Medical Research Archives

ESMED
European Society of Medicine

OPEN ACCESS

Published: May 31, 2023

Citation: Dequidt P, Bourdon P, Tremblais B, et al., 2023. A complete pipeline for glioma grading using intelligible AI on multimodal MRI data. Medical Research Archives, [online] 11(5). https://doi.org/10.18103/mra.v11i5.3793

DOI: https://doi.org/10.18103/mra.v11i5.3793

ISSN: 2375-1924

RESEARCH ARTICLE

A complete pipeline for glioma grading using intelligible AI on multimodal MRI data

Paul Dequidt[4], Pascal Bourdon[1,5], Benoit Tremblais[1,5], Mathieu Naudin[2,3,5], Pierre Fayolle[3,5], Clément Giraud[2,3,5], Carole Guillevin[2,3,5], Christine Fernandez-Maloigne[1,5] and Rémy Guillevin[2,3,5]*

[1]XLIM Laboratory, UMR CNRS 7252, University of Poitiers, Poitiers, France
[2]DACTIM-MIS, LMA, UMR CNRS 7348, University of Poitiers, Poitiers, France
[3]CHU de Poitiers, 86000 Poitiers, France
[4]COMETE, UMRS-S 1075, INSERM
[5]I3M, Common Laboratory CNRS Universitary Hospital and University of Poitiers

*remy.guillevin@chu-poitiers.fr

Abstract

1) Objectives: Machine learning for binary glioma grading have been extensively used on anatomical MRI, especially using the BraTS dataset. The relevance of radiomic criteria based on multimodal imaging, including diffusion, perfusion and spectroscopy data is to be explored, as multimodal datasets are scarce, and there is no common benchmark for performance comparison.

2) Material and methods: Poitiers University Hospital provides 123 multimodal patient data. We computed 124 features and let a recursive feature elimination algorithm (RFE) yield a relevant, reduced subset of features. We trained a SVM classifier on this subset. We proposed a method to adapt the BraTS dataset to allow performance comparison with the literature. We got a performance reference point by training on anatomical data only, and showed improvements when multimodalities were added. We explored the feature relevance through the RFE subset. The RFE subset is not constant and induce variability in the performances. To smooth the variability, we applied the RFE algorithm 100 times and incremented the selected features, resulting in a global feature ranking. We also show the best classifier reached on these 100 trainings and its feature subset.

3) Results: The best classifier reached 86.5% accuracy, with a mean accuracy on 100 trainings of 78.6%. The rankings shows that anatomical and perfusion sequences are the most relevant for glioma grading, especially T1 post-gadolinium, cerebral blood volume and flow. Intensity and texture features are frequently selected, while anisotropic diffusion coefficient, time to peak and mean time transit mappings seem irrelevant.

4) Conclusion: Multimodal radiomics improve the classification and are consistent with the radiological analysis.

1. Introduction

Gliomas are the most common type of brain tumor, linked to a high mortality within a short survival range. Gliomas are highly. heterogeneous tumors, and optimal treatment depends on identifying and locating the highest-grade disease present. This is why the quantitative analysis of images acquired for the diagnosis and treatment of patients with brain tumors has increased significantly in recent years, particularly using machine learning methods and, more specifically, deep learning algorithms[1-3]. The value of imaging in patients with brain tumor can be enhanced if pathologic data can be estimated from imaging input using predictive models, knowing imaging techniques for diagnosis and treatment of patients are often not validated against the histopathologic criterion standard.

According to the World Health Organization (WHO), glioma development is described by their grade, ranging from I to IV, with low grade gliomas (LGG) being grades I-II and high-grade gliomas (HGG) being grades III-IV[4]. But if glioma grading using machine learning is an active field, only a few studies use multimodal Magnetic Resonance (MR) data, including diffusion and perfusion imaging, and magnetic resonance spectroscopy (MRS)[5]. The reference dataset for this task is the MICCAI Brain Tumor Segmentation (BraTS) challenge dataset[6-8], which provides 285 patients with 4 anatomical sequences: T1, T1 contrast-enchanced (T1c), T2 and T2 FLuid Attenuated Inversion Recovery (FLAIR). However, this dataset is not divided into HGG and LGG according to the WHO, but into High Grade Gliomas (glioblastoma multiform, WHO-IV) and Lower Grade Gliomas (WHO-I, II, III). Consequently, most works on this task are not aligned with a WHO-defined system, despite claiming otherwise[9,10]. In a previous study, we proposed a new ground truth to align the BraTS dataset on the WHO definition, to use it as a reference dataset for a WHO-defined binary grading task[11].

