Kontos D, Summers RM, Giger M (2017) Special section guest editorial: radiomics and deep learning. J Med Imaging 4(4)

Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30(9):1234–1248

Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, Cook G (2020) Introduction to radiomics. J Nucl Med 61(4):488–495

Mesci A, Lucien F, Huang X, Wang EH, Shin D, Meringer M, Hoey C, Ray J, Boutros PC, Leong HS, Liu SK (2019) RSPO3 is a prognostic biomarker and mediator of invasiveness in prostate cancer. J Transl Med 17(1):1–11

Moon SH, Kim J, Joung JG, Cha H, Park WY, Ahn JS, Ahn MJ, Park K, Choi JY, Lee KH, Kim BT (2019) Correlations between metabolic texture features, genetic heterogeneity, and mutation burden in patients with lung cancer. Eur J Nucl Med Mol Imaging 46(2):446–454

O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, Boellaard R, Bohndiek SE, Brady M, Brown G, Buckley DL (2017) Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 14(3):169

Papanikolaou N, Matos C, Koh DM (2020) How to develop a meaningful radiomic signature for clinical use in oncologic patients. Cancer Imaging 20:1–10

Park JE, Kim HS (2018) Radiomics as a quantitative imaging biomarker: practical considerations and the current standpoint in neuro-oncologic studies. Nucl Med Mol Imaging 52(2):99–108

Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ (2015) Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. Front Oncol 5:272

Peerlings J, Woodruff HC, Winfield JM, Ibrahim A, Van Beers BE, Heerschap A, Jackson A, Wildberger JE, Mottaghy FM, DeSouza NM, Lambin P (2019) Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test–retest trial. Sci Rep 9(1):1–10

Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, Wahl RL, Kurland BF, Schwarz AJ, Gönen M, Zahlmann G (2015) Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. Stat Methods Med Res 24(1):27–67

Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, Bellomi M (2018) Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp 2(1):1–8

Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, Weigelt B, Vargas HA (2017) Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clin Radiol 72(1):3–10

Song J, Zhang Z (2019) Brain tissue segmentation and Bias field correction of MR image based on spatially coherent FCM with nonlocal constraints. Comput Math Methods Med 2019

Um H, Tixier F, Bermudez D, Deasy JO, Young RJ, Veeraraghavan H (2019) Impact of image preprocessing on the scanner dependence of multi-parametric MRI radiomic features and covariate shift in multi-institutional glioblastoma datasets. Phys Med Biol 64(16):165011

Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, Mak R, Aerts HJ (2016) Exploratory study to identify radiomics classifiers for lung cancer histology. Front Oncol 6:71

Yip SS, Aerts HJ (2016) Applications and limitations of radiomics. Phys Med Biol 61(13):R150

There are two types of learning methods in radiomics: supervised learning and unsupervised learning. In supervised learning, the model is developed depending on the input features and class variable, while in unsupervised learning, the model tries to reveal the correlations on the input features only. Often, radiomics features can be inspected along with other types of data that include proteomics, metabolomics, and others. In fact, inclusion of clinical features with radiomics variable is the new trend in for predictive modeling.

21.11 Conclusion

The requirement for radiomics signature often arises from unfulfilled clinical needs in detection of diseases, staging, treatment response prediction, and chance of survival. The methods applied in radiomics are generic to data science. They also considers the domain-related conditions in the case of smaller dataset that is very common to cancer imaging. Radiomics in cancer needs interdisciplinary strategy that includes knowledge of several fields like oncologists, radiologists, data science, and image processing.

Acknowledgement I would like to thank Dr. Swapan Kumar Basu, Department of Computer Science, Banaras Hindu University, India, for his help and cooperation in preparation of this chapter.

