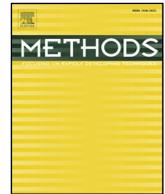




ELSEVIER

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Radiomics in neuro-oncology: Basics, workflow, and applications

Philipp Lohmann^{a,b,*}, Norbert Galldiks^{a,c,d}, Martin Kocher^{a,b,d}, Alexander Heinzl^{a,e},
 Christian P. Filss^{a,e}, Carina Stegmayr^a, Felix M. Mottaghy^{d,e,f}, Gereon R. Fink^{a,c}, N. Jon Shah^{a,g,h},
 Karl-Josef Langen^{a,d,e,g}

^a Institute of Neuroscience and Medicine (INM-3, -4, -11), Research Center Juelich, Wilhelm-Johnen-Str., 52428 Juelich, Germany

^b Department of Stereotaxy and Functional Neurosurgery, Center for Neurosurgery, Faculty of Medicine and University Hospital Cologne, Kerpener Str. 62, 50937 Cologne, Germany

^c Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany

^d Center of Integrated Oncology (CIO), Universities of Aachen, Bonn, Cologne and Duesseldorf, Kerpener Str. 62, 50937 Cologne, Germany

^e Department of Nuclear Medicine and Comprehensive Diagnostic Center Aachen (CDCA), RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany

^f Department of Radiology and Nuclear Medicine, Maastricht University Medical Center (MUMC+), P.Debyeilaan 25, 6229 HX Maastricht, P.O. Box 5800, 6202 AZ Maastricht, the Netherlands

^g JARA – BRAIN – Translational Medicine, Aachen, Germany

^h Department of Neurology, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany

ARTICLE INFO

Keywords:

Artificial Intelligence
 Machine learning
 Deep learning
 Glioma
 Brain metastases
 Multiparametric PET/MRI

ABSTRACT

Over the last years, the amount, variety, and complexity of neuroimaging data acquired in patients with brain tumors for routine clinical purposes and the resulting number of imaging parameters have substantially increased. Consequently, a timely and cost-effective evaluation of imaging data is hardly feasible without the support of methods from the field of artificial intelligence (AI). AI can facilitate and shorten various time-consuming steps in the image processing workflow, e.g., tumor segmentation, thereby optimizing productivity. Besides, the automated and computer-based analysis of imaging data may help to increase data comparability as it is independent of the experience level of the evaluating clinician. Importantly, AI offers the potential to extract new features from the routinely acquired neuroimages of brain tumor patients. In combination with patient data such as survival, molecular markers, or genomics, mathematical models can be generated that allow, for example, the prediction of treatment response or prognosis, as well as the noninvasive assessment of molecular markers. The subdiscipline of AI dealing with the computation, identification, and extraction of image features, as well as the generation of prognostic or predictive mathematical models, is termed radiomics. This review article summarizes the basics, the current workflow, and methods used in radiomics with a focus on feature-based radiomics in neuro-oncology and provides selected examples of its clinical application.

1. Introduction

The diagnosis of brain cancer is predominantly based on neuroimaging findings, and, ultimately, the histomolecular evaluation of tissue samples obtained from tumor resection or biopsy. For decades, mostly anatomical neuroimaging techniques such as contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI) have been used for brain tumor diagnostics, treatment planning, and follow-up. More recently, the increasing number of additional imaging parameters primarily derived from advanced MRI and amino acid PET, as well as technical developments such as the event of hybrid PET/CT and PET/MRI scanners, generate a large amount of complex neuroimaging data in patients with brain tumors.

A timely evaluation of this amount of diagnostic information that potentially can be implemented in clinical routine is costly and hardly feasible without considerable computer support. Here, methods from the emerging field of artificial intelligence (AI) offer new options to support clinicians further. In particular, AI provides the possibilities to partially or fully automate various steps within the diagnostic routine, so that especially time-consuming processes such as the manual detection and segmentation of lesions are performed by a computer and require only a final validation by a clinician. Furthermore, the speed of image processing and analysis can be enhanced using AI-based methods, thereby increasing productivity. Besides, the automated AI-based analysis of imaging data may help to increase the comparability of the obtained results as it is independent of the experience level of the

* Corresponding author at: Institute of Neuroscience and Medicine (INM-4), Research Center Juelich, Wilhelm-Johnen-Str., 52428 Juelich, Germany.

E-mail address: p.lohmann@fz-juelich.de (P. Lohmann).

<https://doi.org/10.1016/j.ymeth.2020.06.003>

Received 8 April 2020; Received in revised form 28 May 2020; Accepted 3 June 2020

1046-2023/ © 2020 Elsevier Inc. All rights reserved.

evaluating clinician.

Moreover, AI offers the potential to extract yet undiscovered features from routinely acquired images. Specifically, quantitative and semi-quantitative image features can be extracted from routinely acquired neuroimaging data, which are usually beyond human perception. Finally, subsets of these image features, combined with patient information such as survival data, molecular markers, or genomics, can be used to develop mathematical models that characterize the underlying brain tumor biology. Subsequently, these models can be used for essential clinical questions, e.g., the assessment of prognosis or treatment response, as well as the noninvasive diagnosis of molecular markers. The computation, identification, and extraction of image features, as well as the generation of prognostic or predictive mathematical models, is summarized under the term radiomics [1–5], a specialized application within the broad field of AI.

This review article summarizes the basics, the current workflow, and methods used in radiomics with a focus on feature-based radiomics in neuro-oncology and provides selected examples of its clinical application.

2. Radiomics

“*Images are more than pictures, they are data*”. This intuitive and precise definition by Robert Gillies and colleagues nicely illustrates the basic idea of radiomics [3]. Radiomics can be subdivided into feature-based and deep learning-based radiomics and is usually applied to routinely acquired imaging data, thereby allowing additional data analysis at a low cost. Since radiomics features are either mathematically predefined (feature-based radiomics) or generated from the data by training computational models (deep learning-based radiomics) using semi- or fully-automated methods from advanced statistics and machine learning, the results are more robust, reliable and reproducible compared to the somewhat subjective and reader dependent clinical evaluation. As mentioned above, the histomolecular characterization of brain tumors is of paramount importance for treatment decisions and prognostication. Radiogenomics, a subdiscipline of radiomics, aims at the non-invasive prediction predominantly of molecular markers, genetic mutations, or chromosomal aberrations, especially in situations where tissue samples are not accessible [6].

2.1. Feature-based radiomics

Feature-based radiomics utilizes a set of mathematically predefined features that are typically extracted from a segmented region-of-interest (ROI) or volume-of-interest (VOI). After feature extraction, a subset of relevant features is determined by feature selection algorithms to avoid overfitting and to generate robust and generalizable predictive models. It should be noted that improper feature selection can also lead to overfitting. For example, if the feature selection is performed based on a very homogenous dataset (same scanner and acquisition protocol), the selected features may not be relevant in other settings. Consequently, using heterogeneous datasets (different scanners and acquisition protocols) for feature selection reduces the probability for selecting only locally relevant features, hence, the risk of overfitting. As the calculated features are defined independently from the data, feature-based radiomics does not necessarily require big datasets, and the computation time is usually low. Also, since the features and their mathematical definitions are known, an interpretation of a biological analogy is conceivable. However, most of the extracted features are quite complex, i.e., a direct link between textural parameters and a physiological or pathophysiological meaning or interpretation is difficult or even not possible through human perception. The most critical processing steps in the feature-based radiomics workflow are summarized in the following.

2.1.1. Image pre-processing

The main objective of radiomics is the generation and use of quantitative features from medical images [1–5]. Furthermore, the extracted features and the generated models should be reproducible and generalizable, especially if data from different scanners and different acquisition protocols are used, which is often the case in clinical routine. To achieve these goals, several pre-processing steps have to be performed. Typical pre-processing steps for radiomics analyses include, but are not limited to, intensity normalization, spatial smoothing, spatial resampling, noise reduction, and corrections of MRI field inhomogeneities [7–9].

2.1.2. Tumor segmentation

For brain tumors, both gliomas and brain metastases, segmentation is usually performed manually on MR or CT images in clinical routine for the planning of radiotherapy or the volumetric assessment of therapy response. Manual, three-dimensional segmentation of brain tumors, including areas of contrast enhancement, necrosis, and peritumoral edema, is laborious and time-consuming. Since the segmented tumor forms the basis for the feature-based radiomics analysis, the contours directly influence the results of the radiomics analysis. To overcome this issue, several algorithms using machine learning techniques, including textural feature analysis and deep learning-based methods, are being developed currently and evaluated for the automated detection and segmentation of brain metastases as well as gliomas [10–15]. However, although these tools can already be used to support tumor segmentation, their reliability and added value still have to be proven before becoming part of clinical routine ultimately.