But anatomical data lack specificity for a robust classification[12]. Multimodal MRI, including perfusion, diffusion and MRS data, gives important information for binary grading[13]. In order to create an alternative to the BraTS dataset, some authors use proprietary datasets, with their own acquisition protocol. They provide more sequences and more MR modalities, but with a lack of patients. For example, Citak-Er et al. use a dataset of 43 patients[14] and Vamvakas et al. use only 40 patients[15]. The lack of a large and open multimodal dataset prevents creating a performance benchmark to compare different publications.

Poitiers University Hospital, France, has led a 3T multimodal pre-operative acquisition protocol with 123 histologically confirmed glioma patients, 46 LGG and 77 HGG. Using these data, we created a complete processing pipeline from the acquisition to the automatic classification. This pipeline involves multimodal registration, brain and tumor segmentation, intensity normalization, feature extraction, feature selection and classification using Support Vector Machines (SVM)[16-17].

In this paper, we will review the current performances in binary glioma grading. We

will describe our acquisition protocol and processing pipeline. We will train and evaluate our classifier on the Poitiers University Hospital multimodal dataset, analyze the feature ranking produced by the feature selection and link this ranking to the radiologic analysis. The conclusion will present some perspectives.

2. Material and methods

2.1. Multimodal imaging

Radiologists can reach a binary estimation of glioma grades from anatomical data[18,19]. Some visual features, such as gadolinium enhancement, necrosis, mass effect, or FLAIR inhomogeneity can be read as clues for a high-grade label. Questions can be raised about the robustness of these features, as it is known that LGG, which are usually non-enhancing after gadolinium injection, can show enhancement patterns[20]. Therefore, multimodal MR data can improve the diagnosis with complementary information[21-26].

Fitting the arterial input function on perfusion imaging give parametric maps such as cerebral blood volume (CBV) or cerebral blood flow (CBF), mean transit time (MTT) and time to peak (TTP). The ratio between the tumor region of interest (ROI) and contro-lateral healthy tissues give the relative CBV (rCBV) and CBF (rCBF) values, showing hypo or hyper-perfusion in the lesion. Diffusion tensor imaging (DTI) generates maps of the anisotropic diffusion coefficient (ADC), fractional anisotropy (FA) and the b1000 trace image. Proton MRS at short (35 ms) and long (135 ms) echo time (TE) give metabolic concentrations such as choline (Cho), creatine (Cr), N-acetylaspartic acid (NAA), lipids (Lip) or lactates (Lac), which can be compared between patients by calculating metabolic ratios.

2.2. Data

We used two datasets: Poitiers University Hospital multimodal dataset and the anatomical BraTS 2018 dataset[1].

Using Poitiers University Hospital Magnetom Skyra 3 Tesla (Siemens Healthineers, Erlangen, Germany), we created a dataset of 123 patients, divided in 46 LGG and 77 HGG, confirmed by histo-molecular analysis. For each patient, we measured the following modalities and sequences: anatomical imaging with T1, T1 post-gadolinium (T1c) and T2 FLAIR; diffusion imaging with ADC, FA and b1000; perfusion imaging, with CBF, CBV, MTT and TTP maps; and proton MRS for short and long TE.

In the existing literature, most authors use the BraTS 2018 dataset. This dataset provides 285 patients with 4 anatomical sequences: T1, T1c, T2 and T2 FLAIR. But this dataset is not divided according to the WHO definition of LGG and HGG. We proposed in a previous study a radiological ground truth to align this dataset on a WHO-based system and showed

[1] https://www.med.upenn.edu/sbia/brats2018/data.html

that our proposed ground truth has little impact on the performances, despite an important class imbalance[11].

We propose to adapt the BraTS dataset to get a performance reference point for our study on our multimodal data. We will train and evaluate our classifier on the BraTS dataset with 4 and 3 anatomical sequences, removing T2 data that we do not have at Poitier University Hospital. We will then use our radiological ground truth to train on a 3 sequence WHO-aligned BraTS dataset. This training provides a reference point to evaluate our performances on Poitiers University Hospital's anatomical data. Finally, we will then train our classifier on the multimodal data and see how multimodality impacts the classification.

2.3. Features

We model the radiologic analysis through two types of features: image computed features and manual measurements. We used Pyradiomics[27] to extract image features aligned with the imaging biomarkers standardization initiative (ISBI)[28].

For each patient, we computed 7 shape features, 6 histogram-based intensity features and 5 texture features. A ROI of tumor segmentation is needed for feature extraction. The shape features are only computed once per patient while the intensity and texture features are computed on each MR sequence. The shape features include length of major and minor axis, maximum 3D diameter, elongation, flatness, sphericity, and surface area. The histogram-based intensity features are mean, skewness, kurtosis, contrast, energy, and entropy. For texture analysis we used the correlation of the gray level co-occurrence matrix (GLCM), coarseness, inverse difference moment (IDM), complexity and strength. This gives us 117 automated image features.

