Comparison of MR Preprocessing Strategies and Sequences for Radiomics-Based MGMT Prediction

Daniel Abler[1(✉)], Vincent Andrearczyk[1], Valentin Oreiller[1], Javier Barranco Garcia[2], Diem Vuong[2], Stephanie Tanadini-Lang[2], Matthias Guckenberger[2], Mauricio Reyes[3], and Adrien Depeursinge[1]

[1] Institute of Information Systems, University of Applied Sciences Western Switzerland (HES-SO), Sierre, Switzerland
{Daniel.Abler,Vincent.Andrearczyk,Valentin.Oreiller,Adrien.Depeursinge}@hevs.ch

[2] Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland
{Javier.Barranco,Diem.Vuong,Stephanie.Tanadini-Lang,Matthias.Guckenberger}@usz.ch

[3] ARTORG Center for Biomedical Research, University of Bern, Bern, Switzerland
Mauricio.Reyes@artorg.unibe.ch

Abstract. Hypermethylation of the O6-methylguanine-DNA-methyltransferase (MGMT) promoter in glioblastoma (GBM) is a predictive biomarker associated with improved treatment outcome. In clinical practice, MGMT methylation status is determined by biopsy or after surgical removal of the tumor. This study aims to investigate the feasibility of non-invasive medical imaging based "radio-genomic" surrogate markers of MGMT methylation status.

The imaging dataset of the RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) challenge allows exploring radiomics strategies for MGMT prediction in a large and very heterogeneous dataset that represents a variety of real-world imaging conditions including different imaging protocols and devices. To characterize and optimize MGMT prediction strategies under these conditions, we examined different image preprocessing approaches and their effect on the average prediction performance of simple radiomics models.

We found features derived from FLAIR images to be most informative for MGMT prediction, particularly if aggregated over the entire (enhancing and non-enhancing) tumor with or without inclusion of the edema. Our results also indicate that the imaging characteristics of the tumor region can distort MR-bias-field correction in a way that negatively affects the prediction performance of the derived models.

Keywords: GBM-MGMT · Radiomics · MR-standardization

This work was supported by the Swiss National Science Foundation (SNSF, grant 205320_179069) and the Swiss Personalized Health Network (SPHN) via the IMAGINE project.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
A. Crimi and S. Bakas (Eds.): BrainLes 2021, LNCS 12963, pp. 367–380, 2022.
https://doi.org/10.1007/978-3-031-09002-8_33

1 Introduction

Brain tumors are a rare condition with about 330 000 yearly incidents (227 000 deaths) worldwide [14]. Despite their comparably low incidence rate, brain cancers cause the highest average number of years of life lost among all cancers [3]. Gliomas represent 75% of primary malignant brain tumors in adults [9]. Glioblastoma (GBM) is the most frequent and most malignant sub-type of glioma, and accounts for about 50% of primary malignant brain cancer. Surgical resection to the maximal safe extent is the primary treatment for newly diagnosed GBM, followed by radiation and chemotherapy. However, due to the tumor's infiltrative growth, resection, even if macroscopically complete, leaves behind tumor cells and the tumor typically recurs at the margins of the resection cavity. Following surgery, addition of Temozolomide (TMZ) chemotherapy to radiotherapy treatment has been the standard of care since 2005. Despite active multi-institutional and international research efforts, GBM patients continue to have a poor prognosis with a 5-year survival rate of about 5.5% [9].

The quest for more personalized oncologic treatment allocation has driven research into genetic and molecular biomarkers. A few of these have been reported to have prognostic and/or predictive implications for GBM tumors. Among those, hypermethylation of the O6-methylguanine-DNA-methyltransferase (MGMT) gene has been shown to be associated with improved outcome in (GBM) and is now considered a favorable prognostic factor and a predictor of chemotherapy response for GBM patients [16]. MGMT encodes for an important DNA repair protein; it's expression may be suppressed by methylation of the promoter region, which results in decreased DNA repair capability and thus increased susceptibility to the damaging effects of DNA-targeting treatments, such as alkylating agents like TMZ and also radiation therapy.

Genetic and molecular profiling is an invasive process that involves tissue extraction from the tumor via stereotactic/needle biopsies; it is inherently prone to sampling bias and therefore may not capture the tumor's spatial heterogeneity, a particular hallmark of GBM. Detection of a tumor's MGMT status from routine clinical imaging, on the other hand, could provide a non-invasive detection approach that would not suffer from the sampling restrictions inherent to surgical biopsies. Recent research into such "radio-genomic" imaging-markers has been fueled by the premise that biomedical images may contain information about the underlying pathophysiology which, even if hidden from the human eye, can be captured via quantitative image analysis [5].