References

Aerts HJ (2016) The potential of radiomic-based phenotyping in precision medicine: a review. JAMA Oncol 2(12):1636–1642

Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006

Altman N, Krzywinski M (2018) The curse (s) of dimensionality. Nat Methods 15(6):399–400

Bedard PL, Hansen AR, Ratain MJ, Siu LL (2013) Tumour heterogeneity in the clinic. Nature 501 (7467):355–364

Chalkidou A, O'Doherty MJ, Marsden PK (2015) False discovery rates in PET and CT studies with texture features: a systematic review. PLoS ONE 10(5):e0124165

Choi ER, Lee HY, Jeong JY, Choi YL, Kim J, Bae J, Lee KS, Shim YM (2016) Quantitative image variables reflect the intratumoral pathologic heterogeneity of lung adenocarcinoma. Oncotarget 7(41):67302

Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H (2014) Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med Image Anal 18 (1):176–196

Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278(2):563–577

Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJ (2018) Artificial intelligence in radiology. Nat Rev Cancer 18(8):500–510

Kim DW, Jang HY, Kim KW, Shin Y, Park SH (2019) Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. Korean J Radiol 20(3):405

classified into three groups: filter, wrapper, and embedded methods. Filter methods ranked the features based on statistical measures. The selection of features has been done depending upon their ranks and characterized them by their computational efficiency, generalization, and robustness to overfitting. Filter methods can be univariate or multivariate. Univariate method provides scores for each of the features by neglecting the relationship between the features, while, in multivariate methods, consider the dependency between the features. In wrapper, identification of subsets with relevant and nonredundant features is performed and evaluates each of the subsets depending upon the performance of the model with the corresponding subset. These types of methods are computationally expensive. Embedded methods are computationally more efficient than wrapper. Both the selection and classification are performed concurrently in embedded methods and create models with more accuracy.

21.10 Training and Performance Measurement

After building the feature selection set of nonredundant, robust and relevant features, it is used for developing a model that generalizes the objective functions. Depending upon the continuous or discrete decision, variable different methods are used. If the decision variables are continuous, then linear regression, regression trees, or other regression methods are used otherwise, and several classification methods such as logistic regression, Naïve Bayes, support vector machines, decision trees, and random forests are used (Parmar et al. 2015).

Performance of the model is estimated for two nonintersect group of patients; among them, one is large and the other is comparative small dataset. The large one is used for the training purpose so that the relative error is minimized, while the other small one is used for validating the model. The small dataset is responsible for estimating more realistic performance metric. Therefore, the radiomics signature should be developed with many institutional data as this generalizes the objective function across all the institutions. But in reality, very few studies are found that considered data belong to different institutions (Kim et al. 2019).

Most of the studies have been done on the single institutional dataset only. In the case of small datasets, the validation method cannot work in right way and provide biasness in the model that results either over optimistic or pessimistic performance. To overcome this problem, cross-validation method is introduced that divides the dataset into k randomly equal subsets (in k-fold cross-validation) (Aerts 2016). One of the k-subsets is retained for validating the model, and the remaining k-1 are used for training the model. This process is repeated k times for every subset that is used once for validation. The method produces total k results, one for every iteration. Then, the average performance and standard deviation are determined that are the final measure, which indicates reproducibility and robustness.

In the case of machine learning, multiple models are compared and one model is chosen among them on the basis of performance. In general, the chosen model has lower standard deviation, which indicates more robustness and reproducibility.

the properties of brightness, color, size, and shape. Wavelet and Laplacian of Gaussian filters are applied on the images to enrich intricate patterns than often cannot quantify through the naked eye.

In deep learning pathway, deep features can be extracted through the use of deep neural networks in the training phase. Deep neural networks are more powerful than the classical machine learning methods in terms of mapping nonlinear representation. It works as a black-box but requires large training data, suffers from low interpretability, and is difficult to conceptualize.

