Cancer Research

Computational Radiomics System to Decode the Radiographic Phenotype

Joost J.M. van Griethuysen[1,2,3], Andriy Fedorov[4], Chintan Parmar[1], Ahmed Hosny[1], Nicole Aucoin[4], Vivek Narayan[1], Regina G.H. Beets-Tan[2,3], Jean-Christophe Fillion-Robin[5], Steve Pieper[6], and Hugo J.W.L. Aerts[1,4]

Abstract

Radiomics aims to quantify phenotypic characteristics on medical imaging through the use of automated algorithms. Radiomic artificial intelligence (AI) technology, either based on engineered hard-coded algorithms or deep learning methods, can be used to develop noninvasive imaging-based biomarkers. However, lack of standardized algorithm definitions and image processing severely hampers reproducibility and comparability of results. To address this issue, we developed *PyRadiomics*, a flexible open-source platform capable of extracting a large panel of engineered features from medical images. *PyRadiomics* is implemented in Python and can be used standalone or using 3D Slicer. Here, we discuss the workflow and architecture of *PyRadiomics* and demonstrate its application in characterizing lung lesions. Source code, documentation, and examples are publicly available at www.radiomics.io. With this platform, we aim to establish a reference standard for radiomic analyses, provide a tested and maintained resource, and to grow the community of radiomic developers addressing critical needs in cancer research. *Cancer Res; 77(21); e104–7. ©2017 AACR.*

Introduction

Medical imaging is considered one of the top innovations that transformed clinical cancer care, as it significantly changed how physicians measure, manage, diagnose, and treat cancer. Imaging is able to noninvasively visualize the radiographic phenotype of a tumor before, during, and after treatment. Radiomics refers to the comprehensive and automated quantification of this radiographic phenotype using data characterization algorithms (1–3). Radiomics can quantify a large panel of phenotypic characteristics, such as shape and texture, potentially reflecting biologic properties like intra- and intertumor heterogeneities (4).

Radiomic technologies, based on artificial intelligence (AI) methods, are either defined using engineered hard-coded features, which often rely on expert domain knowledge, or on deep learning methods, which can learn feature representations automatically from data (5). The potential of radiomics has been shown across multiple tumor types, including brain, head and neck, cervix, and lung cancer tumors. Furthermore, these data,

extracted from MRI, PET, or CT images, were associated with several clinical outcomes, and hence, potentially provide complementary information for decision support in clinical oncology (1).

However, there is a lack of standardization of both feature definitions and image processing, which has been shown to have a substantial impact on the reliability of radiomic data (6–8). Furthermore, many studies use software developed in-house, often not shared with the public, which makes the reproduction and comparison of results difficult.

To address this issue, we developed a comprehensive open-source platform called *PyRadiomics*, which enables processing and extraction of radiomic features from medical image data using a large panel of engineered hard-coded feature algorithms. *PyRadiomics* provides a flexible analysis platform with both a simple and convenient front-end interface in 3D Slicer, a free open-source platform for medical image computing (9), as well as a back-end interface allowing automation in data processing, feature definition, and batch handling. *PyRadiomics* is implemented in Python, a language that has established itself as a popular open-source language for scientific computing, and can be installed on any system.

Here, we discuss the workflow and architecture of *PyRadiomics* and demonstrate its application in characterizing benign and malignant lung lesions. Source code, documentation, instruction videos (see Supplementary Videos S1 and S2), and examples are available at www.radiomics.io/pyradiomics.html. With this resource, we aim to establish a reference standard for radiomic analyses, provide tested and maintained open-source platforms, and raise awareness among scientists of the potential of radiomics technologies.

Platform

The *PyRadiomics* platform can extract radiomic data from medical imaging (such as CT, PET, MRI) using four main

[1]Department of Radiation Oncology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [2]Netherlands Cancer Institute (NKI), Amsterdam, the Netherlands. [3]GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, the Netherlands. [4]Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [5]Kitware, Clifton Park, New York. [6]Isomics, Cambridge, Massachusetts.

Corresponding Author: Hugo J.W.L. Aerts, Dana-Farber Cancer Institute, Brigham and Women's Hospital Harvard Medical School, 450 Brookline Ave, Boston, MA 02115. Phone: 617-525-7156; Fax: 617-582-6037; E-mail: Hugo_Aerts@dfci.harvard.edu

AACR

steps: (i) loading and preprocessing of the image and segmentation maps; (ii) application of enabled filters; (iii) calculation of features using the different feature classes; and (iv) returning results. See Fig. 1A for an illustration of this process.