The manual measurements are the rCBV and rCBF values and 5 metabolic ratios: Cho/Cr, Cho/NAA, NAA/Cr, Lac/Cr at long echo time (TE, 135ms) and Lip/Cr at short TE (35 ms). Therefore, each patient is described by 124 features.

These acquisitions are time and resource consuming. Some sequences were missing for some patients, creating sparse data. We fix these sparse data by assigning the mean value of the corresponding feature.

2.4. Binary classification, current performances

Common evaluation metrics used for this task are precision, sensitivity, and specificity. For binary glioma grading, the accuracy ranges from 84% to 95.5% depending on the method and dataset used[5]. Deep learning methods gives interesting results but lacks intelligibility. We want to use a machine learning methods to have an intelligible classifier. Computing a large set of features increases the dimensionality. This is why many authors use a feature selection algorithm and train on the resulting set.

Computing a large set of features increases the dimensionality. Most authors use a feature selection algorithm, like the SVM Recursive Feature Elimination (SVM-RFE) algorithm, and train on the resulting set[26]. The SVM-RFE returns a subset of features with high informational value and helps reduce

dimensionality. This algorithm gives good results associated with a linear kernel.

2.5. Preprocessing pipeline

We designed a preprocessing pipeline for multimodal sequences, including registration, brain, and tumor segmentation.

Using multimodal MRI data requires the implementation of a complete processing pipeline. These steps include multimodal registration (rigid transformations and oversampling to a common resolution), brain extraction, intensity normalization and tumor segmentation. The Image Tool-Kit[27-30] and FMRIB Software Library (FSL)[31] give tools for registration, brain extraction, oversampling, and intensity normalization. Each step of our pipeline has been visually validated by an expert radiologist.

The rigid multimodal registration groups a translation followed by an affine transform, with a Powell optimizer[32]. We also resampled each image on a common resolution, aligned with the BraTS 2018 acquisitions.

Brain segmentation is performed using the Brain Extraction Tool (BET)[33]. We got the best segmentation on T1c images and applied the segmentation mask on the other registered sequences to get multimodal brain segmentation.

The intensity normalization is computed through histogram matching. We chose the patient with the highest intensity amplitude of the BraTS dataset as the reference histogram.

The features are extracted from the tumor ROI. Good segmentation results are reached using deep learning networks[34], especially with U-net based methods like Isensee et al.[35]. This method is among the best results in the BraTS 2017 challenge with an average Dice score of 0.896.

We have trained the U-net proposed by Isensee et al. on the BraTS data, using only the 3 anatomical sequences available at the Hospital of Poitiers, i.e., T1, T1c and T2 FLAIR. We use the 3D patches extraction proposed by Fenneteau et al.[36] for better memory allocation. With one sequence less than the original implementation by Isensee et al., we reached an average Dice score of 0.804. Our network gives good visual results, and these segmentations has been validated by an expert radiologist.

2.6. Classifier

We want to use an intelligible classifier and analyze its performances through feature selection and feature ranking. We used a 20\% test set and a 5-fold cross-validated SVM-RFE for feature selection, followed by a SVM with grid search to optimize hyperparameters. We tested two kernels (linear and radial) and 5 values of the regularization hyperparameter ranging from 0.1 to 2.

We let the SVM-RFE algorithm choose the best number of features for training. This exploration phase led to a high variability in the classifier performances. We propose to smooth this variability by doing 100 independent training and average the performances (precision, sensitivity, and specificity). We also show the best precision reached on these 100 independent trainings.

Each training gets its own subset of features by the SVM-RFE. We will analyze the selection frequency of each feature to see which feature is relevant for our problem in a global feature ranking.

3. Results

3.1. Performance Analysis

Training on the BraTS 2018 dataset gives a precision of 85.8 ± 4.1%, within performances in the state of the art[8]. Training on 3 sequences gives similar performances, showing that the T2 sequence is not important for binary grading. Using our WHO-aligned ground truth drops slightly the accuracy. As this ground truth has an important class imbalance, these results show that our learning is robust.

We get our performance reference point by training on the anatomical sequences available at Poitiers Hospital. This gives a lower precision with only 77.5 ± 6.8%. The average sensitivity stays stable with 88.5 ± 6.2%, and the specificity drops to 60.2 ± 14.0%. The best classifier on 100 trainings gets an accuracy of 81.1%. We want to see how these values change when we add multimodality.

