2024 14th International Conference on Pattern Recognition Systems (ICPRS) | 979-8-3503-7565-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICPRS62101.2024.10677829

Explainable Machine-Learning for identifying the genetic biomarker MGMT in brain tumors using magnetic resonance imaging radiomics.

Sebastian Ponce
PhD. in Health Sciences and Engineering, *Universidad de Valparaíso*,
Valparaíso, Chile.
Center of Interdisciplinary Biomedical and Engineering Research for Health
Millennium Institute for Intelligent Healthcare Engineering (iHEALTH).
sebastian.ponce@postgrado.uv.cl

Steren Chabert
School of Biomedical Engineering, *Universidad de Valparaíso*,
Valparaíso, Chile
Center of Interdisciplinary Biomedical and Engineering Research for Health
Millennium Institute for Intelligent Healthcare Engineering (iHEALTH).
steren.chabert@uv.cl

Leondry Mayeta
PhD. in Health Sciences and Engineering, *Universidad de Valparaíso*,
Valparaíso, Chile.
Center of Interdisciplinary Biomedical and Engineering Research for Health
Millennium Institute for Intelligent Healthcare Engineering (iHEALTH).
leondry.mayeta@postgrado.uv.cl

Pamela Franco
School of Biomedical Engineering, *Universidad de Valparaíso*,
Valparaíso, Chile
Center of Interdisciplinary Biomedical and Engineering Research for Health
pamela.franco@usach.cl

Francisco Plaza-Vega
Departamento de Matemática y Ciencia de la Computación, *Universidad de Santiago de Chile,*
Santiago, Chile.
francisco.plaza.v@usach.cl

Marvin Querales
School of Medical Technology *Universidad de Valparaíso*,
Valparaíso, Chile.
Center of Interdisciplinary Biomedical and Engineering Research for Health
marvin.querales@uv.cl

Rodrigo Salas
School of Biomedical Engineering, *Universidad de Valparaíso*,
Valparaíso, Chile
Center of Interdisciplinary Biomedical and Engineering Research for Health
Millennium Institute for Intelligent Healthcare Engineering (iHEALTH).
Corresponding Author: rodrigo.salas@uv.cl

***Abstract*—Brain tumors often feature the genetic biomarker O6-Methylguanine-DNA-Methyltransferase (MGMT) associated with a favorable response to chemotherapy and an improved prognosis. Currently, detecting MGMT presence relies on invasive brain biopsy procedures. Thus, this study aims to develop a data mining-based radiomics methodology for non-invasive identifying and evaluating brain tumor genetic biomarkers using radiomics in magnetic resonance images (MRIs). MRIs with segmentation masks were used to extract variables and employ feature selection techniques. Several machine learning models were trained, where Logistic Regression, employing LASSO selection, emerged as the best-performing model, achieving 61% accuracy. Additionally, an explainability module utilizing Shap values identified three significant variables: a T1CE sequence variable related to texture, a FLAIR sequence variable of first-order statistics, and a T1 sequence variable of first-order statistics. This radiomic methodology, with its performance and explainable nature, could offer diagnostic support to clinicians in tumor management.**

***Index Terms*—O6-Methylguanine-DNA-Methyltransferase (MGMT) methylation, genetic biomarkers, machine learning, radiomics, explainability, magnetic resonance imaging, brain tumors.**

979-8-3503-7565-7/24/$31.00 ©2024 IEEE

I. Introduction

Brain tumors comprise lesions classified into two main groups: primary, which originate in the cells of the central nervous system, and secondary, which develop in other parts of the body and metastasize to the brain [1]. Among the most common primary tumors are meningioma and glioblastoma multiforme (GBM). In the United States, between 30,000 and 35,000 primary brain tumors are diagnosed each year, with GBM being the most prevalent. However, the five-year survival rate for patients with this diagnosis is unfavorable, with only 20)% overall survival [2]. Locally, in Chile, mortality related to malignant tumors of the meninges, encephalon, and other parts of the central nervous system was 2.6 per 100,000 inhabitants

in 2013, being the second leading cause of childhood mortality due to cancer [3].