2.1.3. Feature extraction

Different types of quantitative features can be extracted from medical images, most of which reflect tumor heterogeneity. Although hundreds of features can be computed that slightly differ in the way they are mathematically defined, features are usually clustered in four subgroups:

- i) *Shape features* represent geometric relations and properties of the segmented ROI or VOI, such as the maximum diameter, maximum surface area, volume, compactness, or sphericity [7].
- ii) *First-order statistics features* or *histogram-based features* use the image intensity distribution represented by histograms that characterize the distribution of individual pixel or voxel intensity values within the segmented ROI or VOI, without considering their spatial orientation and relationship. Typical histogram-based features are mean, median, minimum, maximum, entropy (randomness), uniformity, asymmetry (skewness), or kurtosis (flatness) [2].
- iii) *Second-order statistics features* or *textural features* quantify the intratumoral heterogeneity. Textural features represent statistical relationships between intensity levels of neighboring pixels or voxels or groups of pixels or voxels. Textural features are not directly computed from the original image, but from different descriptive matrices that already encode specific spatial relations between pixels or voxels in the original image. The most prominent matrix for texture analysis is the gray level co-occurrence matrix (GLCM) [16]. The GLCM represents the number of times that two intensity levels occur in neighboring pixels or voxels within a specific distance along a fixed direction. Based on the GLCM, several textural features can be calculated, such as contrast, energy, correlation, homogeneity, variance, dissimilarity, cluster prominence, cluster tendency, and maximum probability. The neighborhood gray-level different matrix (NGLDM) corresponds to the difference of intensity levels between one voxel and its 26 neighbors in three dimensions. Among others, the textural features coarseness, contrast, and busyness can be computed from the NGLDM [17]. The gray-level run-length matrix (GLRLM) encodes the size of homogeneous runs for each image intensity [18]. Based on the GLRLM,

features such as short-run emphasis (SRE), long-run emphasis (LRE), low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), or run percentage (RP) can be extracted. Several other matrices encoding certain spatial relations between image intensities in the original image exist, from which a large number of textural features can be computed [16].

- iv) *Higher-order statistics features* are computed by statistical methods usually after the application of specific mathematical transformations (filters), e.g., for the identification of repeating patterns, noise suppression, edge enhancement, histogram-oriented gradients, or local binary patterns (LBP). The applied mathematical transformations or filters include Laplacian transforms of Gaussian-filtered images (Laplacian-of-Gaussian), wavelet or Fourier transforms, Minkowski functionals, or fractal analysis [9].

2.1.4. Feature selection

The extracted quantitative features are not equally important for the generation of a predictive or prognostic model from the imaging data. Most of the features are either constant, redundant, duplicated, irrelevant, highly correlated, or contribute to overfitting of the data, which renders the dependent model highly sensitive to image noise. Overfitting describes a methodological mistake in which a generated model corresponds too closely or even matches the analyzed set of (imaging) data. Overfitting results in perfect classification accuracy on the very dataset that has been used for training, but renders the model too specialized to classify new or additional imaging data or reliably predict future observations. One way to lower the risk of overfitting is to perform feature selection before model generation [5]. Several feature set reduction techniques are commonly used in radiomics, which can be classified into supervised and unsupervised feature selection techniques [19]. Unsupervised feature selection techniques do not consider class labels and simply aim at removing redundant features from the feature space. The two most commonly used methods for unsupervised feature selection in radiomics are principal component analysis (PCA) and cluster analysis [5]. Although these methods reduce the risk of overfitting, they usually do not result in the optimal feature subset.

In contrast, supervised feature selection techniques take the relation of the features with the class labels into consideration, leading to feature selection based on their contribution to the classification problem, i.e., the features that contribute most to differentiate between the groups are preferred. There are three commonly used methods for supervised feature set reduction:

- i) *Filter methods* (univariate methods) test the relation between

features and labels without considering their redundancy, i.e., the correlation with each other. Commonly used filter methods include but are not limited to the Wilcoxon rank sum test, the Fisher score, the Chi-squared score, the Student's *t*-test, or the minimum redundancy maximum relevance. Although filtering methods are commonly used for feature selection, correlations and interactions between the features are not considered. [20,21]

- ii) *Wrapper methods* (multivariate methods) overcome this limitation by investigating the entire feature space, taking into account feature relations to others in the dataset. Here, a predictive model is used to score the performance of a subset of features. Each new subset of features is evaluated based on the quality of the performance of a given algorithm. Wrapper models are computationally intensive because they aim to find the subset of features that results in the best performing model. Therefore, wrapper methods are also called greedy algorithms. Prominent examples for wrapper methods are forward feature selection, backward feature elimination, exhaustive feature selection, or bidirectional search. [20,21]
- iii) *Embedded methods* perform the feature selection process within the construction of the machine learning model itself, i.e., the best subset of features is selected during the training of the model. Thereby, embedded methods combine the advantages of filter and wrapper methods. Since the interaction of features is taken into consideration, embedded methods are more accurate than filter methods, faster than wrapper methods, and less prone to overfitting the data. Commonly used embedded methods are ridge regression, tree-based algorithms such as the random forest classifier, or the least absolute shrinkage and selection operator (LASSO) [20,21].

2.1.5. Model generation and evaluation

After feature selection, a mathematical model for the prediction of a known, underlying ground truth such as a particular genetic mutation or presence of tumor recurrence rather than treatment-related tissue changes can be built. In machine learning, several algorithms can be used to generate predictive models, depending on the purpose of the study. The most popular algorithms in radiomics are linear and logistic regression, decision trees (e.g., random forests), support vector machines, neural networks, and the Cox proportional hazards model in case of censored survival data. Generating the models and testing their performance on the same dataset is a methodological mistake that also leads to overfitting. A model that would simply repeat the labels of the training data would, of course, achieve a perfect classification result on the same data but will ultimately fail to predict anything useful on yet-unseen data. Consequently, in supervised machine learning, the available dataset is commonly subdivided into a training dataset and a

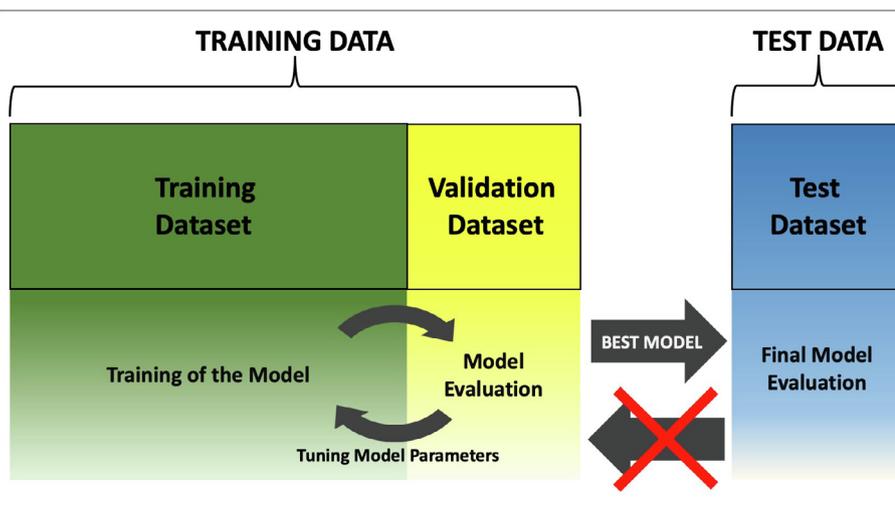
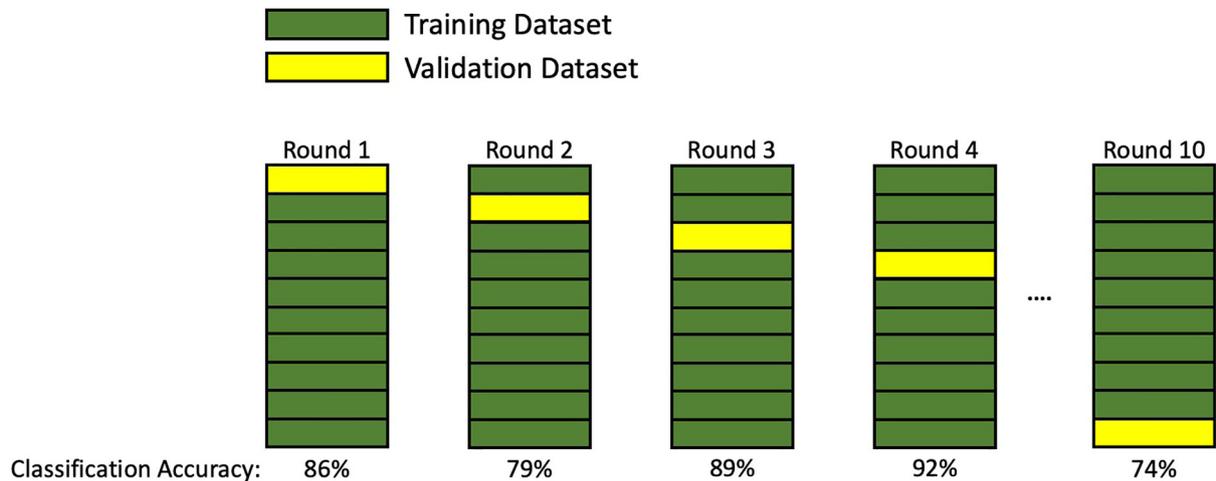