21.9 Feature Selection

The extracted features from the images to develop a radiomic signature are high dimensional in nature. Often, the developed models based on these data suffer from model overfitting (Altman and Krzywinski 2018), which means the model successfully identify the data that belong to the training dataset with high accuracy but fails to classify the unknown data. The reduction of overfitting is a challenging problem in order to develop a robust model that can generalize the objective function and easily identify the unknown samples with high accuracy rate. Several procedures have been used for minimizing the overfitting issue. Usually the models train with large sized of data. There exists a reverse relationship between the sizes of the training datasets and overfitting. If the training dataset is very high then it is more prone to overfitting. In this type of case, dimensionality reduction is the only way to avoid overfitting problem and generalized the objective function by selecting minimum number of features through feature selection methods.

Feature selection methods recognize the redundant, unstable, and irrelevant extracted features and remove them to develop the model. In this way, a robust radiomic signature can be constructed with features that will be highly informative (Aerts et al. 2014; Wu et al. 2016; Aerts 2016; Parmar et al. 2015). The set of features can be evaluated in terms of temporal and spatial stability. Temporal stability is assessed for consistency through test-retest setting, whereas spatial stability is measured in terms robustness to variations in tumor segmentation (Aerts et al. 2014). In the case of split-validation scheme, the distribution of values of every features in the training set and test set is compared. Finally, keep only those features in radiomic signature that show no significance difference between the two and strictly remove the anomalies or system errors. This type of feature selection method is unsupervised in nature as it does not require the class variable during this process. After identifying the stable features, a correlation-based feature removing algorithm is used by developing a correlation heat map to eliminate the redundant features (Wu et al. 2016). Each block in the heat map represents a correlation coefficients. The blocks with correlation coefficient higher than a predefined value are identified and eliminated.

After removing a certain number of features, it is further required to reduce the dimensionality with more sophisticated methods and construct radiomics signature with a least number of features. There exist several methods for this task and are

Normalization of signals is bringing the signal intensities into a common range without altering differences in the range of values. It improves the stability and performance of the model and decreases the training duration (Um et al. 2019).

21.7 Lesion Segmentation

The next phase of the radiomics cascade is lesion segmentation, in which image segmentation is executed on the one or all the fragments of the target lesion. The basic methodology includes the tracing of lesion boundary manually that involves high variability that leads to derivation of unstable radiomic features. To build a stable model, more stable features are required. One of the ways to identify stable features is to execute radiological segmentations on the same lesion at least twice. Then, these two lesions are analyzed using correlation analysis to recognize the stable features (Peerlings et al. 2019).

Nowadays, researchers have been trying to automate the lesion segmentation process. The automation of lesion segmentation reduces the data variability and the efforts of radiologists that lead to analyze large dataset easier. However, verification of the final segmented outcomes by the certified radiologists is crucial. Recently, scientists preferred the delineation approach that resects physiologically distinct regions in the tumor based on functional imaging measurement (especially for MRI). From each of these distinct parts (or regions), radiomic features are extracted. In recent studies, it is also found that radiomic features of peritumoral spaces also produce unique information that reflects the treatment effects and disease outcomes.

21.8 Feature Extraction

One of the objectives of the radiomics is to perform complete evaluation of the lesion images using automated data extraction algorithms. These algorithms determine a several quantitative features that capture a large number of variety of phenotypic traits. These features are divided into two groups: semantic feature and agnostic features (Gillies et al. 2016). Semantic features are used for describing lesion by measuring diameters, volume, and morphology. However, agnostic features are the extracted quantitative measures that are not included in radiologists' lexicon. Agnostic features in the image data are first, second, or even higher-order statistical determinants, shape-based feature, and fractal features. These features are recognized by the computational procedures.

The first-order statistics is often use to represent voxel values by ignoring their spatial relations. The mean and median of the voxel values denote the overall tumor intensity or density or variations. There exist some features that represent shape- or location-specific features that identify the shape characteristics in 3-D space. However, the second-order statistics represents the spatial relationships among the voxels of image contrast. These are mainly referred to as texture features (Depeursinge et al. 2014), which are represented as repetitive elements or patterns on the surface with

important issue for getting meaningful outcomes. For an example, CT data are less varied than MRI data, which is the main reason for wide usage of CT images (Papanikolaou et al. 2020).