It is essential to consider the methylation of the promoter of the O6-methylguanine-DNA-methyltransferase (MGMT) gene in GBM. This gene, located on chromosomal band 10q26, encodes a DNA repair enzyme crucial in preventing carcinogenesis and resistance to chemotherapy with alkylating agents [4]. Determination of MGMT methylation status requires invasive procedures, such as GBM biopsy, followed by DNA, RNA, or protein analysis [5], [6]. However, these invasive methods carry risks and may be affected by tumor heterogeneity, which limits their diagnostic value [7].

Considering these challenges, the possibility of using artificial intelligence techniques in magnetic resonance imaging (MRI) to identify MGMT gene methylation arises. Magnetic resonance images provide distinctive features of GBM, such as necrosis, contrast enhancement, and mass effect, reflecting its macroscopic complexity [8], [9]. In recent studies, with promising results, deep learning and algorithms such as XGBoost have been applied to analyze MR images and predict MGMT methylation [10]. Recently, radiomics has emerged, involving a broad set of computational methods that extract quantitative features from medical images [11], [12]. These features can enhance diagnostic imaging, prognosis, and response to cancer therapy [13]–[16]. Nonetheless, some limitations of this methodology include multicentric acquisition and manual segmentation. Afterward, it is essential to note that few works have incorporated model interpretability [17]–[19], a crucial aspect of understanding which features influence the predictions and generate confidence in the results.

In this context, this work presents an interpretable radiomic methodology for identifying the genetic biomarker MGMT in brain tumors by magnetic resonance imaging using an explainable machine learning model.

II. Methods and Materials

A. Datasets

This study used two preoperative magnetic resonance imaging datasets provided by the BraTS (Brain Tumor Segmentation Challenge) 2021 [20]. This event is jointly sponsored by entities in the field of radiology and medical imaging, such as the RSNA (Radiological Society of North America), the ASNR (American Society of Neuroradiology), and the MICCAI (Medical Image Computing and Computer-Assisted Intervention) conference.

The first dataset corresponds to task 1 of the challenge and is oriented towards evaluating advanced techniques for segmenting heterogeneous brain glioblastoma subregions in MRI images. From this set, segmentation masks have been used to identify distinct tumor areas, including the non-contrast tumor zone, enhanced tumor zone, and peripheral edema. For example, in Cavieres et al. [21]), the authors propose a deep learning method to automatically segment distinct tumor areas from MRI.

The second set, assigned to task 2 of the BraTS 2021 challenge, focuses on evaluating classification methods to predict the methylation status of the MGMT promoter, which is critical in the response to chemotherapy treatment in glioblastomas. Magnetic resonance imaging sequences are as follows: Fluid Attenuated Attenuated Inversion Recovery (FLAIR), T1-weighted pre-contrast (T1), T1-weighted post-contrast (T1CE), and T2-weighted (T2).

B. Exploratory Analysis

As part of the analysis of the variability of the image contrasts, histograms with the means of the intensity values of the pixels of each sequence were made. Once the variables were extracted, a second exploratory analysis was performed using principal component analysis (PCA) and a t-SNE model.

The t-distributed stochastic neighbor embedding (t-SNE) is a nonlinear dimensionality reduction technique for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions maps [22].

C. Image Harmonization

Medical image datasets may be acquired from multiple-centric, multiple vendors, and multiple scanners, inducing the problem of batch effects consisting of considerable variability between batches. Image harmonization methods seek to remove the batch effects and enable increased generalizability and reproducibility of the following processes in the pipeline.

As a processing strategy, image harmonization was performed using histogram matching with a reference image to reduce contrast variability between images within the same sequence. The reference image was chosen based on the histogram analysis, selecting one from the highest frequency range. In image processing, histogram matching enables the transformation of an image to match a specific histogram, a process that proves particularly useful in image normalization. This application is especially relevant when dealing with images sourced from different origins. A review of current harmonization methods with the most common evaluation metrics can be found in Hu et al. [23].