Fig. 1. Purpose of data splitting: The available dataset is divided into training data and test data. The training data is further subdivided into a training (green) and a validation dataset (yellow). The training dataset is used to train the model, whereas the validation dataset is used to evaluate the model. If the model performance in the validation dataset is not satisfying, model parameters can be tuned, and another model is trained on the training dataset. This process can be repeated until a model with an improved performance is generated. Finally, the best model is applied to the test dataset that ideally represents real-world data. Importantly, the test dataset should not be used for tuning the model parameters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Final Accuracy = Averaged Classification Accuracy (Round 1, Round 2,..., Round 10)

Fig. 2. 10-fold cross-validation: The training dataset (green) is partitioned into ten subsets of equal size, and one subset is retained as validation data (yellow), while the remaining nine datasets are used as training data. Afterwards, the process is repeated 10-times with each subset used once as validation data. The classification accuracy from each iteration is then averaged to produce a single estimation of model performance (final accuracy). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

validation dataset using stratified sampling to ensure that the training and validation dataset approximately have the same percentage of samples of each class as the complete set (Fig. 1). The latter is especially crucial for small or unbalanced datasets. After model training and validation, ideally, the model is finally applied to a third dataset, the so-called test dataset. In the best case, the test dataset represents the data the model would face when applied in clinical routine, i.e., data from different institutions, scanners, acquisition protocols, or segmentations. Therefore, the test dataset is the gold standard for evaluation of model performance, robustness, and reliability. Importantly, the test dataset should never be used for tuning the model parameters. However, especially for small datasets, statistical methods such as bootstrapping or cross-validation can be applied to estimate model performance without an external test dataset (Fig. 2).

2.2. Deep learning-based radiomics

Deep learning-based radiomics uses artificial neural networks that imitate the function of the human visual system and automatically extract high-dimensional features from the input images at different levels of scaling and abstraction. Deep learning-based radiomics is especially useful for pattern recognition or the classification of high-dimensional non-linear data [22].

The workflow is fundamentally different from the one described previously. In deep learning-based radiomics, different network architectures, i.e., stacks of linear and non-linear functions, such as convolutional neural networks (CNNs) or auto-encoders are used to find the most relevant features from the input data. A cascaded system of single-layer neural networks is trained to learn and identify structures in the image data that are relevant for classification without a prior definition or selection of the features [23]. Further combinations of these feature structures are then combined to generate features with a higher level of abstraction. Finally, the extracted features can either be further processed by the network for analysis and classification or leave the network and go through the process of model generation similar to the feature-based radiomics approach by using different classifiers such as decision trees, regression models, or support vector machines. Of note, since the networks generate and learn the essential features from

the data, feature selection is rarely performed, but in order to avoid overfitting, techniques such as regularization and dropout of learned connection weights are used. Since the features are highly correlated with the input data, deep learning-based radiomics usually requires larger datasets than feature-based radiomics, which limits its applicability in neuro-oncological research, where the number of available datasets is often limited. However, transfer learning is a method to overcome this limitation by applying neural networks that have already been trained for a different, albeit closely related task, e.g., a neural network that was trained with imaging data for the automated segmentation of gliomas might also be useful for the segmentation of brain metastases [24]. By using the prior knowledge of the network, the amount of data necessary to achieve consistent results, as well as the computational demand, is reduced.

3. Applications of radiomics in Neuro-Oncology

Radiomics in patients with brain tumors is mainly based on the analysis of conventional MRI. Several studies have investigated the usefulness of radiomics for the differentiation of treatment-related changes from tumor progression in patients with gliomas and brain metastases, which is a clinical question of considerable importance. Furthermore, several studies have also evaluated radiomics for the classification as well as the molecular characterization of brain tumors, which is of high relevance in light of the revised World Health Organization (WHO) classification of tumors of the central nervous system (CNS) from 2016. In this classification, the addition of molecular markers to histology is recommended to define brain tumor entities [25]. Furthermore, treatment strategies and decisions for glioma patients are also predominantly based on molecular markers. In particular, the isocitrate dehydrogenase (IDH) genotype [26], the loss of heterozygosity of the 1p/19q chromosome arms [27,28], and the O⁶-methylguanine-DNA methyltransferase (MGMT) promoter methylation status [29] are of high clinical relevance. For the determination of these markers, tissue samples obtained from tumor resection or biopsy are required. Therefore, a noninvasive method for the molecular characterization mentioned above is of scientific and clinical importance.

In the following, critical findings of selected studies using feature-

Table 1
Applications of feature-based MRI radiomics in neuro-oncology.

Study	No. of patients Total (Training/Test)	General purpose	MRI contrast(s)	No. of features		Feature selection method	Classification method	Performance (Training/Test)
				Initial	Final			
Cho et al. [30] (2018)	285 (285/0)	Determination of WHO grades in gliomas	T1, T1-CE, T2, FLAIR	468	5	mRMR	Logistic regression, SVM, random forest	0.90*/n.a. (AUC, averaged)
Hsieh et al. [32] (2017)	107 (107/0)	Determination of WHO grades in gliomas	T1-CE	20	9	Backward elimination	Logistic regression	0.94**/n.a. (AUC)
Tian et al. [33] (2018)	153 (153/0)	Determination of WHO grades in gliomas	T1, T1-CE, T2, DWI, PWI	520	28	SVM-based RFE	SVM	0.97***/n.a. (AUC)
Vamvakas et al. [34] (2019)	40 (40/0)	Determination of WHO grades in gliomas	T1, T1-CE, T2, FLAIR PWI, DTI, MRS	581	21	SVM-based RFE	SVM	0.96*/n.a. (AUC)
Shofty et al. [38] (2018)	47 (47/0)	Prediction of 1p/19q co-deletion status in gliomas	T1-CE, T2, FLAIR	152	9	PCA	Ensemble bagged tree	0.87*/n.a. (AUC)
Han et al. [39] (2018)	277 (184/93)	Prediction of 1p/19q co-deletion in gliomas	T2	647	64	Variance threshold	Random forest	0.89*/0.76 (AUC)
Zhou et al. [40] (2019)	538 (332/206) (197/84) 1p19q	Prediction of IDH genotype & 1p/19q co-deletion status in gliomas	T1-CE, FLAIR	127	15	Increase in prediction error of random forest	Random forest	0.92/0.92 (AUC, IDH)
Lu et al. [41] (2018)	284 (214/70)	Prediction of IDH genotype & 1p/19q co-deletion status in gliomas	T1-CE, T2, DWI	9,809	n.a.	Two sampled t-test with pooled variance estimate	SVM	0.67/0.69 (AUC, 1p19q) 96%/90% (accuracy, IDH)
Li et al. [47] (2018)	193 (133/60)	Prediction of MGMT promoter methylation status in gliomas	T1, T1-CE, T2, FLAIR	1,705	6	Boruta	Random forest	88%/80% (accuracy, 1p/19q) 0.95/0.88 (AUC)
Xi et al. [48] (2018)	118 (98/20)	Prediction of MGMT promoter methylation status in gliomas	T1, T1-CE, T2	1,665	36	Logistic regression with LASSO regularization	SVM	87%***/80% (accuracy)
Li et al. [51] (2017)	117 (78/39)	Evaluation of proliferative activity in gliomas	T2	431	9	SAM	ROC analysis	0.92/0.90 (AUC)
Kim et al. [60] (2019)	95 (61/34)	Differentiation of TRC from tumor progression in gliomas	T1-CE, FLAIR, PWI, DWI	6,472	12	Logistic regression with LASSO regularization	Generalized linear model	0.96/0.85 (AUC)
Hu et al. [61] (2011)	31 (15/16)	Differentiation of TRC from tumor progression in gliomas	T1-CE, T2, FLAIR, DWI, PWI	8	8	n.a.	SVM	n.a./0.94 (AUC, averaged)
Zhang et al. [65] (2018)	87 (87/0)	Differentiation of TRC from tumor progression in gliomas	T1, T1-CE, T2, FLAIR	285	5	CCC	RUSBoost	0.73***/n.a. (AUC)
Peng et al. [66] (2018)	66 (66/0)	Differentiation of TRC from local tumor relapse in BM	T1-CE, FLAIR	51	5	IsoSVM	IsoSVM	0.81***/n.a. (AUC)