Adding multimodal data leads to a 1.1\% gain in averaged accuracy, going from 77.5 ± 6.8% to 78.6 ± 6.6%. The best classifier on 100 trainings reaches 86.5\% accuracy; a strong improvement compared to anatomical data only. We observe a slight decrease in sensitivity, from 88.5% ± 6.2% to 85.9% ± 7.9%. The largest improvement is reached for the average specificity, reaching 67.3 ± 15.0%. Multimodality improves the classification, but these improvements are not as important as we would expect from the state of art, where multimodal SVM often reached 90% accuracy. Extended results are shown in Table 1.

4. Discussion

4.1. Feature ranking

We want to analyze the ranking produced by the SVM-RFE. The best classifier on 100 trainings used a linear kernel and a subset of 23 features, Table 2. This ranking corresponds to the weights of the normal vector of the SVM hyperplane, as described by Guyon et al.[26]. We will also analyze the selection frequency of each feature on 100 trainings. This second ranking gives us a global ranking of all 124 features. Both rankings give information about which features and sequences are relevant in the classification. We will compare the top 23 features of this frequency ranking, Table 3, with the 23 features of Table 2.

These rankings share 14 common features, with homogeneous ranks. Therefore, we will read the ranking on 100 independent trainings as robust global alternative ranking.

Half of the features in Table 2 are anatomical sequences, mostly T1c. In the top ranked features, 3 out of 5 are intensity features computed on anatomical sequences. These results are coherent with the radiological analysis: the T1c contrast is the top feature, and HGG are contrasted lesions with gadolinium enhancement next to necrotic regions, while LGG are homogeneous.

Groundtruth	Accuracy[1]	Sensitivity[1]	Specificity[1]	Best accuracy[2]
BraTS 2018 4 sequences	85,8 ± 4,1%	89,5 ± 4,5%	75,4 ± 9,7%	89,5%
BraTS 2018 3 sequences	86,0 ± 4,2%	89,4 ± 4,5%	75,9 ± 9,6%	87,7%
WHO-aligned BraTS 2018 [10] 3 sequences	82,1 ± 4,1%	84,0 ± 5,5%	67,1 ± 9,7%	84,2%
Poitiers Hospital 3 sequences	77,5 ± 6,8%	88,5 ± 6,2%	60,2 ± 14,0%	81,1%
Poitiers Hospital Multimodal data	78,6 ± 6,6%	85,9 ± 7,9%	67,3% ± 15,0%	86,5%

[1]Mean performances on 100 trainings

[2]Best accuracy reached on 100 trainings

Table 1: Performances results. Each value reached on 100 independant training is given with its standard deviation.

Rank	Feature	Sequence	Modality
1	Contrast	T1c	Anatomical
2	Entropy	T1c	Anatomical
3	rCBV	CBV	Perfusion
4	Skewness	T1	Anatomical
5	Complexity	CBV	Perfusion
6	GLCM correlation	CBV	Perfusion
7	Mean	T2 FLAIR	Anatomical
8	IDM	CBV	Perfusion
9	IDM	T1c	Anatomical
10	Complexity	CBF	Perfusion
11	GLCM correlation	T2 FLAIR	Anatomical
12	IDM	T1	Anatomical
13	Skewness	b1000	Diffusion
14	Mean	T1c	Anatomical
15	Skewness	CBV	Perfusion
16	Lac/Cr	CSI	MRS
17	Energy	CBV	Perfusion
18	GLCM correlation	FA	Diffusion
19	Kurtosis	FA	Diffusion
20	Kurtosis	T2 FLAIR	Anatomical
21	GLCM correlation	TTP	Perfusion
22	Entropy	b1000	Diffusion
23	Entropy	T1	Anatomical

Table 2: Ranking from the best classifier reached on 100 independant training

Rank	Feature	Sequence	Modality	Frequency[1]
1	Contrast	T1c	Anatomical	99
2	rCBV	CBV	Perfusion	94
3	Entropy	T1c	Anatomical	86
4	GLCM correlation	CBV	Perfusion	82
5	IDM	T1	Anatomical	81
6	Lac/Cr	CSI	MRS	78
7	IDM	T1c	Anatomical	69
8	IDM	CBV	Perfusion	67
9	Skewness	T1	Anatomical	63
10	Kurtosis	T1c	Anatomical	62
11	Kurtosis	CBV	Perfusion	57
12	rCBF	CBF	Perfusion	55
13	GLCM correlation	FA	Diffusion	54
14	GLCM correlation	CBF	Perfusion	54
15	Mean	T1c	Anatomical	52
16	IDM	FA	Diffusion	46
17	GLCM correlation	MTT	Perfusion	43
18	Entropy	T1	Anatomical	42
19	Skewness	T1c	Anatomical	41
20	Strength	b1000	Diffusion	41
21	Complexity	b1000	Diffusion	40
22	Mean	T2 FLAIR	Anatomical	39
23	Kurtosis	FA	Diffusion	32