D. Features Extraction

The 3D Slicer software was used [24] to which the brain images and their respective tumor segmentation masks were imported to extract the variables corresponding to the Area of the tumor that does not capture contrast (net), the Area that does capture contrast (et), the Area of tumor edema (ed), the tumor, and finally, cancer together with the peritumoral Area; the latter two were formed based on the first 3. Then, the pyradiomics extension was applied [25], and all available quantitative feature sets were extracted.

E. Features Selection

To reduce the number of features, feature selection techniques were employed. These methods included LASSO (Least Absolute Shrinkage and Selection Operator), Fisher scoring, recursive feature elimination, forward selection, and backward selection (See [12])

F. Machine Learning Models

Subsequently, models were trained using the complete set of variables and the selected subset of variables. Table I shows the employed models, each accompanied by their respective hyperparameter grids and subjected to 5-fold cross-validation.

TABLE I
MODEL HYPERPARAMETER GRID

Model	Hyperparameter grid
Logistic Regression	penalty = {l2, l1, ElasticNet, None}, C = {0.1, 1, 10, 100}
Gradient Boost	learning rate = {0.01, 0.1, 0.5}, n estimators = {50, 100, 500, 1000}, subsample = {0.1, 0.5, 0.9}
Random Forest	n estimators = {100, 500, 1000,2000}, max features = {None, log2, sqrt}
Support Vector Classifier	C = {0.001, 0.1, 1, 10, 100, 1000}, gamma = {0.0001, 0.01, 0.1, 1, 10}, kernel = {rbf, linear, poly, sigmoid}
Xgboost	learning rate = {0.001, 0.1, 0.3, 0.9}, n estimators = {50, 250, 500}
Catboost	depth = {4,6,8,10}, learning rate = {0.01,0.02,0.03,0.04}, iterations = {10,20,50,70,100}

G. SHAP values

Finally, a SHapley Additive exPlanations (SHAP) values analysis was conducted for model explainability using the SHAP library [26] to generate a graph with the ten most influential variables. The features that have the most significant impact on the prediction of MGMT promoter methylation were identified.

III. RESULTS

The proposed framework starts with the exploratory analysis of the considered image dataset. When reviewing the histograms for FLAIR, T1, and T1CE sequences (Figure 1) a greater variety in pixel intensities is observed, with the potential identification of 2 or 3 clusters. These clusters may correspond to different acquisition groups with varying parameters or equipment, with different contrasts (figure 3). The T2 sequence exhibits more uniform values.

After applying the harmonization technique via histogram matching with a reference image, the distribution of the average histograms shifted (Figure 2). In the T1 images, there was an increase in variability. In contrast, in the T1 contrast-enhanced images, a central tendency is observed in most pixels clustered around the peak of the histogram. For the T2 images, the histogram shows higher variability. The pixel intensities were successfully grouped in the FLAIR images, eliminating the identification of two distinct clusters.

Once the variables were extracted, a data frame was created with quantitative values. It comprised 558 rows and 2141 variables, with 535 variables for each image sequence (2140) and one target variable (MGMT). It was decided not to use data balancing techniques, as the target variable was already proportionally distributed with 52.3% for MGMT promoter methylation and 47.7% for non-methylation of the MGMT promoter.

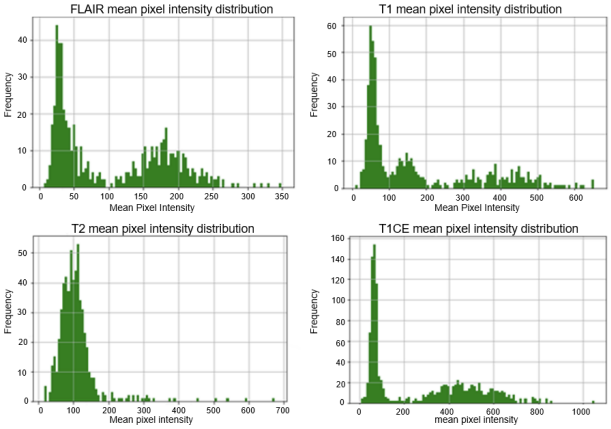

Fig. 1. Histogram of pixel intensities of each sequence before harmonization.