* 5-fold cross-validation; ** leave-one out cross-validation; *** 10-fold cross-validation; AUC: area under the receiver operating characteristic curve; BM: brain metastases; CCC: concordance correlation coefficient; DTI: diffusion tensor imaging; DWI: diffusion-weighted imaging; FLAIR: fluid-attenuated inversion recovery; IDH: isocitrate dehydrogenase; LASSO: least absolute shrinkage and selection operator; MGMT: O⁶-methylguanine-DNA-methyltransferase; MRI: magnetic resonance imaging; mRMR: minimum redundancy maximum relevance algorithm; MRS: magnetic resonance spectroscopy; n.a.: not available; PCA: principal component analysis; PWI: perfusion-weighted imaging; RFE: recursive feature elimination; SAM: significance analysis of microarrays; SVM: support vector machine; T1: T1-weighted MRI; T1-CE: contrast-enhanced T1-weighted MRI; T2: T2-weighted MRI; TRC: treatment-related changes

Table 2
Applications of feature-based PET radiomics in neuro-oncology.

Study	No. of patients Total (Training/ Test)	General purpose	PET tracer	No. of features		Feature selection method	Classification method	Performance (Training/Test)
				Initial	Final			
Pyka et al. [35] (2015)	113 (113/0)	Determination of WHO grades in gliomas	FET	6	3	Discriminant function analysis	Discriminant function analysis	0.83/n.a. (AUC)
Lohmann et al. [42] (2018)	84 (84/0)	Prediction of IDH genotype in gliomas	FET	39	2	Fisher score	Logistic regression	80%*/n.a. (accuracy)
Kong et al. [49] (2019)	107 (71/36)	Prediction of MGMT promoter methylation status in gliomas	FDG	1,561	5	Logistic regression with LASSO regularization	SVM	0.94/0.86 (AUC)
Kong et al. [52] (2019)	123 (82/41)	Evaluation of proliferative activity in gliomas	FDG	1561	9	Logistic regression with LASSO regularization	SVM	0.88/0.76 (AUC)
Lohmann et al. [62] (2019)	35 (25/10)	Differentiation of TRC from tumor progression in gliomas	FET	44	3	Random forest-based RFE	Logistic regression	92%*/86% (accuracy)
Lohmann et al. [68] (2018)	52 (52/0)	Differentiation of TRC from local tumor relapse in BM	FET + MRI (T1-CE)	42	5	Wilcoxon rank sum	Logistic regression	0.86**/n.a. (AUC)

* 10-fold cross-validation; ** leave-one out cross-validation ; AUC: area under the receiver operating characteristic curve; BM: brain metastases; FDG: 2-[¹⁸F]-fluoro-2-deoxy-D-glucose; FET: O-(2-[¹⁸F]fluoroethyl)-L-tyrosine; IDH: isocitrate dehydrogenase; LASSO: least absolute shrinkage and selection operator; MGMT: O⁶-methylguanine-DNA-methyltransferase; MRI: magnetic resonance imaging; n.a.: not available; PET: positron emission tomography; RFE: recursive feature elimination; SVM: support vector machine; T1-CE: contrast-enhanced T1-weighted MRI; TRC: treatment-related changes

Table 3
Applications of deep learning-based MRI radiomics in neuro-oncology.

Study	No. of patients Total (Training/Test)	General purpose	MRI contrast(s)	No. of features		Feature selection method	Classification method	Performance (Training/Test)
				Initial	Final			
Yang et al. [36] (2018)	113 (90/23)	Determination of WHO grades in gliomas	T1-CE	n.a.	n.a.	n.a.	GoogLeNet	87%/95% (accuracy)
Eichinger et al. [43] (2017)	79 (59/20)	Prediction of IDH genotype in gliomas	DTI	101	101	Garson's algorithm ⁺	Neural network	0.92/0.95 (AUC)
Chang et al. [44] (2018)	259 (207/52)	Prediction of IDH genotype & 1p/19q status in gliomas	T1, T1-CE, T2, FLAIR	64	64	PCA ⁺	CNN	n.a./0.91* (AUC, IDH) n.a./0.88* (AUC, 1p/19q)
Li et al. [45] (2017)	119 (85/34)	Prediction of IDH genotype in gliomas	T1-CE, FLAIR	16,384	494	Fisher score	SVM	n.a./0.96 (AUC)
Chang et al. [44] (2018)	259 (207/52)	Prediction of MGMT promoter methylation status in gliomas	T1, T1-CE, T2, FLAIR	64	64	PCA ⁺	CNN	n.a./0.81 (AUC)
Korfiatis et al. [50] (2017)	155 (110/45)	Prediction of MGMT promoter methylation status in gliomas	T2	n.a.	n.a.	n.a.	ResNet	n.a./95% (accuracy)
Jang et al. [63] (2018)	78 (59/19)	Differentiation of TRC from tumor progression in gliomas	T1-CE	n.a.	n.a.	n.a.	CNN	0.72**/0.83 (AUC)
Li et al. [64] (2020)	84 (84/0)	Differentiation of TRC from tumor progression in gliomas	DTI	n.a.	n.a.	AlexNet	SVM	0.95**/n.a. (AUC)

⁺ Calculation of relative feature importance of input variables in the neural network; * 5-fold cross-validation; ** 10-fold cross-validation; AUC: area under the receiver operating characteristic curve; CNN: convolutional neural network; DTI: diffusion tensor imaging; FLAIR: fluid-attenuated inversion recovery; IDH: isocitrate dehydrogenase; MGMT: O⁶-methylguanine-DNA-methyltransferase; MRI: magnetic resonance imaging; n.a.: not available; PCA: principal component analysis; ResNet: residual neural network; SVM: support vector machine; T1: T1-weighted MRI; T1-CE: contrast-enhanced T1-weighted MRI; T2: T2-weighted MRI; TRC: treatment-related changes

based, as well as deep learning-based radiomics in patients with brain tumors are summarized. An overview of the discussed studies, the central processing steps, and the key results are summarized in Tables 1–3.

3.1. Determination of WHO grades in patients with newly diagnosed gliomas

Although the genotypic parameters are increasingly gaining importance for the classification of brain tumors according to the revised WHO classification [25], the initial phenotypical evaluation of tissue samples by histology still plays an important role. Therefore, several studies have investigated radiomics for the determination of WHO grades in patients with newly diagnosed gliomas.

A study from Cho and colleagues [30] investigated the role of feature-based radiomics for glioma grading. They used 285 datasets from the Brain Tumor Segmentation Challenge (BraTS) [31] comprising pre- and post-contrast T1-weighted, T2-weighted, and fluid-attenuated inversion recovery (FLAIR) MRI and extracted a total of 468 radiomics features. Feature selection was performed by the minimum redundancy maximum relevance algorithm followed by three different classifiers for model generation. Five features were selected, and the random forest classifier showed the highest area under the receiver operating characteristic (ROC) curve (AUC) of 0.92 for evaluation of glioma grade after 5-fold cross-validation. To account for differences in acquisition parameters and scanner types used in clinical routine, Hsieh and colleagues [32] used intensity invariant MRI features, so-called LBPs, for the differentiation of WHO grade IV glioblastoma from WHO grade II and III gliomas in 107 patients. Contrast-enhanced T1-weighted MRI scans from The Cancer Imaging Archive (TCIA) of the National Cancer Institute (NCI) were used. Imaging features were calculated and combined in a logistic regression classifier for tumor grading, and model performance was tested using leave-one out cross-validation. The LBP features achieved a high diagnostic accuracy of 93% (sensitivity, 97%; negative predictive value, 99%; AUC, 0.94). Interestingly, the LBP features were significantly better than conventional texture features (accuracy, 84%; sensitivity, 76%; negative predictive value, 89%, AUC, 0.89). However, this promising model lacks further validation on an external dataset.

Tian and colleagues [33] used conventional MRI, diffusion-weighted imaging (DWI), and perfusion parameters obtained from arterial spin labeling (ASL) for glioma grading in a group of 153 patients. Also, a support vector machine classifier was used for model generation, and model accuracy was determined by ROC analysis. The multiparametric model using texture information from all MR contrasts yielded the highest AUC of 0.97 for the classification of WHO grade III and IV and WHO grade II gliomas. A multiparametric approach was also conducted by Vamvakas and colleagues [34]. Textural as well as histogram features were calculated from conventional MRI and advanced MRI using perfusion-weighted imaging (PWI), diffusion tensor imaging (DTI), and magnetic resonance spectroscopy (MRS). In total, 581 features for each patient were obtained. After a support vector machine classifier was applied for feature selection and classification, twenty-one features were used for training and testing the classifier using leave-one-out cross-validation resulting in an accuracy of 96% (sensitivity, 96%; specificity, 96%; AUC, 0.96). The authors concluded that especially multiparametric MRI combined with radiomics analysis yields great potential for the pre-therapeutic classification of patients with gliomas.

