



Original Article

MITK Phenotyping: An open-source toolchain for image-based personalized medicine with radiomics

Michael Götz^{a,c,*}, Marco Nolden^b, Klaus Maier-Hein^a^a Medical Image Analysis, Division Medical Image Computing; ^b Medical Imaging Components, Division Medical Image Computing, DKFZ Heidelberg, Germany; ^c National Center for Tumor Diseases (NCT), Heidelberg, Germany

ARTICLE INFO

Article history:

Received 4 July 2018

Received in revised form 26 October 2018

Accepted 29 November 2018

Available online 31 December 2018

Keywords:

Radiomics

Image phenotyping

Personalized medicine

Open source software

Radiology image processing

Machine learning

ABSTRACT

Radiomics – The extraction of quantitative features from radiologic images – shows increasing potential in contributing to modern personalized medicine approaches. MITK Phenotyping is an openly distributed radiomics framework implementing an exhaustive set of features, adhering to most recent international standards, and supporting a variety of different user interfaces and programming languages.

© 2018 The Authors. Published by Elsevier B.V. Radiotherapy and Oncology 131 (2019) 108–111 This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Radiologic imaging plays a crucial role in modern oncology. Images are commonly acquired at various time points during disease progression, providing in-vivo and spatially resolved insights into otherwise hidden processes and structures of the human body. For example, images are used for diagnostic purposes in order to detect and characterize lesions. They are used for treatment planning in order to locate tumorous tissue and structures of risk and during follow-up examinations to assess tumor growth and therapy outcome.

Radiomics is an approach that allows extracting well-defined quantitative features from underlying unstructured information in the raw images. The predictive information contained in these features sets can then complement the subjective evaluation by radiologists. The features can be either manually engineered or learned automatically. Other than the images themselves, they can be directly compared across different images and thus can be used to identify predictive patterns, i.e. phenotypes, to be used either alone or in combination with complementary information from other biomarkers. The high potential of such approaches has already been shown for different entities and modalities [1–3].

Radiomics research is often carried out in a clinical environment by medical experts who cannot invest sufficient resources

in the programming part of experiments. But it had been shown that already small differences in the implementation of the radiomics features can lead to significantly altered feature values [4,5]. An obvious solution would be the application of an existing tool, but many of these academic tools require knowledge in a specific programming language for correct usage, which accounts for additional barriers in this area. In addition, the heterogeneous landscape of radiomics tools and corresponding feature sets can be a severe problem when considering the reproducibility of published Radiomic models, trying to pool feature values from different sites, or applying existing radiomic models in the clinical routine.

The ‘Image Biomarker Standardisation Initiative (IBSI)’ [6] has taken on the problem of different implementations, aiming to precisely define a common subset of handcrafted features to ensure a consistent meaning and implementation between different toolkits and programming languages. It includes closed and open source frameworks like pyRadiomics [7], written in python, Moddicom [8], written in R, and IBEX [9], written in Matlab and C++. To build upon this effort and take it one step further, we developed MITK Phenotyping, an open-source Radiomic framework that we built with three main goals in mind: first, the inclusion of an exhaustive set of features with various feature types and parametrization, adhering to the standards set by IBSI. Second, the support of different user interfaces and programming languages with one implementation, including an easy-to-use graphical interface. Third, the open distribution of a stable and well-tested implementation.

* Corresponding author at: DKFZ Heidelberg, Im Neuenheimer Feld 581, 69120 Heidelberg, Germany.

E-mail address: m.goetz@dkfz-heidelberg.de (M. Götz).

Feature overview

MITK Phenotyping supports all image-related tasks in a Radiomics study within a single framework. Fig. 1 shows the image related steps of a common Radiomics pipeline.

Data input/output

MITK Phenotyping offers a unified interface for loading and saving different data types like images, meshes, pointsets, and others by combining and extending other public frameworks like ITK, VTK, and GDCM. The modular reader and writer solution of MITK automatically selects best-suited reader if multiple frameworks could be used. This structure also enables adding new reader/writer by providing a shared library. DICOM images are loaded using an advanced heuristic for reading images from different vendors and more complex data like Diffusion Weighted Images or DICOM RT-Structs.