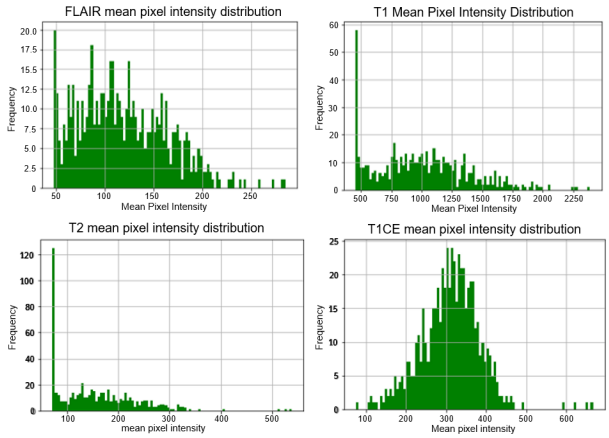

Fig. 2. Histogram of pixel intensities of each sequence after harmonization.

For the models trained using all variables, the Area Under the Curve (AUC) values were as follows: Logistic Regression achieved an AUC of 0.59, Gradient Boosting recorded 0.54, Random Forest yielded 0.52, Support Vector Machine registered 0.55, CatBoost attained 0.51, and XGBoost also reached 0.51. Subsequently, among the models trained with reduced variables, the best-performing model was Logistic Regression, which utilized LASSO for feature selection (Table II).

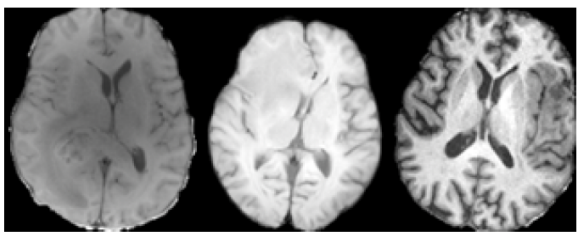

Fig. 3. Contrast difference in T1 sequences

TABLE II
MODEL WITH BEST PERFORMANCE.

Model	Logistic regression	
Feature selection	LASSO (153 features)	
Class	0	1
Precision	0.62	0.60
Recall	0.60	0.61
F1-score	0.61	0.60
Accuracy	0.61	
ROC-AUC	0.67	

The ROC curve graphs illustrate the performance metrics of all models, encompassing both those that incorporated all variables (Figure 4) and those with a reduced number of variables (Figure 5). Notably, for most models, a variable reduction leads to enhanced performance, with the Logistic Regression model exhibiting the best overall performance.

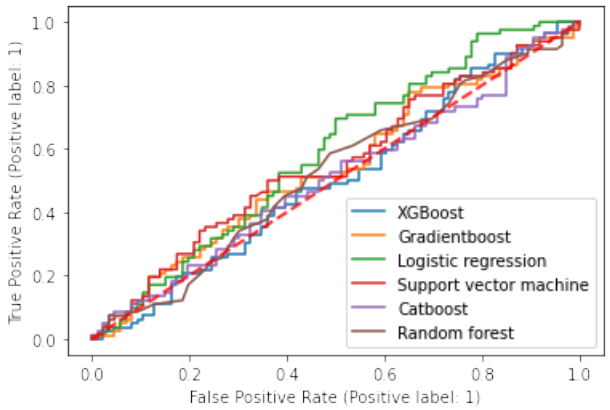

Fig. 4. ROC curves using all features extracted.

