(2):340-55.
4. Heller N, Sathianathen N, Kalapara A, Walczak E, Moore K, Kaluzniak H, Rosenberg J, Blake P, Rengel Z, Oestreich M, Dean J. The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint arXiv:1904.00445. 2019 Mar 31.
5. Zhao B, Schwartz LH, Kris MG. Data from RIDER_Lung CT. The Cancer Imaging Archive. 2015.
6. Lindner L, Narnhofer D, Weber M, Gsaxner C, Kolodziej M, Egger J. Using synthetic training data for deep learning-based GBM segmentation. 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 Jul 23 (pp. 6724-6729).