Loading and preprocessing

In this step, medical images (e.g., CT, PET, MRI) and segmentation maps (e.g., performed by radiologist) are to be loaded into the platform. The large majority of image handling is done using *SimpleITK*, which provides a streamlined
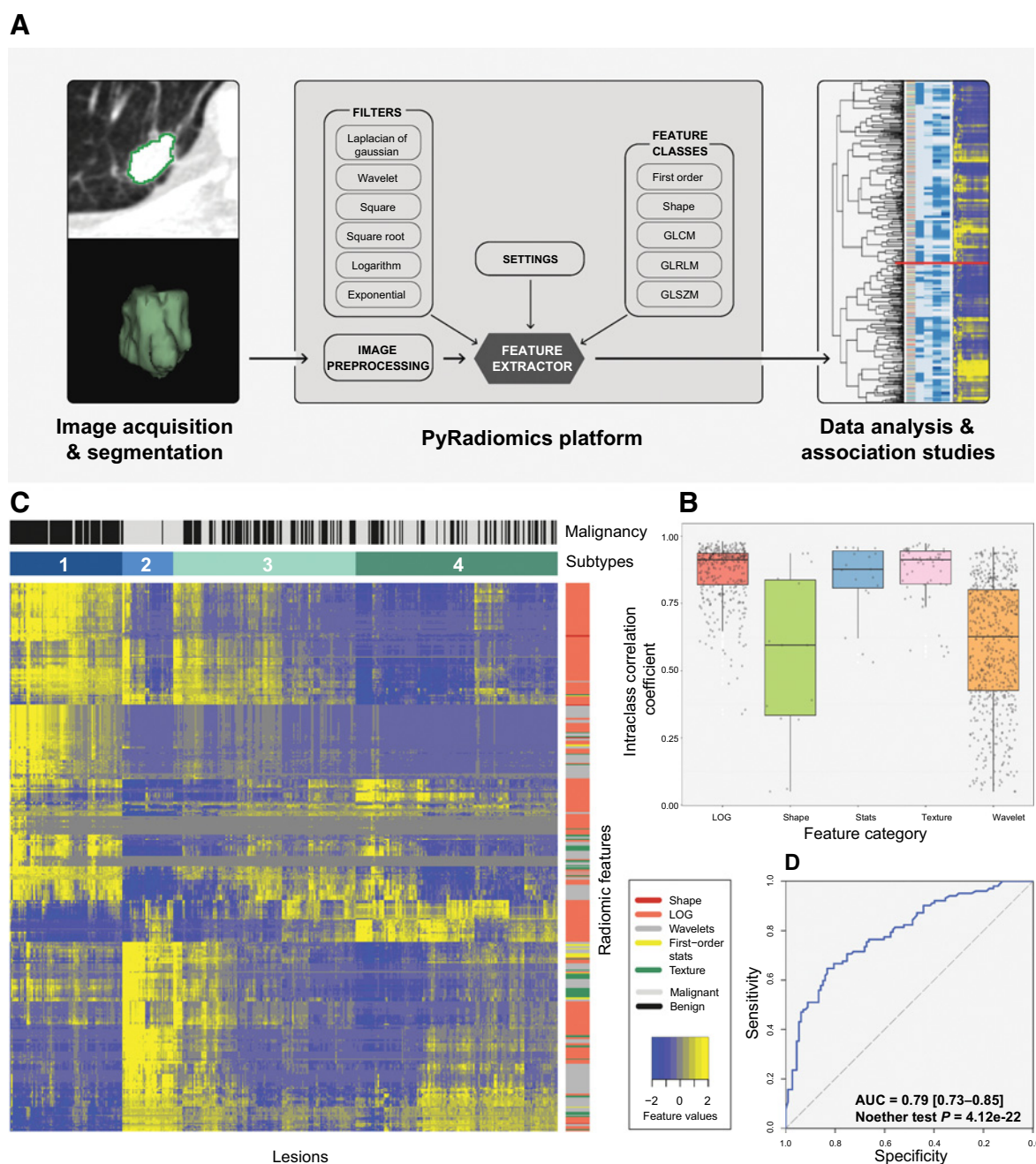
Figure 1.
A, Overview figure of the process of *PyRadiomics*. First, medical images are segmented. Second, features are extracted using the *PyRadiomics* platform, and third, features are analyzed for associations with clinical or biologic factors. **B,** Stability of radiomics features for variation in manual segmentations by expert radiologists. **C,** Heatmap showing expression values of radiomics features (rows) of 429 lesions (columns). Note the four subtypes that could be identified from the expression values and their associations with malignancy. **D,** Area under the curve (AUC) showing the performance of the multivariate biomarker to predict malignancy of nodules.