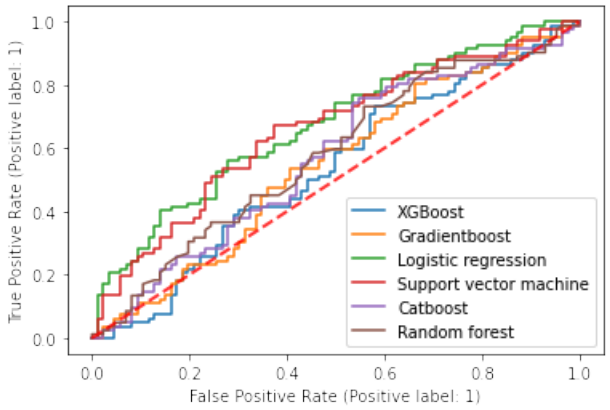

Fig. 5. ROC curves a reduced number of features extracted.

Finally, the explainability of the Model was evaluated with the Shap values (Figure 6). The three most significant variables identified were: a T1CE sequence variable related to texture, a FLAIR variable associated with first order statistics, and a T1 variable of first order statistical type. Among the top ten most prominent variables, texture-related features and first-order

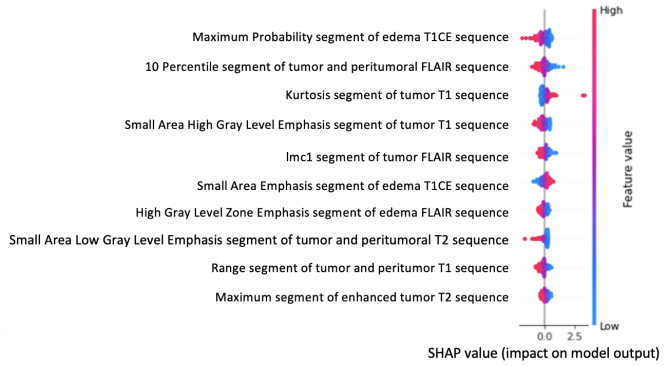

Fig. 6. Shap values results. The local explanation summary shows the direction of the relationship between a variable and class study.

statistics predominated, indicating that shape-related variables did not rank within the top ten in importance.

It was also noted that six out of the ten variables were derived from more extensive segments, specifically those encompassing the entire tumor (the sum of both contrast-enhancing and non-enhancing segments) and segments including both the tumor and edema. Furthermore, it is essential to highlight that eight of the ten variables have a negative impact on the prediction when their values are high, suggesting they significantly influence the likelihood of the genetic biomarker's absence.

IV. DISCUSSION

Research aimed at predicting MGMT gene methylation based on MR imaging is scarce. An alternative involves feature extraction using radiomic methodology, as was performed in this work. However, an important fact is to evaluate the use of all or part of these features (the most relevant ones) to improve classification performance. Suboptimal performance was observed when utilizing all extracted variables across various models, highlighting the necessity of applying feature selection techniques to enhance model effectiveness. These findings are consistent with the study of Lee et al. [27], which also reports improved performance upon employing such techniques.

The model demonstrating the best performance was the Logistic Regression with LASSO feature selection, achieving an AUC of 0.67. This slightly surpasses the results of Palsson et al. [10], which achieved an AUC of 0.598 using a Random Forest model and LASSO selection technique (with 15 variables selected). The enhancement in performance could be attributed to expert's manual tumor area segmentation, in contrast to the study above, which utilized automatic segmentation and image harmonization. Furthermore, this result outperforms the work of Baba, the winner of a Kaggle competition [28], who achieved a roc-auc of 0.62173 using a ResNet model without employing a radionics-based methodology.

In Le et al.'s study [29], using a dataset of 53 patients, an AUC of 0.896 was achieved using an XGBoost model with variable selection via F-score (selecting nine variables).

Meanwhile, the study by Xie et al. [6] attained an accuracy of 85.59% using a Support Vector Machine model with LASSO selection on their dataset, which comprised 98 patient images. Compared to both studies, our model demonstrated lower performance. The superior performance in these studies can be attributed to their use of image sets from a single clinical center, which minimizes the variability often present in multicentric studies. These studies face more significant challenges in achieving high performance due to addressing this variability through specific techniques or algorithms, such as image harmonization techniques.