Evidence for imaging markers of MGMT promoter methylation in GBM, however, remains mixed. A systematic review and meta-analysis [17] of 22 studies (published before March 2018) that investigated MR-imaging features linked to MGMT promoter methylation identified reduced edema, elevated apparent diffusion coefficient (ADC) and low perfusion as likely imaging characteristics of promoter methylated MGMT. The review found increased performance for studies using ADC or perfusion measures and recommends including the corresponding MR sequences in imaging protocols.

More recent studies indicate AUCs between 0.74 to 0.87 [6–8,10,19] and accuracies of 0.76 to 0.83 [4,12] for a variety of "classical radiomics" and deep learning models.

These studies were performed on single-institution datasets [7,12], public data collections [4,6], or a combination of both [8,10] and were all limited to fewer than 300 patients across training and test sets. The most heterogeneous dataset appears to have been investigated in [10] which employed a cohort of 133 patients (from the public TCIA GBM collection and from Guangzhou General Hospital) for training and another cohort of 60 patients (from two further hospitals) for independent validation.

This study presented in this manuscript was performed on the imaging dataset of the 2021 RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) challenge which aggregates imaging information from five public repositories and an undisclosed number of institutional data collections. Images were acquired "under standard clinical conditions, but with different equipment and imaging protocols, resulting in a vastly heterogeneous image quality reflecting diverse clinical practice across different institutions" [1].

This diversity allows exploring the performance of radiomic strategies for MGMT prediction in the presence of significant heterogeneity (protocols, device-manufacturers and generations) that must be expected in real-world applications. To this end, we first examined different data preprocessing approaches and their effect on average prediction performance of simple radiomics models. Using the most promising preprocessing strategy and most informative MR-sequence, we then investigated which tumor compartments and groups of radiomic features drive MGMT prediction.

2 Materials and Methods

2.1 Data

This study used the public dataset of the RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge 2021 [1,2,13]. The private test set of the RSNA-ASNR-MICCAI challenge was not considered in the analysis because our image processing workflow relied on tumor segmentation tools that could not be automatized fully for deployment as Jupyter notebook as required for submission via the Kaggle portal, see Sect. 2.2 for further details.

Data for the MGMT prediction task of the BraTs challenge (task 2) consisted in pre-operative MR-imaging for 585 patients with GBM; brain tumor, as well as information about the tumor's MGMT promoter methylation. MR imaging comprised four MR sequences per patient: T1-weighted (T1), T1 contrast (T1c) after administration of Gadolinium contrast agent, T2-weighted MRI (T2) and FLAIR (FLAIR). All images were provided in DICOM format, after application of a skull-stripping technique that preserved space and resolution of the original images [1]. The tumors' MGMT status was provided as a binary variable that indicates either absence or presence of promoter methylation of the MGMT gene.

The imaging of three of the 585 patients was incomplete with missing information in at least a subset of the four MR-sequences. These patients (IDs 109, 123, 709) were excluded from further processing.

The imaging dataset was highly heterogeneous, comprising MRIs collected from multiple public and institutional resources. Consequently, imaging protocols and quality varied greatly across the dataset. For example, MR slice spacing (thickness) varied between 0.22 to 11.00 mm (0.43 to 6.00 mm) with different acquisition planes across MR sequences and patients.

2.2 Image Pre-processing

All images were processed according to the workflow depicted in Fig. 1: First, complete patient imaging datasets containing all four MRI sequences (T1, T1c, T2, FLAIR) were selected and DICOM images converted to the NIfTI format. Brain masks were created by removing the zero-background from the skull-stripped images. Independently, the tumors present in each patient's imaging dataset were segmented, as described in this section. Exploiting the availability of brain masks and tumor regions of interest (ROIs), the original images were subjected to further pre-processing followed by feature extraction.

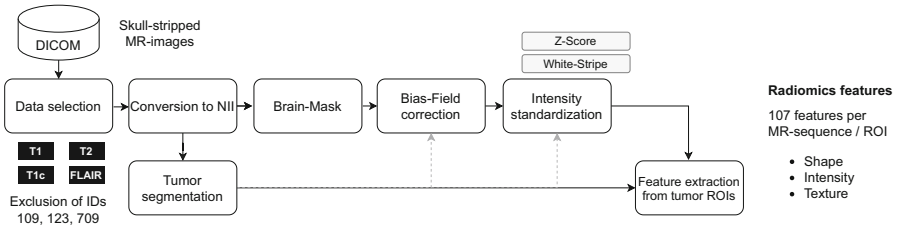


Fig. 1. Image processing workflow.