21.5 Image Acquisition

It is an impossible task to know every specific problem regarding the image acquisition a priori. Therefore, a standard protocol should be required for image acquisition for ensuring accuracy, repeatability, and reproducibility. As the level of standardization is not known initially, hence, a trial and error procedure is required to estimate such requirement. Primarily, it is advised to be more liberal for collecting images with a certain amount of heterogeneity. If it is not able to locate useful features in the collected data, then certain restrictions are imposed like company, sequence parameters, MRI field strength, etc. as the standardization.

Image acquisition protocol changes with time due to the competency of scanners that is getting improved due to upgradation and updation (Park and Kim 2018). The model is trained based on the group of heterogeneous patients' data obtained throughout the year, while testing part is done with the recently acquired exam data. In this way, the model becomes robust although the method needs a higher number of patient data. In the case of multicentric or in the event of multicompany single center studies, mainly two strategies are followed. Either use the model train with the data from one site (or company) and test with data from other sites (or companies) or use mixed data for performing training as well as validation. Effective strategies between these two are decided by basis of the highest performance and generalization.

21.6 Preprocessing of Images

Preprocessing in radiomics is required for removing noise and artifacts in the data. It enhances quality of the data to improve the outcome of the model (Kontos et al. 2017). Several filtering mechanisms have been used to reduce the noise in the images. However, these filters may cause information loss. Therefore, one should use these filters carefully. For example, motion correction methods can handle 4-D data like diffusion weighted images or dynamic contrast-enhanced MRI and CT images of lung by removing patient motion from the images collected from different modalities. But the use of filters and motion correction techniques should be used on the images as a last option in the image acquisition phase. MRI images often suffer from the heterogeneous spatial signals that cannot represent the properties of biological tissue but the patient body habitus, the geometric characteristics of external surface coils, the imperfect profile of rf pulse, and etc. Therefore, to remove such heterogeneity, bias field correction algorithms can be applied (Song and Zhang 2019). Normalization must be a part of preprocessing, while the collected images are from the different modalities in which an underlying heterogeneity exists in the data.

The size, quality, and diversity of the data are very important for justification of the specified hypothesis. In today's world, although imaging data are quiet easily available, still it is very difficult to collect and curate high-quality comprehensive image data. The complexity of the clinical problem is the major criteria to decide the minimum number of patients that are required to develop a radiomics signature (Kumar et al. 2012). For example, in disease detection problems, image contrast resolution is the sufficient factor to distinguish the abnormal tissue from the image data. As the problem is straightforward, it requires fewer patient's record. However, forecasting of treatment response or prediction of disease-free survival is more complex problems and these require a large number of patient records. According to Chalkidou et al. (2015), in two class classification problem, only 10–15 patient records are required for every features to train the model. If the developed radiomics model is based on small data, then there are high chances of over fitting and instability. The available data should be divided into 60:40 or 70:30 ratio for developing a radiomics model where 60–70% of the data are used for training purpose and rest 40–30% of data are used for validating/testing purpose.

To develop a radiomics signature, the sample size is an important criterion. Sample to develop radiomics signature defines the group of patients or representatives of the target population in terms of occurrence of the disease. The diseases that occur with lower possibility in a society require a large sample size to develop radiomics signature that is useful to prediction tool. In this context, a priori knowledge of the class size of different subtypes of the diseases is needed that will be classified by the analysis of radiomics signature. This would be helpful to observe whether the investigation produces the useful information or not. If the investigation is not useful, then sample or the group of patients should be improved with sufficient cases of each subtype of the disease. However, in practical scenario, several subtypes may often consist of a limited number of cases for radiomics analysis. Therefore, it is an important task to validate the classification model with the available dataset before analyzing the full performance analysis of the system.