Registration and resampling

Choosing the right registration and resampling algorithm is problem-specific and a topic of ongoing research. Therefore, we included MatchPoint [10] which provides unified access for a wide range of registration and resampling algorithms from libraries like ITK, Plastimatch, and others. The most common algorithms are already included and new algorithms can also be included in MatchPoint.

Segmentation

MITK Phenotyping includes basic segmentation algorithms like thresholding, boolean, and morphological operations. The main goal of these algorithms is to provide means for an automated cleaning and sanity checking of manually created segmentations as commonly performed in Radiomic studies [11]. In addition, the graphical user interface inherited from the MITK workbench offers several semi-automatic segmentation tools. Since they rely on data visualization and user interaction, they are not directly available from other programming languages.

Preprocessing

Commonly used preprocessing steps are made available by wrapping existing ITK methods or providing new implementations. This includes simple operations like smoothing, mathematical operations and more complex operations like LoG or Wavelet based image transformations. All filters are implemented at the C++ layer and therefore are available on all supported platforms.

Additional preprocessing methods can be added with only a few lines of code.

Feature extraction

At the heart of MITK Phenotyping is the extraction of feature values from images. The features are organized in feature classes which can be independently calculated. This includes common feature classes like first order features, volumetric statistics, and co-occurrence based features, but also some less known feature classes like curvature statistics, and grey level distance features. All features defined in the 'Image Biomarker Standardisation Initiative (IBSI)' are supported [6].

The calculated feature values are stored in a list of name-value pairs. The name of each feature is a unique string that encodes the class of feature and the actual feature names. By this, it is possible to use similarly named features within different feature classes, while being able to distinguish these features. The parameters used for the calculation of the features can be encoded in the feature code in order to enable reproducibility of experiments as part of good scientific practice.

Design principles

The main design goal of MITK Phenotyping is to provide a unified framework for different user applications. Depending on the given scenario, the framework can be accessed either by a graphical user interface (GUI, c.f. Fig. 2), a command line tool (CLI), a scripting language like python or a programming language like C++. This is achieved through a layer-based software design. The core is written in C++ and provides all functionality of MITK Phenotyping, while the outer layers provide different interfaces to other programming languages, like python. The software is designed as a part of the open-source MITK framework [12], which is commonly used in academia and in industry. This C++ framework provides unified building and accessing to other well-known libraries like GDCM, ITK or VTK, and extending these third-party frameworks with own functionalities.

Using MITK as the foundation for MITK Phenotyping provides access to a large, well-tested and documented set of common functionalities like image reading, writing, or registering. Automated testing and regular releases including MITK Phenotyping ensure a constantly high software quality. An automated testing environment reveals unexpected changes in the calculated feature values. All features are also tested for plausibility and cross-checked against other implementations, for example in collaboration with the IBSI. An extensive documentation provides descriptions and mathematical foundations of each feature, the software design and the used core software.

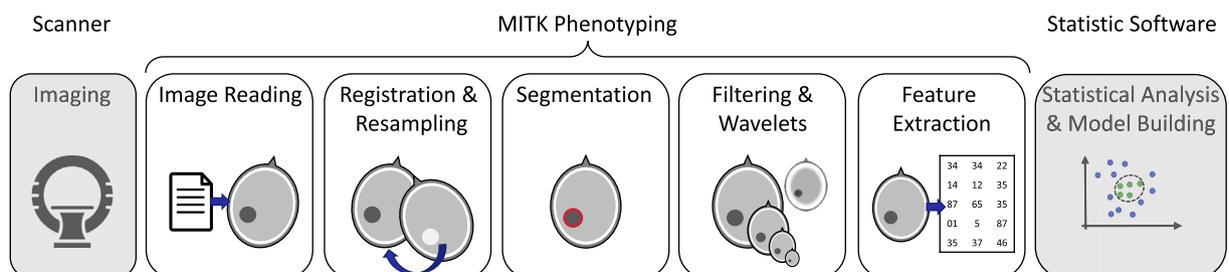


Fig. 1. Common Radiomic research pipeline. The aim of MITK Phenotyping is to support all image-related tasks from image acquisition and to model building which can be done in specialized languages like R, Matlab, or Python allowing to apply the same MITK Phenotyping pipeline without changes during exploration, model building or in clinical routine.