Tumor Segmentation. Each patient's tumor was segmented using the automatic GBM segmentation software DeepBraTumIA[1]. DeepBraTumIA requires T1, T1c, T2 and FLAIR sequences as input to identify the enhancing tumor (ET), non-enhancing tumor (NET) and edema subregions. Masks for each of these ROIs are output in a co-registered reference frame, as well as in the original frame of the respective MR original sequence. At the time of writing, the tool required user interaction via a GUI to initiate the automatic segmentation process. As this process could not be fully automatized, submission through the challenge portal via Python Jupyter notebook was not possible. Instead, validation was performed on subsets of the public data as described in Sect. 2.3.

[1] https://www.nitrc.org/projects/deepbratumia/, as of November 2021.

Table 1. Investigated MR-image preprocessing approaches.

Config ID	Bias-field correction	Standardization	Mask
A	Yes	white-stripe	Brain
B	Yes	z-score	Brain
C	No	z-score	Brain
D	No	z-score	Brain w/o tumor
E	Yes	z-score	Brain w/o tumor

Intensity Standardization. Non-quantitative MR imaging lacks a physical/anatomical reference-scale that would allow direct quantitative comparison of image intensity across vendors and imaging protocols. To enable comparability in the absence of such a reference, statistical standardization techniques are employed that aim at aligning the intensity distributions found in anatomical reference structures across individual image acquisitions. Frequently used reference-regions in the context of brain-MRI are the brain's white-matter component (white-stripe standardization) or the entire brain anatomy (z-score standardization over brain mask). Both approaches operate on the assumption that each constituent of the brain anatomy results in imaging voxels of a characteristic MR intensity and that the relative proportion of those constituents remains relatively stable across individuals. The latter assumption is clearly violated in the case of GBM patients whose tumor lesion may replace a substantial portion of healthy brain tissue and represents a source of strong inter-individual variability in the imaging appearance.

Given the strong heterogeneity of the RSNA-ASNR-MICCAI dataset, we considered effective standardization to be critical to extracting a robust MGMT-related signal from the imaging data. For this reason, we investigated a range of preprocessing strategies that included different combinations of (a) application of bias-field correction, (b) standardization method and (c) the brain regions to which those two steps were applied.

Table 1 summarizes the tested strategies. Bias-field correction and z-score standardization were performed using the respective SimpleITK [11] functions. For white-stripe standardization we relied on a modified version of the implementation in [15][2]. Processing with "brain" mask included all non-zero voxels of the skull-stripped images. For processing with "brain w/o tumor" mask, all tumor subregions detected by the automatic segmentation approach were removed from the "brain" mask.

Radiomics Feature Extraction. In addition to the tumor subregions provided by the automatic segmentation approach, further masks were created corresponding to different combinations of those subregions. We only included combinations of ROIs that typically result in a contiguous volume, and therefore

[2] https://github.com/jcreinhold/intensity-normalization, as of November 2021.

did not consider the combination of non-enhancing tumor and edema. Table 2 summarizes all relevant ROIs.

Table 2. Tumor sub-regions used for feature extraction

ROI name	Tumor subregion(s)
ROI_1	Edema
ROI_2	Enhancing tumor (ET)
ROI_3	Non-enhancing tumor (NET)
ROI_1-2	Edema & ET
ROI_2-3	ET & NET
ROI_1-2-3	ET & NET & edema

We used pyradiomics [18] to extract shape-based, intensity and texture features (n=107) from each ROI (Table 2) and MR-sequence resulting from the different pre-processing configurations (Table 1). The preprocessed images were resampled to a spatial resolution of 1.00 mm × 1.00 mm × 1.00 mm prior to feature extraction. Image intensity values were shifted to positive values and the "bin_width" parameter for gray value discretization was chosen in such a way that the full range of intensity values after standardization could be well captured by approximately 64 bins.

2.3 Modeling Experiments

This study aimed at identifying (a) the most suitable pre-processing approach for the provided imaging dataset and (b) sets of radiomics features, ROI and MR-modality combinations that carry predictive information of MGMT-status. To address these goals, we performed two sets of modeling experiments that will be reported here.