Pyka and colleagues [35] demonstrated the feasibility of amino acid PET radiomics using the tracer O-(2-[¹⁸F]fluoroethyl)-L-tyrosine FET for the differentiation between WHO grade III and IV gliomas. Textural features calculated from the GLCM in combination with the FET PET-based metabolic tumor volume yielded a diagnostic accuracy of 85%.

Yang and colleagues [36] investigated the usefulness of deep learning-based radiomics using CNNs for glioma grading in a group of

113 patients based on contrast-enhanced T1-weighted MRI. Two commonly used CNNs (AlexNet and GoogLeNet) were explored. The pre-trained GoogLeNet showed the best classification accuracy in the test dataset (AUC, 0.94). This study demonstrates that transfer learning using pre-trained CNNs can be useful for clinical decision-making in patients with gliomas, especially when the number of available datasets is low.

3.2. Prediction of the IDH genotype and the 1p/19q co-deletion status in gliomas

The revised WHO classification of CNS tumors emphasizes the importance of molecular parameters such as the IDH genotype or the loss of heterozygosity of the 1p/19q chromosome arms for the classification of gliomas [25]. This results in two heterogeneous groups of tumors, which differ especially in their clinical prognosis [37]: IDH-mutant gliomas, typically astrocytomas (without 1p/19q co-deletion) or oligodendrogliomas (harboring 1p/19q co-deletion), usually have a better prognosis than IDH-wildtype gliomas, such as astrocytomas or glioblastomas. However, the determination of these parameters requires tissue samples. Therefore, several groups have investigated the role of radiomics for the non-invasive prediction of these important molecular markers in patients with gliomas.

Shofty and co-workers [38] achieved an accuracy of 87% after 5-fold cross-validation by an ensemble bagged tree classifier based on 152 radiomics features from conventional MRI for the determination of a chromosomal 1p/19q co-deletion. Similarly, Han and colleagues [39] used a random forest classifier to generate a radiomics signature based on conventional MRI of 277 patients with WHO grade II and III gliomas. The final model achieved an AUC of 0.89 in the training and 0.76 in the test cohort and outperformed a model solely based on clinical parameters. Interestingly, a combined model using radiomics and clinical parameters did not result in an improved prediction of the chromosomal 1p/19q co-deletion. Zhou and colleagues [40] used conventional MRI data from a large multicenter trial, including more than 500 patients for the prediction of the IDH genotype and the 1p/19q co-deletion status. After feature extraction, a random forest classifier was used for model generation, and the final model was tested in another set of MR images from TCIA and resulted in an AUC of 0.92 in the training and the test cohort. For a subset of patients with IDH mutant gliomas, another model was trained to predict the 1p/19q co-deletion status. The model achieved a moderate AUC of 0.67 in the training and 0.69 in the test cohort. Lu and colleagues [41] used conventional MRI data from 214 patients with malignant gliomas from TCIA. The test dataset consisted of 70 patients with MRI data from different institutions. The IDH genotype and the 1p/19q status could be predicted using a multi-level machine learning model based on a support vector machine classifier with diagnostic accuracies of 90% and 80%, respectively.

FET PET radiomics for the prediction of the IDH genotype was evaluated by Lohmann and colleagues [42] in a cohort of 84 glioma patients. Following feature selection, a simple two-parameter logistic regression model could be identified that achieved a diagnostic accuracy of 80% after 10-fold cross-validation. A subgroup of patients was examined on a high-resolution BrainPET scanner, which, due to its higher spatial resolution and sensitivity, was expected to yield better results in a radiomics analysis. Indeed, a subgroup analysis of the 28 patients examined on the BrainPET revealed the highest diagnostic accuracy of 86% after 10-fold cross-validation. The results have to be confirmed in a more extensive dataset.

LBP features extracted from T2-weighted MRI and DTI were used in a study by Eichinger and colleagues [43] for the training of a CNN with a single hidden layer. The IDH genotype could be predicted with high diagnostic accuracy of 95% in the test dataset. Chang and colleagues [44] trained a CNN on conventional MRI data from 259 glioma patients from TCIA and extracted the relevant imaging features by PCA. Both, the IDH genotype and the 1p/19q co-deletion status could be predicted

with high diagnostic accuracies of 94% and 92%, respectively. Li and colleagues [45] used post-contrast T1 and FLAIR MRI for the prediction of the IDH genotype in a cohort of 119 patients with WHO grade II gliomas. A CNN was used for tumor segmentation, and more than 16,000 features were generated for each case. Finally, a subset of 494 features was used in a support vector machine classifier yielding an AUC of 0.96. Interestingly, the authors compared the results with a feature-based radiomics approach, in which a significantly lower AUC of 0.86 was achieved for the prediction of the IDH genotype.

3.3. Determination of the MGMT promoter methylation status in glioblastoma

The methylation status of the MGMT promoter is of significant clinical value to predict the response to alkylating chemotherapy in patients with glioblastoma [29,46]. Consequently, several studies applied radiomics for the noninvasive determination of the MGMT promoter methylation status.

Li and colleagues [47] predicted the MGMT promoter methylation status in patients with glioblastoma from conventional MRI after calculating more than 1700 radiomic features. The final random forest classifier used a subset of six features and yielded an AUC of 0.88. Here, the combination of radiomic features with clinical parameters did not further improve the accuracy. Similarly, Xi and co-workers [48] extracted more than 1600 radiomics features from conventional MRI. The final model achieved a diagnostic accuracy of 87% in the validation dataset and 80% in the test dataset by utilizing a subset of 36 features.

Kong and colleagues [49] extracted more than 1,500 features from 2-¹⁸F-fluoro-2-deoxy-D-glucose (FDG) PET scans of 107 patients with primary glioma. After feature selection by sequential application of the Wilcoxon rank sum test and logistic regression with LASSO regularization, five features were used for model generation using a support vector machine classifier. The model achieved an AUC of 0.94 in the training and 0.86 in the test dataset.

Chang and colleagues [44] predicted the MGMT promoter methylation status based on conventional MRI from 259 glioma patients from TCIA using a CNN with an accuracy of 83%. Three different residual deep neural network (ResNet) architectures were compared in a study by Korfiatis and colleagues [50] for the prediction of the MGMT promoter methylation status. Conventional MRI data from 155 patients was used, and the authors reported that the ResNet50, i.e., a ResNet with a 50 layer architecture, showed the best performance with an accuracy of almost 95% in the test dataset.

3.4. Evaluation of the proliferative activity of gliomas

Ki-67 is a marker of tumor cell proliferation, and its level of expression in brain tumor tissue samples is usually assessed by immunohistochemistry obtained from tissue samples. The evaluation of the proliferative activity in gliomas using Ki-67 may be of value for neuropathological differential diagnosis and treatment decisions. Noninvasive methods for the assessment of Ki-67 expression levels are currently under investigation.

Li and colleagues [51] used T2-weighted MR images from 117 patients with WHO grade II and III gliomas for this purpose. Using Ki-67 immunohistochemistry as reference, a radiomics signature comprising nine features for the prediction of Ki-67 expression levels could be identified, yielding an accuracy of almost 89% in the validation dataset.

Kong and colleagues [52] used FDG PET radiomics for the prediction of Ki-67 expression levels. More than 1,500 radiomics features were extracted from FDG PET scans from 123 patients with primary glioma. Nine radiomics features were selected, and the final model was constructed by a support vector machine resulting in a moderate accuracy in the test cohort of 73%. However, the authors conclude that a combination of FDG PET and MRI radiomics might increase diagnostic accuracy.

3.5. Differentiation of treatment-related changes from tumor progression in patients with malignant gliomas

Early differentiation of neoplastic tissue from treatment-related changes such as pseudoprogression is of considerable clinical relevance in patients with malignant gliomas [53,54]. Pseudoprogression is characterized by the occurrence of a progressive enhancing lesion on MRI, usually within 12 weeks after radiotherapy or chemoradiation with concurrent temozolomide chemotherapy in patients with malignant gliomas, with a spontaneous improvement of MRI findings without any treatment change [54–56]. Although several studies have already demonstrated the value of amino acid PET as well as PWI for the diagnosis of pseudoprogression [54,57–59], radiomics might add additional information to improve the diagnostic accuracy further.