interface to the widely used open-source Insight Toolkit (ITK; ref. 10). This enables *PyRadiomics* to support a wide variety of image formats, while also ensuring that much of the low-level functionality and basic image processing is thoroughly tested and maintained. For texture and shape features, several resampling options are included to ensure isotropic voxels with equal distances between neighboring voxels in all directions.

Filtering

Features can be calculated on the original image or on images preprocessed using a choice of several built-in filters. These include wavelet and Laplacian of Gaussian (LoG) filters, as well as several simple filters, including square, square root, logarithm, and exponential filters. For the application of the wavelet and LoG filter, the platform makes use of the *PyWavelets* and *SimpleITK*, respectively. The remaining filters are implemented using *NumPy*.

Feature calculation

The platform contains five feature classes: a class for first-order statistics, a class for shape descriptors, and texture classes gray level cooccurrence matrix (11), gray level run length matrix (12, 13), and gray level size zone matrix (14). All statistic and texture classes can be used for feature extraction from both filtered and unfiltered images. Shape descriptors are independent from intensity values and therefore can only be extracted from unfiltered images. Feature extraction is supported for both single slice (2D) and whole volume (3D) segmentations.

PyRadiomics modules

Calculated features are stored and returned in an ordered dictionary. Every feature is identified by a unique name consisting of the applied filter, the feature class, and feature name. Besides the calculated features, this dictionary also contains additional information on the extraction, including current version, applied filters, settings, and original image spacing.

To enhance usability, *PyRadiomics* has a modular implementation, centered around the *featureextractor* module, which defines the feature extraction pipeline and handles interaction with the other modules in the platform. All feature classes are defined in separate modules. Furthermore, all are inherited from a base feature extraction class, providing a common interface. Finally, the platform contains two helper modules, *generalinfo* that provides additional extraction information included in the returned result and the *imageoperations* module that implements the functions used during image preprocessing and filters.

Aside from interactive use in Python scripts through the *featureextractor* module, *PyRadiomics* supports direct usage from the command line. There are two scripts available, *pyradiomics* and *pyradiomicsbatch*, for single-image and batch processing, respectively. For both scripts, an additional parameter file can be used to customize the extraction, and results can be directly imported into many statistical packages for analysis, including R and SPSS. In addition, a convenient front-end interface for *PyRadiomics* is provided as the "radiomics" extension within 3D Slicer. All code, including the Slicer extension, documentation, frequently asked questions, and instruction videos (see Sup-

plementary Videos S1 and S2) are available at www.radiomics.io/pyradiomics.html. In the Supplementary Information, detailed descriptions of feature definitions, dataset, and analyses can be found.

Case Study

In a case study, we demonstrated an application of *PyRadiomics* for lung lesion characterization to discriminate between benign and malignant nodules. We used the publicly available cohort of the Lung Image Database Consortium (15), which consists of diagnostic and lung cancer screening CT scans along with marked-up annotated lesions and per-lesion malignancy rating (i.e., if a nodule is benign or malignant) from experienced radiologists (Supplementary Methods S1). From 302 patients, we included 429 distinct lesions in our analysis, each with four volumetric segmentations and malignancy ratings. In total, 1,120 radiomic features (14 shape features, 19 first-order intensity statistics features, 60 texture features, 395 LoG features, and 632 wavelet features) were extracted from all four delineations of every lesion (Supplementary Methods S2–S4).

To assess the effect variations in the manual segmentations on radiomic feature values, we calculated the stability for each of the features extracted from four segmentations performed by expert radiologists. This stability was calculated using intraclass correlation coefficient (ICC; Fig. 1B). High stability (median ± SD: ICC > 0.8) was observed for LoG (ICC = 0.91 ± 0.11), first-order intensity statistics (ICC = 0.88 ± 0.13), and texture features (ICC = 0.91 ± 0.11), whereas shape (ICC = 0.60 ± 0.31) and wavelet (ICC = 0.63 ± 0.23) features showed moderate stability, which indicates their sensitivity toward delineation variability.