In a first experiment, we investigated the change of global performance trends with the preprocessing approach. The second experiment, focused on identifying the specific features and ROIs that provide most relevant information for the globally best performing pre-processing combination.

As indicated in Sect. 2.1, all experiments performed in this study relied exclusively on the public data set of the challenge which was divided into subsets (stratified by MGMT status) for model development (training-set: 80%) and testing of the final prediction model on unseen data (test-set: 20%), respectively. Initial experiments showed the final test performance to be highly sensitive to those splits, probably due to the high degree of heterogeneity of the imaging data. To obtain representative performance results despite this variability, we repeated model development and final evaluation steps multiple times using different train/test splits of the dataset, as illustrated in Fig. 2. The specific feature and model selection strategies employed for each of the experiments are further detailed in the following subsections.

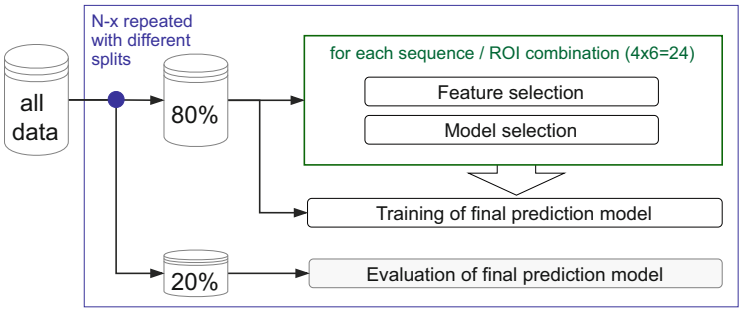


Fig. 2. Model development and final evaluation was performed multiple times using different training/test splits of the full dataset ($N = 50$ for experiment-1, $N = 15$ for experiment-2).

Experiment-1: Performance Trends in Function of Preprocessing Strategy. The first experiment aimed at uncovering global performance trends in function of the chosen preprocessing strategy.

For each combination of MR sequence and ROI (24 combinations), a set of features was selected by choosing the 20 (individually) most predictive features, identified by univariate F-score ranking, and discarding strongly correlated features (Pearson correlation coefficient exceeding 0.7) from this group. The selected features were standardized and the hyperparameters of a family of linear classifiers (including support-vector classifiers with linear kernel, logistic regression, with L2, L1 and combined penalties) were optimized to maximize AUC in 5-fold cross-validation, resulting in a single "best model" per MR-sequence and ROI. Finally, the combination of MR-sequence and ROI that yielded the highest-performing "best model" in cross-validation was selected as final prediction model, and evaluated on the unseen split of the dataset.

This process was repeated for 50 different train/test splits by varying the random seed of the the stratified split procedure, resulting in 50 independently selected combinations of prediction model, MR-sequence and ROI.

The entire feature selection, training and evaluation process was repeated with features derived from the five different pre-processing configurations in Table 1.

Experiment-2: Identification of Predictive Features and ROIs for Optimal Preprocessing Strategy and Most Informative MR-sequence. The second experiment aimed at identifying the specific set of radiomic features that carry predictive information of MGMT status for the different ROIs of the best-performing preprocessing approach and MR sequence resulting from the first experiment.

The feature selection approach in this experiment relied on repeated hyperparameter optimization of Elastic-Net classifiers in 5-fold cross-validation. Each such optimization resulted in a set of models with varying cross-validation per-

formance and varying number of selected features. From this set of models, we chose the best-performing model with no more than 20 features and extracted its features. For a given training set, this feature selection process was repeated 15 times by varying the random seed for cross-validation splitting, resulting in 15 feature sets. The features from those sets were then aggregated and ordered by their selection frequency, resulting in a ranking of the most-frequently selected features across the 15 selection repetitions.

From the top-ranked (most frequently selected) features, we built a prediction model by 5-fold cross-validation and grid search over the space of linear classifiers and hyper-parameters. The most promising of these models (highest cross-validation performance) was selected as final prediction model. This prediction model was evaluated on the test-split which had not been used in the feature selection and training process.

To obtain robust estimates of test-performance, we repeated the above process (feature selection, model training, model evaluation) for 15 different train/test splits by varying the random seed of the stratified splitting procedure. This resulted in 15 independently selected feature sets, prediction models and their corresponding test performances for each ROI and MR-sequence combination.