Kim and co-workers [60] used a multiparametric model incorporating conventional and advanced MRI (DWI and PWI) and achieved an AUC of 0.85 in an external test dataset. The multiparametric model was superior compared to other models based on single imaging contrasts. Similarly, Hu and colleagues [61] combined PWI and DWI with conventional MRI for the differentiation of pseudoprogression from tumor progression in a group of 31 patients. The generated support vector machine model used eight features and yielded an AUC of 0.94 (sensitivity, 90%; specificity, 94%). Both studies emphasize the value of multiparametric imaging for radiomics and machine learning approaches in neuro-oncology.

Lohmann and colleagues [62] used FET PET scans from 35 patients. After using two different segmentation methods to increase the number of available samples artificially, the final logistic regression model used three textural features and yielded a diagnostic accuracy of 92% in the training dataset and 86% in the test dataset. Again, the model has to prove its generalizability in a more extensive cohort study.

Jang and colleagues [63] developed a CNN that incorporated both clinical and imaging features from post-contrast MR images from 78 patients. The network achieved an AUC of 0.83 in the test dataset. Interestingly, the authors also investigated the performance of CNNs solely based on clinical or imaging features, which showed inferior performances in both cases. A novel feature learning method based on deep convolutional generative adversarial networks (DCGAN) and a CNN (AlexNet) termed DC-AL GAN was recently introduced by Li and colleagues [64]. The set of discriminative features extracted from DTI that was identified by the DC-AL GAN was used for classification by a support vector machine classifier and yielded a diagnostic accuracy of 92% (AUC, 0.95) in the validation dataset.

3.6. Differentiation of treatment-related changes from local tumor relapse in patients with brain metastases

Besides gliomas, the differentiation of tumor recurrence from treatment-related changes is also of great importance in patients with brain metastases, since patients with brain metastases are frequently treated with stereotactic radiosurgery. Consequently, radiation injuries (especially radiation necrosis) may not infrequently occur after radiosurgery that are often indistinguishable from local brain metastasis relapse using conventional MRI alone.

Zhang and colleagues [65] calculated 285 radiomics features from conventional MRI data of 87 patients after gamma knife radiosurgery. Since follow-up imaging data were also available, the authors could also investigate feature reproducibility and identify a feature subset with reproducible values. Additionally, changes in radiomics features (delta radiomics) were evaluated for the diagnosis of tumor progression. Finally, the model built by an ensemble classifier yielded a diagnostic accuracy of 73% in the validation dataset. Unfortunately, no test dataset was available. Peng and colleagues [66] extracted radiomics features from conventional MRI data of 66 patients. The diagnostic model was generated using the IsoSVM algorithm that performs both features selection and classification [67]. The model yielded an AUC of

0.81 (sensitivity, 87%; specificity, 65%). Interestingly, experienced radiologists could only correctly classify 73% (sensitivity, 97%; specificity, 19%).

Lohmann and colleagues [68] performed a radiomics analysis on multimodal FET PET/MRI data from 52 patients for the differentiation of treatment-related changes from brain metastases recurrence. After feature selection using the Wilcoxon rank sum test, logistic regression models were generated for the combined FET PET/MRI feature set as well as for each modality separately. The highest diagnostic accuracy with an AUC of 0.86 after leave-one out cross-validation was achieved by the combined FET PET/MRI model. Consequently, radiomics using multimodal imaging seems to achieve higher accuracies than either modality alone.

4. Conclusions

In summary, feature-based and deep learning-based radiomics is increasingly evaluated in the field of neuro-oncology. Radiomics should be considered as an additional source of diagnostic information that, especially in combination with clinical, histopathological, molecular, and conventional imaging parameters, has a great potential to significantly improve the diagnostics and management of patients with brain tumors.

Notwithstanding, most studies lack further validation of the generated models in more extensive, multi-institutional datasets, which currently hampers the applicability and potential translation of these machine learning techniques into clinical routine. Furthermore, there is a need for standardization of image acquisition and the radiomics analysis workflow. Although various software packages and algorithms for feature extraction and data evaluation such as pyradiomics [69], MaZda [70], and LifeX [71] are gradually becoming established, the workflow used in most studies is still complex, often utilizing highly specialized and self-tuned algorithms that prevent other researchers from understanding the details, not to mention reproducing the results. Consequently, the impact of different image acquisition protocols, image reconstruction, or preprocessing parameters on the radiomics signatures and the computed models warrants more consideration in future studies.

Feature-based and deep learning-based radiomics have a great potential to add important diagnostic information to many highly relevant clinical questions in brain tumor patients. Fortunately, efforts to overcome the limitations for clinical translation of this promising field of research are currently ongoing.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [Project number 428090865] (P.L. and N.G.).

References

- [1] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, C.M. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J. Aerts, Radiomics: extracting more information from medical images using advanced feature analysis, *Eur. J. Cancer* 48 (4) (2012) 441–446.
- [2] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, S. Sanduleanu, R. Larue, A.J.G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F.M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: the bridge between medical imaging and personalized medicine, *Nat. Rev. Clin. Oncol.* 14 (12) (2017) 749–762.
- [3] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures they are data, *Radiology* 278 (2) (2016) 563–577.
- [4] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, *Nat. Commun.* 5 (2014) 4006.
- [5] S. Rizzo, F. Botta, S. Raimondi, D. Oraggi, C. Fanciullo, A.G. Morganti, M. Bellomi, Radiomics: the facts and the challenges of image analysis, *Eur. Radiol. Exp.* 2 (1) (2018) 36.
- [6] M.A. Mazurowski, Radiogenomics: what it is and why it is important, *J. Am. Coll. Radiol.* 12 (8) (2015) 862–866.
- [7] V. Kumar, Y. Gu, S. Basu, A. Berglund, S.A. Eschrich, M.B. Schabath, K. Forster, H.J. Aerts, A. Dekker, D. Fenstermacher, D.B. Goldhof, L.O. Hall, P. Lambin, Y. Balagurunathan, R.A. Gatenby, R.J. Gillies, Radiomics: the process and the challenges, *Magn. Reson. Imaging* 30 (9) (2012) 1234–1248.
- [8] S.S. Yip, H.J. Aerts, Applications and limitations of radiomics, *Phys. Med. Biol.* 61 (13) (2016) R150–R166.
- [9] M. Zhou, J. Scott, B. Chaudhury, L. Hall, D. Goldhof, K.W. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert, R. Gatenby, Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches, *AJNR Am. J. Neuroradiol.* 39 (2) (2018) 208–216.
- [10] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T.L. Jones, T.R. Barrick, F.A. Howe, X. Ye, Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels, *Comput. Methods Programs Biomed.* 157 (2018) 69–84.
- [11] W. Deng, Q. Shi, K. Luo, Y. Yang, N. Ning, Brain tumor segmentation based on improved convolutional neural network in combination with non-quantifiable local texture feature, *J. Med. Syst.* 43 (6) (2019) 152.
- [12] A. Selvapandian, K. Manivannan, Fusion based Glioma brain tumor detection and segmentation using ANFIS classification, *Comput. Methods Programs Biomed.* 166 (2018) 33–38.
- [13] P. Prasanna, A. Karnawat, M. Ismail, A. Madabhushi, P. Tiwari, Radiomics-based convolutional neural network for brain tumor segmentation on multiparametric magnetic resonance imaging, *J Med Imaging (Bellingham)* 6(2) (2019) 024005.
- [14] O. Charron, A. Lallemand, D. Jarnet, V. Noblet, J.B. Clavier, P. Meyer, Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network, *Comput. Biol. Med.* 95 (2018) 43–54.
- [15] E. Grovik, D. Yi, M. Iv, E. Tong, D. Rubin, G. Zaharchuk, Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI, *J. Magn. Reson. Imaging* 51 (1) (2020) 175–182.
- [16] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Trans. Syst. Man. Cybern.* 3 (6) (1973) 610–621.
- [17] M. Amadasun, R. King, Textural features corresponding to textural properties, *IEEE Trans. Syst. Man Cybern.* 19 (5) (1989) 1264–1274.
- [18] D.-H. Xu, A.S. Kurani, J.D. Furst, D.S. Raicu, Run-length encoding for volumetric texture, *Conference Proceedings from The 4th IASTED International Conference on Visualization, Imaging and Image Processing: VIP* (2004).
- [19] Y. Zhang, A. Oikonomou, A. Wong, M.A. Haider, F. Khalvati, Radiomics-based prognosis analysis for non-small cell lung cancer, *Sci. Rep.* 7 (2017) 46349.
- [20] V. Parekh, M.A. Jacobs, Radiomics: a new application from established techniques, *Expert Rev. Precise Med. Drug Dev.* 1 (2) (2016) 207–226.
- [21] M. Kuhn, K. Johnson, *Applied Predictive Modeling*, Springer, New York, 2013.
- [22] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in Neural. Inform. Process. Syst.* (2012).
- [23] Y.J. Cha, W.I. Jang, M.S. Kim, H.J. Yoo, E.K. Paik, H.K. Jeong, S.M. Yoon, Prediction of response to stereotactic radiosurgery for brain metastases using convolutional neural networks, *Anticancer Res.* 38 (9) (2018) 5437–5445.
- [24] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, A Survey on Deep Transfer Learning, *arXiv e-prints*, 2018.
- [25] D.N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W.K. Cavenee, H. Ohgaki, O.D. Wiestler, P. Kleihues, D.W. Ellison, The 2016 world Health organization classification of tumors of the central nervous system: a summary, *Acta Neuropathol.* 131 (6) (2016) 803–820.
- [26] C. Hartmann, B. Hentschel, W. Wick, D. Capper, J. Felsberg, M. Simon, M. Westphal, G. Schackert, R. Meyermann, T. Pietsch, G. Reifenberger, M. Weller, M. Loeffler, A. von Deimling, Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas, *Acta Neuropathol.* 120 (6) (2010) 707–718.
- [27] M.C. Chamberlain, D. Born, Prognostic significance of relative 1p/19q codeletion in oligodendroglial tumors, *J. Neurooncol.* 125 (2) (2015) 249–251.
- [28] C.K. Speirs, J.R. Simpson, C.G. Robinson, T.A. DeWees, D.D. Tran, G. Linette, M.R. Chicoine, R.G. Dacey, K.M. Rich, J.L. Dowling, E.C. Leuthardt, G.J. Zipfel, A.H. Kim, J. Huang, Impact of 1p/19q codeletion and histology on outcomes of anaplastic gliomas treated with radiation therapy and temozolomide, *Int. J. Radiat. Oncol. Biol. Phys.* 91 (2) (2015) 268–276.
- [29] M.E. Hegi, L. Liu, J.G. Herman, R. Stupp, W. Wick, M. Weller, M.P. Mehta, M.R. Gilbert, Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity, *J. Clin. Oncol.* 26 (25) (2008) 4189–4199.
- [30] H.H. Cho, S.H. Lee, J. Kim, H. Park, Classification of the glioma grading using radiomics analysis, *PeerJ* 6 (2018) e5982.
- [31] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M.A. Weber, T. Arbel, B.B. Avants, N. Ayache, P. Buendia, D.L. Collins, N. Cordier, J.J. Corso, A. Criminisi, T. Das, H. Delingette, C. Demiralp, C.R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K.M. Iftekharuddin, R. Jena, N.M. John, E. Konukoglu, D. Lashkari, J.A. Mariz, R. Meier, S. Pereira, D. Precup, S.J. Price, T.R. Raviv, S.M. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H.C. Shin, J. Shotton, C.A. Silva, N. Sousa, N.K. Subbanna, G. Szekely, T.J. Taylor, O.M. Thomas, N.J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D.H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, K. Van Leemput, The multimodal brain tumor image segmentation benchmark (BRATS), *IEEE Trans. Med. Imaging* 34 (10) (2015) 1993–2024.
- [32] K. Li-Chun Hsieh, C.Y. Chen, C.M. Lo, Quantitative glioma grading using transformed gray-scale invariant textures of MRI, *Comput. Biol. Med.* 83 (2017)