21.4 Image Modalities for Radiomics

One of the main issues in image analysis is the absence of any standard guideline for image acquisition. Therefore, high variability exists in quantitative imaging (Raunig et al. 2015; O'Connor et al. 2017). Moreover, all the imaging modalities available in markets are developed only for generating images but not for assessing the images quantitatively. These imaging instruments are designed for radiologist to assess the images through their perceptions. Therefore, the quantitative assessment varies due to human perception. As the devices, made by the different companies, are not able to quantitative analysis of images, this leads to serious problems in treatment. The reason behind this fact is the quantitative measurement of images varies due to the usage of different hardware and software platform in the scanner. Furthermore, hybrid imaging systems are now available in the market, which produce images with varying contrasts. Hence, a careful selection of the imaging modalities is an

21.2 Radiomics Process Cascade

Radiomics is a medical image investigation process that consists of multiple phases (Fig. 21.1). At the primary stage, a group of oncologists define a clinical problem that the developed model is taken care of. They have also decided that what types of images are preferred to develop the radiomics model. After that, image acquisition and preprocessing part have been done optimally and it should be ensured that the quality of images is high enough for image analysis. If the size of the data is very large, then deep learning techniques may useful over the conventional machine learning processes. At last, these data are used for training and validation purposes.

From the next subsection, we are going to discuss every phase of the radiomics in detail.

21.3 Defining the Clinical Problem

In the primary phase of radiomics, an appropriate clinical problem is defined by a group of physician. The problem must fulfill some specified criteria. It should have addressed an unfulfilled need in cancer treatment, and the successful radiomic signature could perform a treatment plan for the patient.