2.4 Statistical Analysis

Average AUC performance on the test-sets was compared across the different preprocessing approaches in experiment-1, and across tumor ROIs in experiment-2. To test the null hypothesis that two sets of performance measurements originate from the same distribution, an independent 2-sample t-test was performed when both sets of measurements were normally distributed and of equal variance. Welch's t-test was performed in case of normal distribution and un-equal variance, and Mann-Whitney U test in case of non-normal distribution. P-values < 0.05 were interpreted as "significant" rejection of the null hypothesis, and are reported in the manuscript.

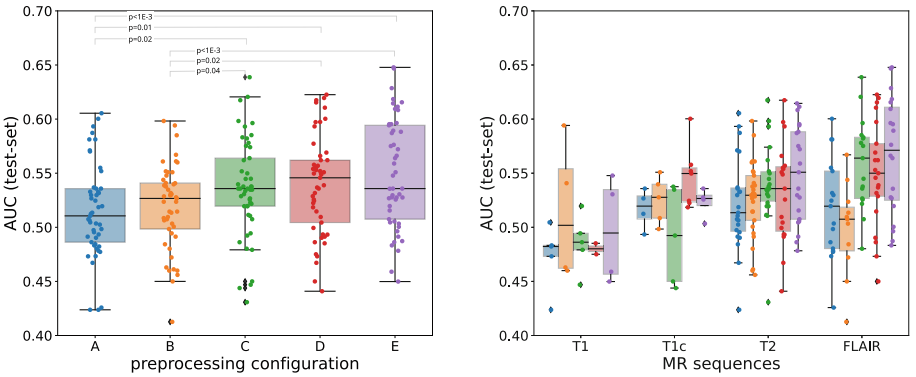
3 Results

3.1 Experiment-1: Identification of Favorable Preprocessing Strategies and Most Informative MR-sequence

Figure 3 summarizes the test performance of the single-sequence, single-ROI prediction models evaluated in experiment-1, Sect. 2.3, in function of pre-processing configuration.

Models based on features derived from preprocessing configurations {C, D, E} performed significantly better ($p < 0.05$) compared to models based on preprocessing configurations {A and B}, Fig. 3(a), with mean AUC of 0.54±0.05 to 0.55±0.05 vs. 0.52±0.04, Table 3. Within these sets, no statistically significant performance differences were observed. Both configuration sets differ in their

approach towards bias-field correction: For preprocessing configurations {A, B}, the bias field was estimated over the entire brain volume; configurations {C, D, E} either did not attempt to correct for the bias-field, or excluded the tumor from bias-field estimation. The individually highest performing models from configurations {C, D, E} were based on the MR-FLAIR sequence, Fig. 3(b).



(a) Preprocessing (b) MR-sequence

Fig. 3. Distribution of test-AUC of 50 best performing classifiers resulting from experiment-1, Sect. 2.3. Models based on features derived from preprocessing configurations {C, D, E} perform significantly better ($p < 0.05$) than those based on features derived from configurations {A, B}, Fig. 3(a). Within these groups, no significant differences in overall performance were observed. Highest performing models from configurations {C, D, E} were based on MR-FLAIR Fig. 3(b)

Table 3. Average cross-validation and test performance (AUC) of 50 best performing classifiers resulting from experiment-1 in function of preprocessing approach.

Config ID	A	B	C	D	E
Validation AUC	0.60±0.02	0.61±0.01	0.61±0.01	0.61±0.02	0.62±0.01
Test AUC	0.52±0.04	0.52±0.04	0.54±0.05	0.54±0.05	0.55±0.05

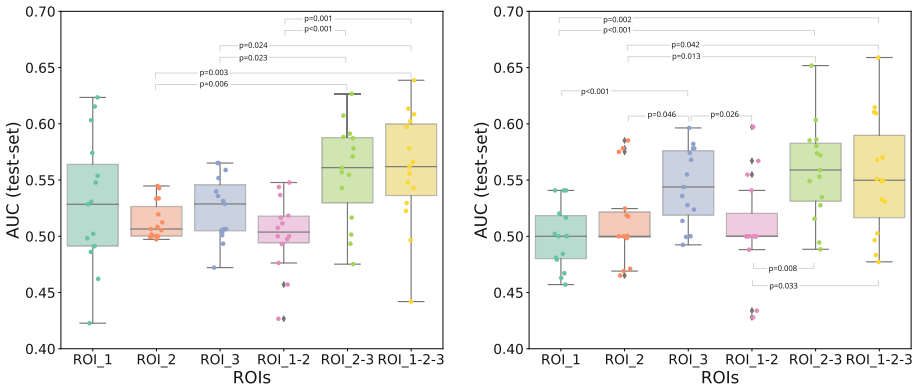
3.2 Experiment-2: Identification of Predictive Feature Groups and ROIs

Figure 4 summarizes the prediction performance of experiment-2 using features derived from MR-FLAIR sequences preprocessed according to approach C, Fig. 4(a), and approach D, Fig. 4(b), respectively.