- 102–108.
- [33] Q. Tian, L.F. Yan, X. Zhang, X. Zhang, Y.C. Hu, Y. Han, Z.C. Liu, H.Y. Nan, Q. Sun, Y.Z. Sun, Y. Yang, Y. Yu, J. Zhang, B. Hu, G. Xiao, P. Chen, S. Tian, J. Xu, W. Wang, G.B. Cui, Radiomics strategy for glioma grading using texture features from multiparametric MRI, *J. Magn. Reson. Imaging* 48 (6) (2018) 1518–1528.
- [34] A. Vamvakas, S.C. Williams, K. Theodorou, E. Kapsalaki, K. Fountas, C. Kappas, K. Vassiou, I. Tsougas, Imaging biomarker analysis of advanced multiparametric MRI for glioma grading, *Phys. Med. Biol.* 60 (2019) 188–198.
- [35] T. Pyka, J. Gempt, D. Hiob, F. Ringel, J. Schlegel, S. Bette, H.J. Wester, B. Meyer, S. Forster, Textural analysis of pre-therapeutic [18F]-FET-PET and its correlation with tumor grade and patient survival in high-grade gliomas, *Eur. J. Nucl. Med. Mol. Imaging* 43 (1) (2016) 133–141.
- [36] Y. Yang, L.F. Yan, X. Zhang, Y. Han, H.Y. Nan, Y.C. Hu, B. Hu, S.L. Yan, J. Zhang, D.L. Cheng, X.W. Ge, G.B. Cui, D. Zhao, W. Wang, Glioma grading on conventional MR images: a deep learning study with transfer learning, *Front. Neurosci.* 12 (2018) 804.
- [37] J.E. Eckel-Passow, D.H. Lachance, A.M. Molinaro, K.M. Walsh, P.A. Decker, H. Sicotte, M. Pekmezci, T. Rice, M.L. Kosel, I.V. Smirnov, G. Sarkar, A.A. Caron, T.M. Kollmeyer, C.E. Praska, A.R. Chada, C. Halder, H.M. Hansen, L.S. McCoy, P.M. Bracci, R. Marshall, S. Zheng, G.F. Reis, A.R. Pico, B.P. O'Neill, J.C. Buckner, C. Giannini, J.T. Huse, A. Perry, T. Tihan, M.S. Berger, S.M. Chang, M.D. Prados, J. Wiemels, J.K. Wiencke, M.R. Wrensch, R.B. Jenkins, Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors, *N. Engl. J. Med.* 372 (26) (2015) 2499–2508.
- [38] B. Shofity, M. Artzi, D. Ben Bashat, G. Liberman, O. Haim, A. Kashanian, F. Bokstein, D.T. Blumenthal, Z. Ram, T. Shahar, MRI radiomics analysis of molecular alterations in low-grade gliomas, *Int. J. Comput. Assist. Radiol. Surg.* 13 (4) (2018) 563–571.
- [39] Y. Han, Z. Xie, Y. Zang, S. Zhang, D. Gu, M. Zhou, O. Gevaert, J. Wei, C. Li, H. Chen, J. Du, Z. Liu, D. Dong, J. Tian, D. Zhou, Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas, *J. Neurooncol.* 140 (2) (2018) 297–306.
- [40] H. Zhou, K. Chang, H.X. Bai, B. Xiao, C. Su, W.L. Bi, P.J. Zhang, J.T. Senders, M. Vallieres, V.K. Kavouridis, A. Boaro, O. Arnaout, L. Yang, R.Y. Huang, Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas, *J. Neurooncol.* 142 (2) (2019) 299–307.
- [41] C.F. Lu, F.T. Hsu, K.L. Hsieh, Y.J. Kao, S.J. Cheng, J.B. Hsu, P.H. Tsai, R.J. Chen, C.C. Huang, Y. Yen, C.Y. Chen, Machine learning-based radiomics for molecular subtyping of gliomas, *Clin. Cancer Res.* 24 (18) (2018) 4429–4436.
- [42] P. Lohmann, C. Lerche, E.K. Bauer, J. Steger, G. Stoffels, T. Blau, V. Dunkl, M. Kocher, S. Viswanathan, C.P. Filss, C. Stegmayr, M.I. Ruge, B. Neumaier, N.J. Shah, G.R. Fink, K.J. Langen, N. Galldiks, Predicting IDH genotype in gliomas using FET PET radiomics, *Sci. Rep.* 8 (1) (2018) 13328.
- [43] P. Eichinger, E. Alberts, C. Delbridge, S. Trebeschi, A. Valentinitz, S. Bette, T. Huber, J. Gempt, B. Meyer, J. Schlegel, C. Zimmer, J.S. Kirschke, B.H. Menze, B. Wiestler, Diffusion tensor image features predict IDH genotype in newly diagnosed WHO grade II/III gliomas, *Sci. Rep.* 7 (1) (2017) 13396.
- [44] P. Chang, J. Grinband, B.D. Weinberg, M. Bardis, M. Khy, G. Cadena, M.Y. Su, S. Cha, C.G. Filippi, D. Bota, P. Baldi, L.M. Poisson, R. Jain, D. Chow, Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas, *AJNR Am. J. Neuroradiol.* 39 (7) (2018) 1201–1207.
- [45] Z. Li, Y. Wang, J. Yu, Y. Guo, W. Cao, Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma, *Sci. Rep.* 7 (1) (2017) 5467.
- [46] M.E. Hegi, A.C. Diserens, T. Gorlia, M.F. Hamou, N. de Tribolet, M. Weller, J.M. Kros, J.A. Hainfellner, W. Mason, L. Mariani, J.E. Bromberg, P. Hau, R.O. Mirimanoff, J.G. Cairncross, R.C. Janzer, R. Stupp, MGMT gene silencing and benefit from temozolomide in glioblastoma, *N. Engl. J. Med.* 352 (10) (2005) 997–1003.
- [47] Z.C. Li, H. Bai, Q. Sun, Q. Li, L. Liu, Y. Zou, Y. Chen, C. Liang, H. Zheng, Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: a multicentre study, *Eur. Radiol.* 28 (9) (2018) 3640–3650.
- [48] Y.B. Xi, F. Guo, Z.L. Xu, C. Li, W. Wei, P. Tian, T.T. Liu, L. Liu, G. Chen, J. Ye, G. Cheng, L.B. Cui, H.J. Zhang, W. Qin, H. Yin, Radiomics signature: a potential biomarker for the prediction of MGMT promoter methylation in glioblastoma, *J. Magn. Reson. Imaging* 47 (5) (2018) 1380–1387.
- [49] Z. Kong, Y. Lin, C. Jiang, L. Li, Z. Liu, Y. Wang, C. Dai, D. Liu, X. Qin, Y. Wang, Z. Liu, X. Cheng, J. Tian, W. Ma, (18F)F-FDG-PET-based Radiomics signature predicts MGMT promoter methylation status in primary diffuse glioma, *Cancer Imaging* 19 (1) (2019) 58.
- [50] P. Korfiatis, T.L. Kline, D.H. Lachance, I.F. Parney, J.C. Buckner, B.J. Erickson, Residual deep convolutional neural network Predicts MGMT methylation status, *J. Digit. Imaging* 30 (5) (2017) 622–628.