[1]Number of occurence in the RFE subset on 100 independant trainings

Table 3: Ranking from 100 independant training

Perfusion features are less frequent than anatomical features. Most of them are from CBV and CBF maps. The rCBV value is among the 3 best features in both rankings. The rCBF value appears in Table 3 and has been selected 55 times. These ranking underline the relevance of these manual measurements for the classification task. TTP and MTT maps are only present once in both rankings, so we can question their relevance for binary grading.

Regarding diffusion imaging, we note that the ADC is absent from both rankings, which is surprising as ADC is visually discriminative, as shown in Figure 1.

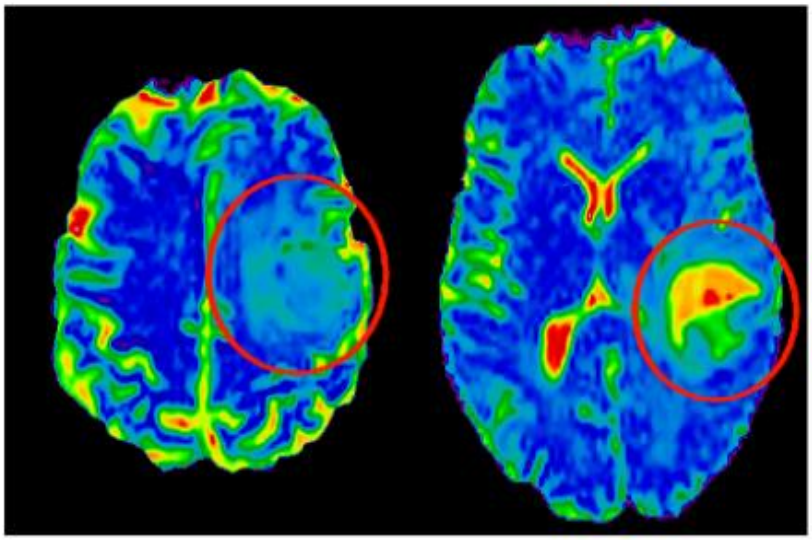

Figure 1: ADC maps of (left) a LGG patient and (right) a HGG patient. The HGG patients shows high intra-tumoral ADC values which are not visible for a LGG patient. Therefore, ADC should appear as a discriminative sequence.

Most features in both rankings are intensity-based features, followed by texture features. No shape feature appears in both rankings. Every type of intensity features appears, mostly from anatomical T1c and T1 sequences, which underlines the relevance of such features. Dissimilarly, coarseness and strength features are always absent. These texture features are not interesting for binary grading.

Using 100 independent training allow us to get a global feature ranking, as seen in Figure 2. This Figure shows the number of occurrences of each 124 features in the SVM-RFE subset on 100 independent trainings. From the shape of this histogram, we can tell that some features were never selected and that some others were highly solicited. The most selected feature, T1c contrast, has been selected 99 times out of 100. 15 features were never selected in the SVM-RFE subset. Most of them were texture features, namely strength and coarseness. This result shows very clearly that these texture features are not interesting for glioma grading.

Two other features were never selected: the T2 FLAIR contrast and CBV contrast. The absence of the T2 FLAIR contrast is surprising, as HGG can have a heterogeneous aspect in T2 FLAIR, resulting in an increased contrast within the lesion.