Highest test performance was achieved by models derived from ROI_2-3 and ROI_1-2-3, reaching AUCs of 0.56±0.04 to 0.56±0.05 (preprocessing configuration C) and 0.55±0.05 to 0.56±0.04 (preprocessing configuration D), respectively, Table 4. In both preprocessing configurations, the performance of models

from these ROIs was significantly higher than performance of models from ROI_2 alone. For configuration C, the performance of models from these ROIs was also significantly higher than performance of models from ROI_3, but not from ROI_1. The opposite was found for configuration D.

Figure 5 shows the 15 features that were most frequently included in the 15 independently developed ROI-specific prediction models resulting from experiment-2. Features selected for ROI_3, ROI_2-3, ROI_1-2-3 and preprocessing configurations C and D are displayed. Multiple features were selected consistently across both best-performing ROIs (ROI_2-3 and ROI_1-2-3): GLDM-DependenceNonUniformityNormalized and shape (SurfaceVolumeRatio vs Sphericity) for preprocessing configuraton C, and GLCM-ClusterProminence and shape (SurfaceVolumeRatio vs Sphericity) for preprocessing configuration D. Also for ROI_3, the same three features (GLRLM-GrayLevelNonUniformity, GLCM-ClusterShade, GLSZM-SmallAreaEmphasis) were selected consistently across preprocessing configurations.



(a) Pre-processing (C) (b) Pre-processing (D)

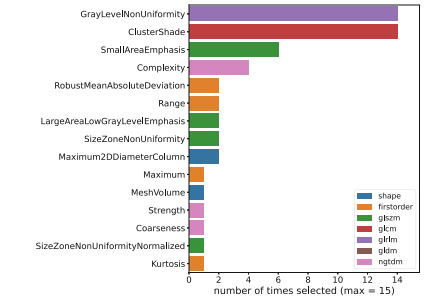
Fig. 4. Distribution of test-AUC for ROI-specific models resulting from the model selection and evaluation approach proposed in experiment-2. Features were derived from MR-FLAIR images preprocessed according to configuration C, Fig. 4(a), and D, Fig. 4(b), respectively. Models based on features derived from ROI_2-3 and ROI_1-2-3 performed significantly best across the tested preprocessing configurations.

Table 4. Average cross-validation and test performance (AUC) of 15 repetition of experiment-2 Sect. 2.3 based on MR-FLAIR images preprocessed according to configuration C and D.

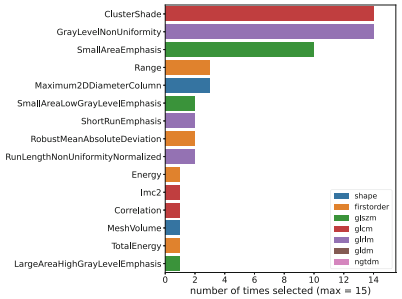
AUC	ROI_1	ROI_2	ROI_3	ROI_1-2	ROI_2-3	ROI_1-2-3
(C) Validation	0.58±0.02	0.57±0.02	0.59±0.03	0.57±0.02	0.59±0.03	0.60±0.02
(C) Test	0.53±0.06	0.51±0.02	0.52±0.03	0.50±0.03	0.56±0.04	0.56±0.05
(D) Validation	0.56±0.03	0.57±0.04	0.59±0.03	0.55±0.02	0.60±0.03	0.60±0.02
(D) Test	0.50±0.03	0.51±0.04	0.54±0.03	0.51±0.04	0.56±0.04	0.55±0.05

4 Discussion

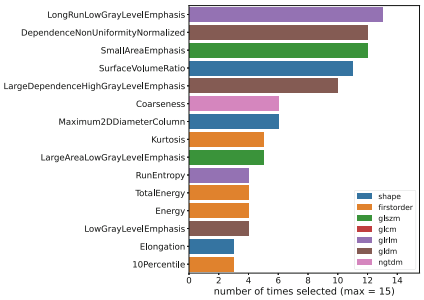
This study investigated the effect of different preprocessing strategies on MGMT prediction performance. The observed significant performance difference between preprocessing configurations {A, B} vs {C, D, E} indicates that inclusion of the tumor region in the bias-field estimation process deteriorates the resulting models' ability to predict MGMT status. This negative effect on modeling outweighed any other potential gains that bias-field correction may have conferred to the