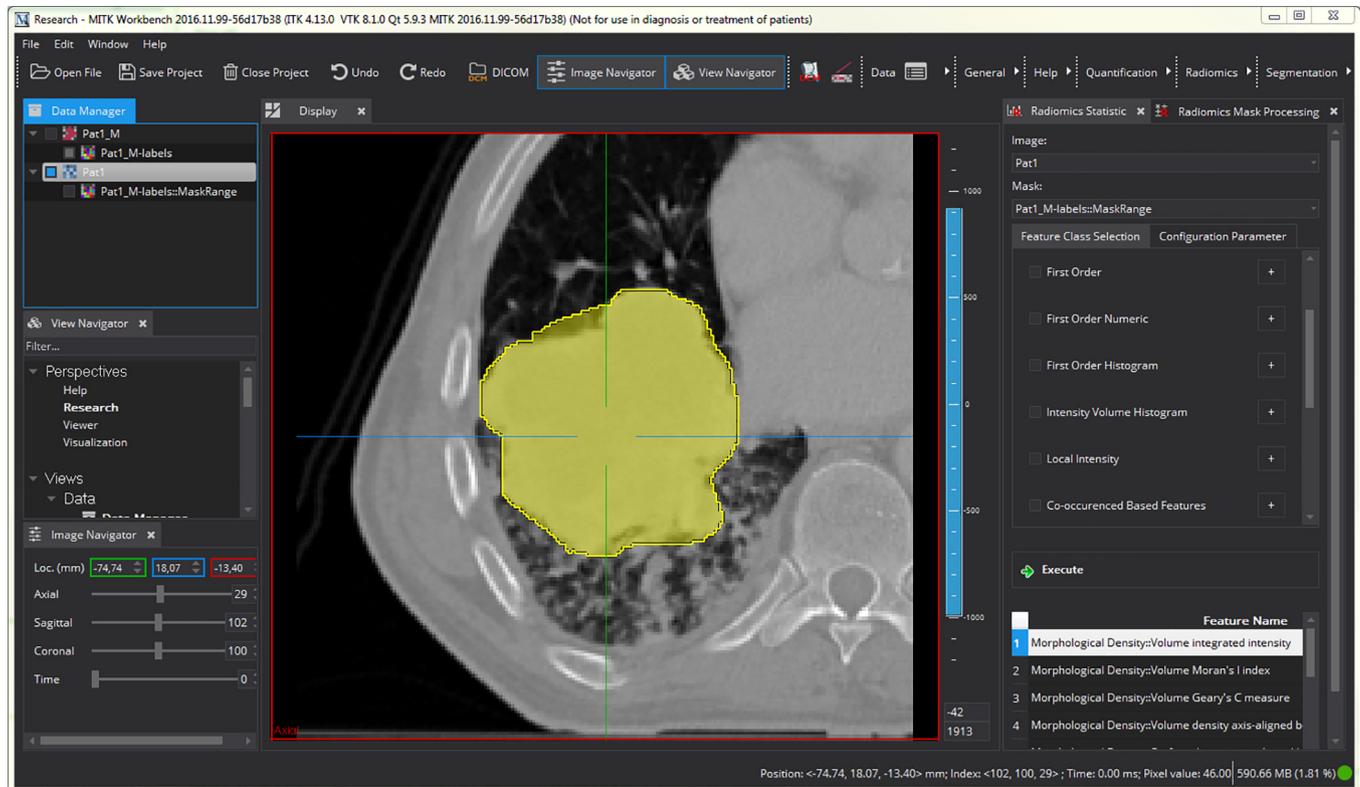


Fig. 2. Graphical interface of MITK Phenotyping. Besides the already known features of the MITK Workbench, like semiautomatic segmentation and 3D visualization, it is now possible to calculate Radiomic Features as well as performing common preprocessing steps. The radiomics-related functionality is also accessible via other interfaces like command line tools (CLI), scripting languages, and compiled languages like C++.