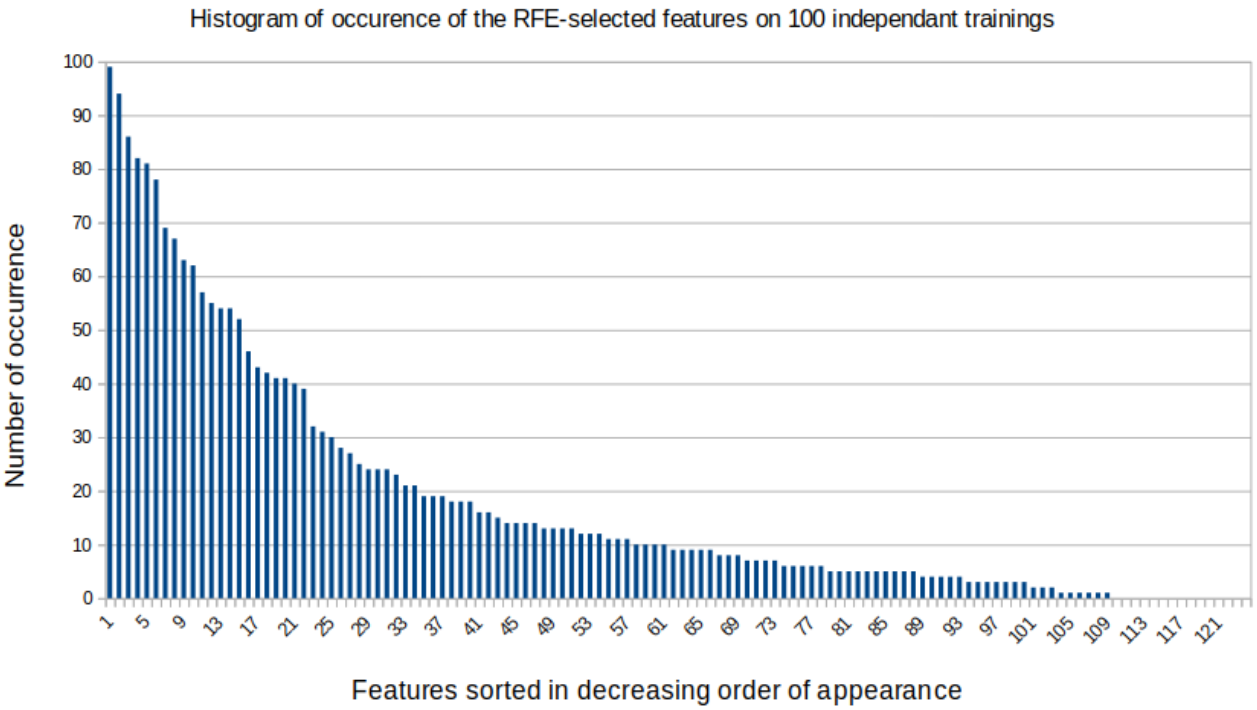

Figure 2: Features sorted in decreasing order of appearance

5. Conclusion

Using multimodal MRI data from Poitiers University Hospital, we proposed a complete processing pipeline ranging from the acquisition to the automatic classification. This pipeline includes automatic multimodal registration, brain extraction, intensity normalization, tumor segmentation, and features computation. Two types of features are still based on manual measurements: metabolic SRM ratios and relative contralateral perfusion values.

To the best of our knowledge, our work is the only one to use multimodal data and provide a performance reference point with the BraTS database, which is the most popular dataset used in glioma grading. This allows us to compare our classifier result with the current state-of-art and show the impact of multimodal data on our performances. Adding multimodal data, we observed a small gain in mean intensity compared to learning on anatomical data only. The best impact was on specificity. Therefore, we can say that adding multimodal information has a positive effect on the performances.

We proposed an analysis of our classifier through feature ranking. We analyzed two rankings: one produced by 100 independent trainings and the other by the best classifier reached on these trainings. Both rankings produced comparable results, and we used the analysis over 100 trainings as a global ranking for all 124 features. This global ranking allows us to deepen the feature analysis, showing that strength and coarseness are not interesting features for this classification task. The same result appears for ADC, MTT and TTP sequences.

Our ranking also produced interesting result for radiologists, as it highlights anatomical intensity-based features, such as T1c contrast, as relevant for binary grading. Therefore, we showed how machine learning can be used as an exploration tool and produce a consistent analysis of the radiological process.

To extend this study, our current work deals with comparison of this approach with convolutional neural networks classification, and by considering the explainability of the algorithms, essential point for health data[37]. Moreover, we now use imaging coming from a Siemens 7Tesla MRI Scanner, intalled in Poitiers since December 2019. Our Siemens Magnetom 7Tesla Ultra High Field 7 Tesla Magnetic Resonance Imaging scanner is one of only three in France, and the lonely one for clinical use as well as for research. Indeed, MAGNETOM Terra is the first 7T scanner released for clinical use in Europe and the US and the Universitary Hospital in Poitiers, was the first in the world, in december 2021, to accommodate over 200O patients for clinical use. This is a huge opportunity to develop our AI systems for the benefit of patients and physicians.

Corresponding Author
Rémy Guillevin
DACTIM-MIS, LMA, UMR CNRS 7348,
University of Poitiers, Poitiers, France.
CHU de Poitiers, 86000 Poitiers, France.
I3M, Common Laboratory CNRS Universitary Hospital and University of Poitiers.
Email: remy.guillevin@chu-poitiers.fr

Conflicts of Interest statement
None.

Funding information
None.

Acknowledgements
None.