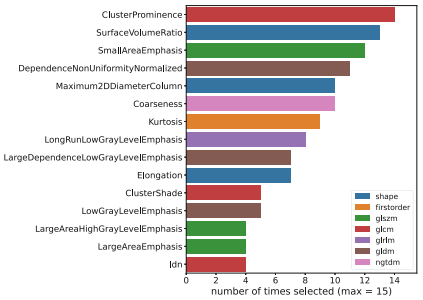
(a) ROI_3 (C)



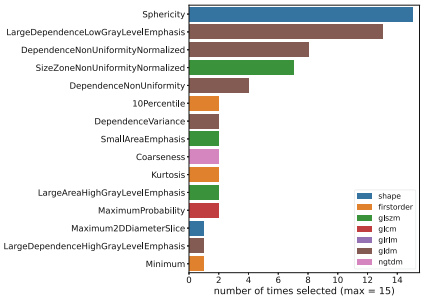
(b) ROI_3 (D)



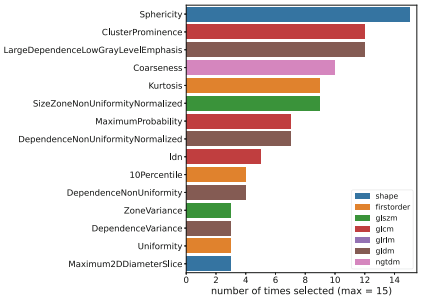
(c) ROI_2-3 (C)



(d) ROI_2-3 (D)



(e) ROI_1-2-3 (C)



(f) ROI_1-2-3 (D)

Fig. 5. Selected features

prediction task. Our experiments did not provide clear indications for the relative superiority of different standardization techniques on overall performance (white-stripe vs z-score based on intensity statistics of the entire brain with or without tumor regions). However, when disaggregated by MR-sequence, predictors based on preprocessing approaches {C, D, E} applied to FLAIR sequences appeared to outperform all other approaches.

Focusing on FLAIR sequences, we then investigated predictive performance and relevant features separately for each ROI and for two of the three well performing preprocessing strategies (C, D). Results showed higher predictive performance for models based on ROI_2-3 (ET, NET) and ROI_1-2-3 (entire tumor: edema, ET, NET) compared to individual tumor sub-regions (ROI_1, ROI_2), potentially indicating that the transition zones between the individual tumor compartments carry relevant information for the prediction task. This confirms the findings of earlier studies that identified features from the tumor core and the whole tumor region, but not the edema alone, to be most predictive [6,10,12] across different MR-sequences.

Experiment-2 selected various features consistently across the best performing ROIs and preprocessing configurations C and D. These include shape measures (Sphericity, SurfaceVolumeRatio) of the core tumor alone (ROI_2-3) and with edema (ROI_1-2-3), as well as texture features (GLSZM-SmallAreaEmphasis, GLDM-DependenceNonUniformityNormalized, GLDM-LargeDependenceLowGrayLevelEmphasis, GLCM-ClusterProminence). Shape measures of the active tumor, and edema region have previously been shown to carry relevant information for MGMT prediction [6]. Another study [8] reported the highest prediction performance to be achieved by T2 texture features (FLAIR was not investigated) with GLCM-ClusterProminence among the best-performing features.

Overall performance of the identified single-sequence and single-ROI models remained relatively low with maximum average test-AUC of 0.56 ± 0.04 for ROI_2-3 in preprocessing configurations C and D. Large variability in performance measures was observed across models trained and validated against different train/test splits despite ensuring stratification with regard to MGMT status. For example, the performance of individual train test splits varied between AUC 0.47 to 0.63 and 0.48 to 0.65 for ROI_2-3 and preprocessing configurations C and D, respectively.

This variability indicates the presence of other strongly differentiating factors in the imaging dataset that were not controlled for by stratification with regard to MGMT status. Among the potential candidates are specific characteristics of the imaging protocols used by the different imaging centers that contributed to this datasets. As a next step, we therefore seek to investigate the effect of imaging characteristics on MGMT prediction performance by subdividing the dataset into subgroups that share a more homogeneous set of image acquisition and quality parameters.