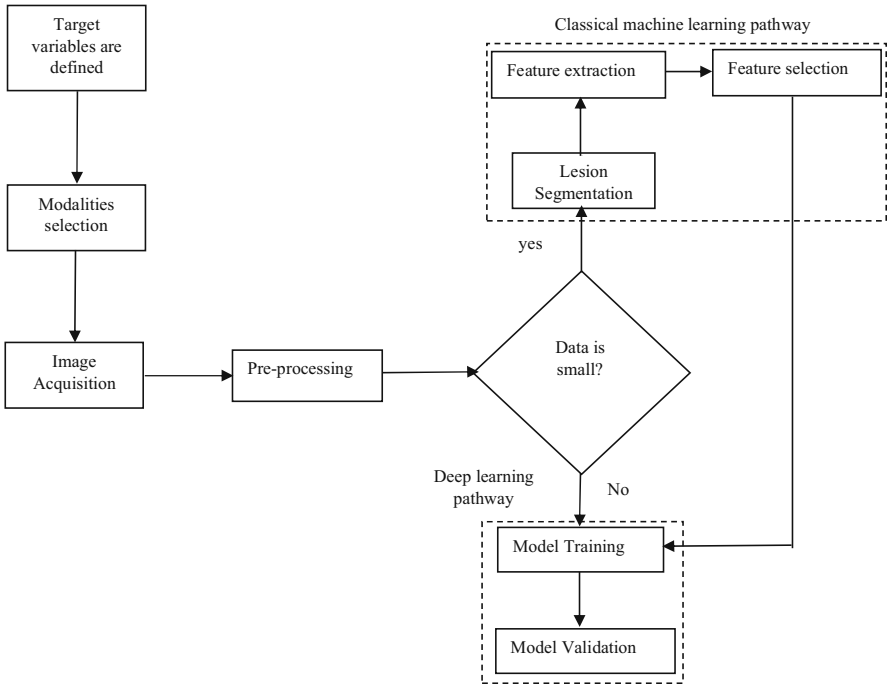

Fig. 21.1 Different phases of radiomics

medicine is not fully utilized as the medical practitioners have been drawn their conclusions by perceiving and recognizing image patterns and the past experiences. Hence, a degree of subjectivity and variability exists in medical image interpretation (Papanikolaou et al. 2020). To reduce the variability, the concept of quantitative evaluation of medical imaging has come into the picture to investigate the images and record more objective measurements. On the other hand, clinicians often profile tumors through invasive biopsy and molecular assay to capture its biological diversity and disease evolution (Bedard et al. 2013). However, repetitive tumor sampling is expensive as well as burdensome for patients. While biopsies often produce erroneous results as they capture the heterogeneity within a small portion of the tumor, radiomics captures the heterogeneity of entire volume of a tumor (Mayerhoefer et al. 2020). Researchers also found that the extracted features through radiomics are strongly correlated with heterogeneity at the cellular level (Moon et al. 2019; Choi et al. 2016). It also provides the expression of the genotype, the tumor microenvironment, and susceptibility of the cancer treatments (Mesci et al. 2019; Sala et al. 2017) and easily identifies the substantial phenotypic differences in tumors (Gillies et al. 2016; Hosny et al. 2018).

The extracted radiomic features may predict patterns of tumor evolution, progression, treatment responses, and survival chances of the patient with the combination of genomic, transcriptomic, or proteomic characteristics (Gillies et al. 2016; Sala et al. 2017; Yip and Aerts 2016). This prediction can be done through mining algorithms from the sufficiently large historical data. The datasets are mostly unstructured from different sources like, computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) (Rizzo et al. 2018), which may be combined with the clinical, laboratory, histological, and genomic data. The data mining algorithms accept the medical images and other unstructured data as input, carefully check the quality, and then find the optimal parameters as reliable and robust output (Mayerhoefer et al. 2020).

However, a lot of challenges exist in the multiphase radiomics processes. Nowadays, scanners are mainly designed for acquiring images rather than quantitative image analysis (Papanikolaou et al. 2020). Most of the acquired images lack desired microscopic resolutions; also, some are affected by ionizing radiation exposure. As a result, the outcome cumulatively varies significantly. Moreover, extraction of quantitative parameters is time-consuming process. Therefore, it is very clear that improvements are required regarding reproducibility and diagnostic accuracy, which are the main driving forces of radiomics. In this chapter, an overview of this multistep process is presented; specifically, the considerations and the best practices for developing radiomics signature are discussed that produce clinical value to cancer patients.

The rest of this chapter elaborates the processes of cancer cell motility and metastasis in detail and is organized as follows: Section 21.2 gives the idea of the workflow of radiomics; Sections 21.3–21.10 illustrate different phases of radiomics; Section 21.11 concludes the chapter.

Radiomics: Cropping More from the Images 21

Sounak Sadhukhan

Abstract

To reduce the variability in cancer diagnosis from radiological images, quantitative methods are introduced in radiology. In recent years, quantitative imaging biomarkers are augmented with the visual assessment by the radiologists that produces a better outcome. The advancement in machine learning processes gives an extra mileage to the rise of radiomics as it can extract quantitative information from the medical images. In this chapter, we focus on the methodology and the critical issues in the workflow to develop a clinically meaningful radiomics signature.

Keywords

Radiomics · Quantitative image analysis · Machine learning · Deep learning

21.1 Introduction

Radiomics is an emerging translational field of research that finds relationships between qualitative and quantitative features extracted from clinical images and data using characterization algorithms, with or without associated gene expression to support evidence-based clinical decision-making (Gillies et al. 2016).

In the past few decades, medical imaging has been used as a tool for detecting, staging, and evaluating treatment response of cancer. It has a great ability to visualize cancer's appearance, such as macroscopic tumor heterogeneity for primary as well as metastatic tumor in noninvasive way. However, its potential for precision

S. Sadhukhan (✉)
Department of Computer Science, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India

S. K. Basu et al. (eds.), *Cancer Diagnostics and Therapeutics*,
https://doi.org/10.1007/978-981-16-4752-9_21