- [51] Y. Li, Z. Qian, K. Xu, K. Wang, X. Fan, S. Li, X. Liu, Y. Wang, T. Jiang, Radiomic features predict Ki-67 expression level and survival in lower grade gliomas, *J. Neurooncol.* 135 (2) (2017) 317–324.
- [52] Z. Kong, J. Li, Z. Liu, Z. Liu, D. Zhao, X. Cheng, L. Li, Y. Lin, Y. Wang, J. Tian, W. Ma, Radiomics signature based on FDG-PET predicts proliferative activity in primary glioma, *Clin. Radiol.* 74 (10) (2019) 815e15–815.e23.
- [53] K.J. Langen, N. Galldiks, E. Hattingen, N.J. Shah, Advances in neuro-oncology imaging, *Nat. Rev. Neuro.* 13 (5) (2017) 279–289.
- [54] N. Galldiks, M. Kocher, K.J. Langen, Pseudoprogression after glioma therapy: an update, *Expert Rev. Neurother.* 17 (11) (2017) 1109–1115.
- [55] R.J. Young, A. Gupta, A.D. Shah, J.J. Graber, Z. Zhang, W. Shi, A.I. Holodny, A.M. Omuro, Potential utility of conventional MRI signs in diagnosing pseudoprogression in glioblastoma, *Neurology* 76 (22) (2011) 1918–1924.
- [56] P.Y. Wen, D.R. Macdonald, D.A. Reardon, T.F. Cloughesy, A.G. Sorensen, E. Galanis, J. Degroot, W. Wick, M.R. Gilbert, A.B. Lassman, C. Tsien, T. Mikkelsen, E.T. Wong, M.C. Chamberlain, R. Stupp, K.R. Lamborn, M.A. Vogelbaum, M.J. van den Bent, S.M. Chang, Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group, *J. Clin. Oncol.* 28 (11) (2010) 1963–1972.
- [57] N. Galldiks, V. Dunkl, G. Stoffels, M. Hutterer, M. Rapp, M. Sabel, G. Reifenberger, S. Kebir, F. Dorn, T. Blau, U. Herrlinger, P. Hau, M.I. Ruge, M. Kocher, R. Goldbrunner, G.R. Fink, A. Drzezga, M. Schmidt, K.J. Langen, Diagnosis of pseudoprogression in patients with glioblastoma using O-(2-[18F]fluoroethyl)-L-tyrosine PET, *Eur. J. Nucl. Med. Mol. Imaging* 42 (5) (2015) 685–695.
- [58] P. Patel, H. Baradaran, D. Delgado, G. Askin, P. Christos, A. John Tsiouris, A. Gupta, MR perfusion-weighted imaging in the evaluation of high-grade gliomas after treatment: a systematic review and meta-analysis, *Neuro Oncol* 19 (1) (2017) 118–127.
- [59] C.P. Filss, F. Cicone, N.J. Shah, N. Galldiks, K.J. Langen, Amino acid PET and MR perfusion imaging in brain tumours, *Clin. Transl. Imag.* 5 (3) (2017) 209–223.
- [60] J.Y. Kim, J.E. Park, Y. Jo, W.H. Shim, S.J. Nam, J.H. Kim, R.E. Yoo, S.H. Choi, H.S. Kim, Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients, *Neuro Oncol.* 21 (3) (2019) 404–414.
- [61] X. Hu, K.K. Wong, G.S. Young, L. Guo, S.T. Wong, Support vector machine multiparametric MRI identification of pseudoprogression from tumor recurrence in patients with resected glioblastoma, *J. Magn. Reson. Imaging* 33 (2) (2011) 296–305.
- [62] P. Lohmann, M.A. Elahmadawy, J. Werner, M. Rapp, G. Ceccon, G.R. Fink, N.J. Shah, K. Langen, N. Galldiks, OS9.6 Diagnosis of pseudoprogression using FET PET radiomics, *Neuro-Oncology* 21(Supplement_3) (2019) iii19-iii19.
- [63] B.S. Jang, S.H. Jeon, I.H. Kim, I.A. Kim, Prediction of pseudoprogression versus progression using machine learning algorithm in glioblastoma, *Sci. Rep.* 8 (1) (2018) 12516.
- [64] M. Li, H. Tang, M.D. Chan, X. Zhou, X. Qian, DC-AL GAN: pseudoprogression and true tumor progression of glioblastoma multiforme image classification based on DCGAN and AlexNet, *Med. Phys.* 47 (3) (2020) 1139–1150.
- [65] Z. Zhang, J. Yang, A. Ho, W. Jiang, J. Logan, X. Wang, P.D. Brown, S.L. McGovern, N. Guha-Thakurta, S.D. Ferguson, X. Fave, L. Zhang, D. Mackin, L.E. Court, J. Li, A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images, *Eur. Radiol.* 28 (6) (2018) 2255–2263.
- [66] L. Peng, V. Parekh, P. Huang, D.D. Lin, K. Sheikh, B. Baker, T. Kirschbaum, F. Silvestri, J. Son, A. Robinson, E. Huang, H. Ames, J. Grimm, L. Chen, C. Shen, M. Soike, E. McTyre, K. Redmond, M. Lim, J. Lee, M.A. Jacobs, L. Kleinberg, Distinguishing true progression from radionecrosis after stereotactic radiation therapy for brain metastases with machine learning and radiomics, *Int. J. Radiat. Oncol. Biol. Phys.* 102 (4) (2018) 1236–1243.
- [67] M. Spitzer, S. Lorkowski, P. Cullen, A. Sczyrba, G. Fuellen, IsoSVM—distinguishing isoforms and paralogs on the protein level, *BMC Bioinf.* 7 (2006) 110.
- [68] P. Lohmann, M. Kocher, G. Ceccon, E.K. Bauer, G. Stoffels, S. Viswanathan, M.I. Ruge, B. Neumaier, N.J. Shah, G.R. Fink, K.J. Langen, N. Galldiks, Combined FET PET/MRI radiomics differentiates radiation injury from recurrent brain metastasis, *Neuroimage Clin.* 20 (2018) 537–542.
- [69] 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. Aerts, Computational radiomics system to decode the radiographic phenotype, *Cancer Res.* 77 (21) (2017) e104–e107.
- [70] P.M. Szczypinski, M. Strzelecki, A. Materka, A. Klepaczko, MaZda—a software package for image texture analysis, *Comput. Methods Programs Biomed.* 94 (1) (2009) 66–76.
- [71] C. Nioche, F. Orhac, S. Boughdad, S. Reuze, J. Goya-Outi, C. Robert, C. Pellot-Barakat, M. Soussan, F. Frouin, I. Buvat, LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity, *Cancer Res.* 78 (16) (2018) 4786–4789.