References:

[1] Bahar Ryan C., Merkaj Sara, Cassinelli Petersen Gabriel I., Tillmanns Niklas, Subramanian Harry, Brim Waverly Rose, Zeevi Tal, Staib Lawrence, Kazarian Eve, Lin MingDe, Bousabarah Khaled, Huttner Anita J., Pala Andrej, Payabvash Seyedmehdi, Ivanidze Jana, Cui Jin, Malhotra Ajay, Aboian Mariam S., "Machine Learning Models for Classifying High- and Low-Grade Gliomas: A Systematic Review and Quality of Reporting Analysis", Frontiers in Oncology, vol 12 , 2022, DOI: 10.3389/fonc.2022.856231, ISSN=2234-943X

[2] Menze B, Isensee F, Wiest R, Wiestler B, Maier-Hein K, Reyes M, Bakas S., "Analyzing magnetic resonance imaging data from glioma patients using deep learning", Comput Med Imaging Graph. 2021 Mar;88:101828. doi: 10.1016/j.compmedimag.2020.101828. Epub 2020 Dec 2. PMID: 33571780; PMCID: PMC8040671.

[3] Lo, Chung-Ming, Yu-Chih Chen, Rui-Cian Weng, and Kevin Li-Chun Hsieh. 2019. "Intelligent Glioma Grading Based on Deep Transfer Learning of MRI Radiomic Features" *Applied Sciences* 9, no. 22: 4926. https://doi.org/10.3390/app9224926

[4] 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. 14260 Ellison, The 2016 world health organization classification of tumors of the central nervous system: a summary, Acta neuropathologica 131 (6) (2016) 803–820.

[5] P. Dequidt, P. Bourdon, O. B. Ahmed, B. Tremblais, C. Guillevin, M. Naudin, C. Fernandez-Maloigne, R. Guillevin, Recent advances in glioma grade classification using machine and deep learning on MR data, in: 2019 Fifth International Conference on Advances in Biomedical Engi- neering (ICABME), IEEE, 2019, pp. 1–4.

[6] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (BraTS), IEEE transactions on medical imaging 34 (10) (2014) 1993–2024.

[7] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, C. Davatzikos, Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features, Scientific data 4 (2017) 170117.

[8] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, M. Rozycki, et al., Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BraTS challenge, arXiv preprint arXiv:1811.02629.

[9] C. Ge, Q. Qu, I. Y.-H. Gu, A. S. Jakola, 3d multi-scale convolutional net- works for glioma grading using mr images, in: 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018, pp. 141–145.

[10] G. Cui, J. Jeong, B. Press, Y. Lei, H.-K. Shu, T. Liu, W. Curran, H. Mao, X. Yang, Machine-learning-based classification of lower-grade gliomas and high-grade gliomas using radiomic features in multi-parametric MRI, arXiv preprint arXiv:1911.10145.

[11] P. Dequidt, P. Bourdon, B. Tremblais, C. Guillevin, B. Gianelli, C. Boutet, J.-P. Cottier, J.-N. Vallée, C. Fernandez-Maloigne, R. Guillevin, Assigning a new glioma grade label ground-truth for the brats dataset using radiologic criteria, in: 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, 2020, pp. 1–6.

[12] N. Upadhyay, A. Waldman, Conventional MRI evaluation of gliomas, The British journal of radiology 84 (special issue 2) (2011) S107–S111.

[13] J.-L. Dietemann, Neuro-imagerie diagnostique, Elsevier Health Sciences, 2018.

[14] F. Citak-Er, Z. Firat, I. Kovanlikaya, U. Ture, E. Ozturk-Isik, Machine- learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3t, Computers in biology and medicine 99 (2018) 154 – 300

[15] A. Vamvakas, S. Williams, K. Theodorou, E. Kapsalaki, K. Fountas, C. Kappas, K. Vassiou, I. Tsougos, Imaging biomarker analysis of advanced multiparametric MRI for glioma grading, Physica Medica 60 (2019) 188–198.

[16] Jainy Sachdeva, Vinod Kumar, Indra Gupta, Niranjan Khandelwal, Chirag Kamal Ahuja, A package-SFERCB- "Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors", Applied Soft Computing,Volume 47,2016,Pages 151-167,ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2016.05.020.

[17] Latif G, Ben Brahim G, Iskandar DNFA, Bashar A, Alghazo J. Glioma Tumors' Classification Using Deep-Neural-Network-Based Features with SVM Classifier. Diagnostics (Basel). 2022 Apr 18;12(4):1018. doi: 10.3390/diagnostics12041018. PMID: 35454066; PMCID: PMC9032951.

[18] B. L. Dean, B. P. Drayer, C. R. Bird, R. A. Flom, J. A. Hodak, S. W. Coons, R. G. Carey, Gliomas: classification with MR imaging., Radiology 174 (2) (1990) 411–415.