Selecting all features with high stability (ICC > 0.8) resulted in 535 radiomic features (5 shape features, 14 first-order intensity statistics features, 48 texture features, 310 LoG features, and 158 wavelet features). Figure 1C displays unsupervised clustering of the standardized expression values of the 535 stable radiomic features (rows) in 429 nodules (columns). We observed four distinct clusters of lesions with similar expression values. Comparing these clusters with lesion malignancy status, we observed significant difference between them ($P = 2.56e-24$, χ^2 test). Ninety-two percent ($n = 81$) of the samples of cluster S1 ($n = 88$) were malignant, whereas 95% ($n = 38$) of the samples of cluster S2 ($n = 40$) were benign. For clusters S3 ($n = 143$) and S4 ($n = 158$), the proportion of malignant samples was 54% ($n = 78$) and 34% ($n = 53$), respectively. These results demonstrate associations between imaging-based subtypes and malignancy status of lung lesions.

To evaluate the performance of a multivariate imaging biomarker, we divided the cohort into training ($n = 214$) and validation ($n = 215$). Using minimum redundancy maximum relevance, we selected 25 stable radiomic features from the training cohort (Supplementary Table S1). A multivariate biomarker was developed by fitting selected features into a random forest classifier, based on the training data. The biomarker demonstrated strong and significant performance to characterize lung nodules [area under the curve = 0.79 (0.73–0.85); Noether test, $P = 4.12e-22$] on the validation cohort (Fig. 1D). More details on features extraction and analysis methods are provided in the Supplementary Material.

Conclusions

PyRadiomics provides a flexible radiomic quantification platform, with a simple and convenient front-end interface in 3D Slicer, as well as a back-end interface within Python allowing automation in data processing, feature definition, and batch handling. By providing a tested and maintained open-source radiomics platform, we aim to establish a reference standard for radiomic analyses promoting reproducible science within the quantitative imaging field, to raise awareness among scientists of this platform to support their work, and to provide a practical go-to resource. By doing so, we hope to grow the community of radiomic technology developers to address critical needs in cancer research.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: J.J.M. van Griethuysen, A. Fedorov, V. Narayan, R.G.H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H.J.W.L. Aerts
Development of methodology: J.J.M. van Griethuysen, A. Fedorov, V. Narayan, S. Pieper, H.J.W.L. Aerts

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A. Hosny, H.J.W.L. Aerts
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.J.M. van Griethuysen, C. Parmar, A. Hosny, S. Pieper, H.J.W.L. Aerts
Writing, review, and/or revision of the manuscript: J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H.J.W.L. Aerts
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): V. Narayan, R.G.H. Beets-Tan, H.J.W.L. Aerts
Study supervision: A. Fedorov, R.G.H. Beets-Tan, H.J.W.L. Aerts
Other (development of the software described in the article): A. Fedorov, J.-C. Fillion-Robin

References

1. Aerts HJWL. The potential of radiomic-based phenotyping in precision medicine: a review. JAMA Oncol 2016;2:1636–42.
2. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
3. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48:441–6.
4. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 2012;12:323–34.
5. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
6. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol 2016;61:R150–66.
7. Orlhac F, Soussan M, Maisonobe J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. J Nucl Med 2014;55:414–22.
8. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. J Nucl Med 2012; 53:693–700.
9. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging 2012;30: 1323–41.
10. Johnson HJ, Mccormick MM. The ITK Software Guide Book 2: Design and Functionality Fourth Edition Updated for ITK version 4. 10. 2016.
11. Haralick R, Shanmugan K, Dinstein I. Textural features for image classification [Internet]. IEEE Trans Syst Man Cybern 1973;SMC-3:610–21.
12. Galloway MM. Texture analysis using gray level run lengths. Comput Gr Image Process 1975;4:172–9.
13. Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run lengths for texture analysis. Pattern Recognit Lett 1990;11:415–9.
14. Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. Pattern Recognition and Information Processing (PRIP) 2009;140–5. Available from: http://www.thibault.biz/Doc/Publications/TextureIndexesAndGrayLevelSizeZoneMatrixApplicationToCellNucleiClassification_PRIP2009.pdf.
15. Armato SG 3rd, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 2011;38:915–31.