Discussion

MITK Phenotyping provides a stable, well tested, and extensible framework for the development of Radiomic analysis pipelines. It includes methods for most common preprocessing steps and all features defined in the IBSI initiative. Its layer-structure enables software access through different interfaces and programming languages, thus allowing straightforward translation from exploration to research and then to product development since a change of the programming language is possible without adapting to a new Radiomics framework. All information for a later reproduction of results can be easily stored for the sake of good scientific practice. The current version of MITK Phenotyping, as well as a more detailed documentation can be found at <http://www.mitk.org/Phenotyping>.

The Graphical User interface (GUI) of MITK Phenotyping is intended to remove existing barriers related to performing the steps of a Radiomics study. While it is still possible to use a programming language to access all relevant functions, this is not necessary anymore since everything can as well be done using the GUI. This is beneficial for users without or with limited experience in programming and for users who like to visualize the results of each step, for example to understand the implications of each preprocessing step.

The availability of the source code and a permissive open-source license allows the fast implementation of an individual graphical application based on MITK Phenotyping. This can be useful in different scenarios, for example to reduce the necessary work to perform a study with custom requirements. Another relevant setting would be the integration of a known radiomic signature within a graphical application to make it accessible to clinical users. Such applications could then be used to facilitate a broader

validation of signatures or even enhance the clinical decision process directly.

References

- [1] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5. <https://doi.org/10.1038/ncomms5006>.
- [2] Bickelhaupt S, Paech D, Kickingereder P, Steudle F, Lederer W, Daniel H, et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography. *J Magn Reson Imaging JMRI* 2017;46:604–16. <https://doi.org/10.1002/jmri.25606>.
- [3] Kickingereder P, Götz M, Muschelli J, Wick A, Neuberger U, Shinohara RT, et al. Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for stratifying anti-angiogenic treatment response. *Clin Cancer Res* 2016. <https://doi.org/10.1158/1078-0432.CCR-16-0702>.
- [4] Larue RTHM, Defraene G, Ruysscher DD, Lambin P, van Elmt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol* 2017;90:20160665. <https://doi.org/10.1259/bjr.20160665>.
- [5] Lv W, Yuan Q, Wang Q, Ma J, Jiang J, Yang W, et al. Robustness versus disease differentiation when varying parameter settings in radiomics features: application to nasopharyngeal PET/CT. *Eur Radiol* 2018. <https://doi.org/10.1007/s00330-018-5343-0>.
- [6] Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative – feature definitions. *CoRR* 2016. [abs/1612.07003](https://arxiv.org/abs/1612.07003).
- [7] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77. <https://doi.org/10.1158/0008-5472.CAN-17-0339>.
- [8] Dinapoli N, Alitto AR, Vallati M, Gatta R, Autorino R, Boldrini L, et al. Moddicom: a complete and easily accessible library for prognostic evaluations relying on image features. In: 2015 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC, Milan, IEEE; 2015. p. 771–4. <https://doi.org/10.1109/EMBC.2015.7318476>.
- [9] Lifei Zhang, Fried David V, Fave Xenia J, Hunter Luke A, Jinzhong Yang, Court Laurence E. ibex: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys* 2016;42:1341–53. <https://doi.org/10.1118/1.4908210>.

- [10] Floca R. MatchPoint: on bridging the innovation gap between algorithmic research and clinical use in image registration. In: Dössel O, Schlegel WC, editors. World Congr. Med. Phys. Biomed. Eng. Sept. 7 – 12 2009 Munich Ger.. Springer Berlin Heidelberg; 2010. p. 1105–8.
- [11] Vallières M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol* 2015;60:5471–96. <https://doi.org/10.1088/0031-9155/60/14/5471>.
- [12] Nolden M, Zelzer S, Seitel A, Wald D, Müller M, Franz AM, et al. The medical imaging interaction toolkit: challenges and advances: 10 years of open-source development. *Int J Comput Assist Radiol Surg* 2013;8:607–20. <https://doi.org/10.1007/s11548-013-0840-8>.