[19] Gates EDH, Lin JS, Weinberg JS, Prabhu SS, Hamilton J, Hazle JD, Fuller GN, Baladandayuthapani V, Fuentes DT, Schellingerhout D. Imaging-Based Algorithm for the Local Grading of Glioma. AJNR Am J Neuroradiol. 2020 Mar;41(3):400-407. doi: 10.3174/ajnr.A6405. Epub 2020 Feb 6. PMID: 32029466; PMCID: PMC7077885.

[20] R. Cavaliere, M. B. S. Lopes, D. Schiff, Low-grade gliomas: an update on pathology and therapy, The Lancet Neurology 4 (11) (2005) 760–770.

[21] R. Guillevin, Metabolic-oncological MR imaging of diffuse low-grade glioma: A dynamic approach, in: Diffuse Low-Grade Gliomas in Adults, Springer, 2013, pp. 219–234.

[22] Inano R, Oishi N, Kunieda T, Arakawa Y, Yamao Y, Shibata S, Kikuchi T, Fukuyama H, Miyamoto S., Voxel-based clustered imaging by multiparameter diffusion tensor images for glioma grading., Neuroimage Clin. 2014 Aug 7;5:396-407. doi: 10.1016/j.nicl.2014.08.001. eCollection 2014.PMID: 25180159

[23] Chengjun Yao, Shunzeng Lv, Hong Chen, Weijun Tang, Jun Guo, Dongxiao Zhuang, Nikos Chrisochoides, Jinsong Wu, Ying Mao & Liangfu Zhou (2016) The clinical utility of multimodal MR image-guided needle biopsy in cerebral gliomas,International Journal of Neuroscience, 126:1, 53-61, DOI: 10.3109/00207454.2014.992429

[24] Ning Z, Luo J, Xiao Q, Cai L, Chen Y, Yu X, Wang J, Zhang Y. Multi-modal magnetic resonance imaging-based grading analysis for gliomas by integrating radiomics and deep features. Ann Transl Med. 2021 Feb;9(4):298. doi: 10.21037/atm-20-4076. PMID: 33708925; PMCID: PMC7944310

[25] J. J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H. J. Aerts, Computational radiomics system to decode the radiographic phenotype, Cancer research 77 (21) (2017) e104–e107.

[26] A. Zwanenburg, S. Leger, M. Vallières, S. Löck, et al., Image biomarker standardisation initiative-feature definitions, arXiv preprint arXiv:1612.07003.

[27] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Machine learning 46 (1-3) (2002) 389–422.

[28] R. Beare, B. Lowekamp, Z. Yaniv, Image segmentation, registration and characterization in r with simpleitk, Journal of statistical software 86.

[29] Z. Yaniv, B. C. Lowekamp, H. J. Johnson, R. Beare, Simpleitk image analysis notebooks: a collaborative environment for education and reproducible research, Journal of digital imaging 31 (3) (2018) 290–303.

[30] B. C. Lowekamp, D. T. Chen, L. Ibáñez, D. Blezek, The design of simpleitk, Frontiers in neuroinformatics 7 (2013) 45.

[31] P. Bandettini, M. Jenkinson, C. Beckmann, T. Behrens, M. Woolrich, S. Smith, FSL, NeuroImage 62 (2) (2012) 782–790.

[32] M. Naudin, Visualisation et aide à la décision pour la neuro-navigation per-opératoire, Ph.D. thesis, Poitiers (2018).

[33] S. M. Smith, Fast robust automated brain extraction, Human brain mapping 17 (3) (2002) 143–155.

[34] T. Zhou, S. Ruan, S. Canu, A review: Deep learning for medical image segmentation using multi-modality fusion, Array 3 (2019) 100004.

[35] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K. H. Maier-Hein, Brain tumor segmentation and radiomics survival prediction: Contribution to the BraTS 2017 challenge, in: International MICCAI Brainlesion Workshop, Springer, 2017, pp. 287–297.

[36] A. Fenneteau, P. Bourdon, D. Helbert, C. Fernandez-Maloigne, C. Habas, R. Guillevin, Investigating efficient CNN architecture for multiple sclerosis lesion segmentation, Journal of Medical Imaging 8 (1) (2021) 014504.

[37] Pascal Bourdon, Olfa Ben Ahmed, Thierry Urruty, Khalifa Djemal, Christine Fernandez-Maloigne, Explainable AI for medical imaging: knowledge matters, pages 267-292, chapter of Multi-faceted Deep Learning, Models and Data, Editors: Benois-Pineau Jenny, Zemmari Akka, Springer 2021