References

1. Baid, U., et al.: The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification. arXiv:2107.02314, September 2021
2. Bakas, S., et al.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data **4**(1), 1–3 (2017). https://doi.org/10.1038/sdata.2017.117
3. Burnet, N.G., Jefferies, S.J., Benson, R.J., Hunt, D.P., Treasure, F.P.: Years of life lost (YLL) from cancer is an important measure of population burden–and should be considered when allocating research funds. Br. J. Cancer (2005). https://doi.org/10.1038/sj.bjc.6602321
4. Chang, P., et al.: Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. Am. J. Neuroradiol. **39**(7), 1201–1207 (2018). https://doi.org/10.3174/ajnr.A5667
5. Gillies, R.J., Kinahan, P.E., Hricak, H.: Radiomics: images are more than pictures. Data. Radiol. **278**(2), 563–577 (2016). https://doi.org/10.1148/radiol.2015151169
6. Hajianfar, G., et al.: Noninvasive O6 methylguanine-DNA methyltransferase status prediction in glioblastoma multiforme cancer using magnetic resonance imaging radiomics features: univariate and multivariate radiogenomics analysis. World Neurosurg. **132**, e140–e161 (2019). https://doi.org/10.1016/j.wneu.2019.08.232
7. Kihira, S., et al.: Multiparametric MRI texture analysis in prediction of glioma biomarker status: added value of MR diffusion. Neuro-Oncol. Adv. **3**(1), vdab051 (2021). https://doi.org/10.1093/noajnl/vdab051
8. Korfiatis, P., et al.: MRI texture features as biomarkers to predict MGMT methylation status in glioblastomas: MRI texture features to predict MGMT methylation status. Med. Phys. **43**(6Part1), 2835–2844 (2016). https://doi.org/10.1118/1.4948668
9. Lapointe, S., Perry, A., Butowski, N.A.: Primary brain tumours in adults. The Lancet **392**(10145), 432–446 (2018). https://doi.org/10.1016/S0140-6736(18)30990-5
10. Li, Z.-C., et al.: Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: a multicentre study. Eur. Radiol. **28**(9), 3640–3650 (2018). https://doi.org/10.1007/s00330-017-5302-1
11. Lowekamp, B.C., Chen, D.T., Ibáñez, L., Blezek, D.: The design of SimpleITK. Front. Neuroinform. **7**, 1–8 (2013). https://doi.org/10.3389/fninf.2013.00045
12. Lu, Y., et al.: Machine learning-based radiomic, clinical and semantic feature analysis for predicting overall survival and MGMT promoter methylation status in patients with glioblastoma. Magn. Reson. Imaging **74**, 161–170 (2020). https://doi.org/10.1016/j.mri.2020.09.017
13. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging **34**(10), 1993–2024 (2015). https://doi.org/10.1109/TMI.2014.2377694
14. Patel, A.P., et al.: Global, regional, and national burden of brain and other CNS cancer, 1990–2016: a systematic analysis for the global burden of disease study 2016. The Lancet Neurol. **18**(4), 376–393 (2019). https://doi.org/10.1016/S1474-4422(18)30468-X

15. Reinhold, J.C., Dewey, B.E., Carass, A., Prince, J.L.: Evaluating the impact of intensity normalization on MR image synthesis. In: Angelini, E.D., Landman, B.A. (eds.) Medical Imaging 2019: Image Processing. p. 126. SPIE, San Diego, United States, March 2019. https://doi.org/10.1117/12.2513089
16. Rivera, A.L., et al.: MGMT promoter methylation is predictive of response to radiotherapy and prognostic in the absence of adjuvant alkylating chemotherapy for glioblastoma. Neuro-Oncol. **12**(2), 116–121 (2010). https://doi.org/10.1093/neuonc/nop020
17. Suh, C., Kim, H., Jung, S., Choi, C., Kim, S.: Clinically relevant imaging features for MGMT promoter methylation in multiple glioblastoma studies: a systematic review and meta-analysis. Am. J. Neuroradiol. **39**, 1439–1445 (2018). https://doi.org/10.3174/ajnr.A5711
18. van Griethuysen, J.J., et al.: Computational radiomics system to decode the radiographic phenotype. Cancer Res. **77**(21), e104–e107 (2017). https://doi.org/10.1158/0008-5472.CAN-17-0339
19. Xi, Y.b., et al.: Radiomics signature: a potential biomarker for the prediction of MGMT promoter methylation in glioblastoma: GBM radiomics features reflect MGMT. J. Magn. Reson. Imaging **47**(5), 1380–1387 (2018). https://doi.org/10.1002/jmri.25860