One of the most relevant modules of the methodology proposed in this research corresponded to the use of Shap values to evaluate the explainability of the best machine learning model used. This methodology allowed us to understand which extracted features had the most significant impact on predicting the different classes. This result is relevant for clinical end users because, as proposed in previous research [30], understanding how an artificial intelligence algorithm generates predictions increases reliability. It gives more significant support to the specialist during diagnosis. It should be noted that this is one of the few works on MGMT methylation identification that includes an interpretability module, and future works may involve new variables such as clinical aspects or expert opinion.

The characteristics that can influence feature extraction from images include artifacts generated by metallic elements and motion artifacts, which can modify the pixel intensities or cause some pixels to represent incorrect values. Furthermore, obtaining multicentric images can increase variability because different equipment is used or different acquisition protocols are followed at each center, resulting in images that have different contrasts. The image database of this study includes images from multiple centers, suggesting the use of various protocols and equipment; this diversity is evident not only in the visual inspection of the images, as seen in the T1 sequence group, where images from the same sequence show different contrasts, but also in the histogram analysis (figure (3). Additionally, motion artifacts are detected in some images.

Compared to one obtained locally, one of the limitations of this database is the variability of the images. In addition, the image acquisition parameters are unknown, and there is a lack of relevant clinical data such as age, sex, blood test results, and radiological findings; these parameters are factors that can enhance the performance of radiomics models [31]. Diffusion magnetic resonance imaging and apparent diffusion coefficient (ADC) maps could also be incorporated, as they are common in most brain MRI protocols and provide quantitative information, which could improve performance, in the study by Ladenhauf et al.'s [32], it was observed that in the peritumoral white matter, both absolute and normalized ADC values were significantly higher in patients with unmethylated MGMT tumors compared to those with methylated MGMT tumor.

V. Conclusion

This research proposed a radiomic-based explainable artificial intelligence methodology for determining MGMT gene methylation. The best performance obtained reached an accuracy of 61% and ROC-AUC of 0.67; it was superior to other investigations using the same dataset. This methodology could be an additional tool for clinical specialists, mainly because it involves an explainable module that allows an understanding of how the predictions are generated and which features significantly impact them.

Although one of the main advantages of non-invasive methodologies over invasive ones is the ability to obtain information more quickly, which facilitates the determination of the recommended treatment, it is also true that some tumors, due to their location and accessibility, are not safe to diagnose through biopsy, as this could damage critical brain areas. Another benefit of opting for non-invasive methodologies is that the patient's overall health status and ability to tolerate a surgical procedure are important to consider, including evaluating preexisting medical conditions, anesthetic risk, and age. Therefore, using non-invasive methodologies offers the opportunity to assess patients who do not meet the requirements for invasive tests, providing a safe alternative.

This research employed a simple image processing algorithm, namely histogram harmonization with a reference image, which marginally improved performance compared to not utilizing this technique. Therefore, future studies could explore more sophisticated methods. Additionally, future research could consider employing a different target variable, such as the direct efficacy of chemotherapy, and identify biomarkers in images based on that. This is because there may be more effective strategies than seeking a biomarker in images for MGMT promoter methylation.

For future studies, it is recommended to not only incorporate additional machine learning models and a broader grid of hyperparameters, along with various feature selection techniques but also to explore further elements of explainability, such as Local Interpretable Model-agnostic Explanations (LIME), which provide explanations for the predictions of any model and feature maps, which visualize areas of an image identified as significant for the model's predictions. Implementing these methods could enhance the explainability and interpretability of the models, improving greater transparency and confidence in decision-making.

Acknowledgment

This work is supported by the Agencia Nacional de Investigación y Desarrollo de Chile, through the grant ANID—Millennium Science Initiative Program ICN2021_004, ANID Fondecyt Postdoctorado 3240078, ANID FONDECYT Regular 1221938, and ANID FONDECYT Regular 1231268.

References

[1] S. Grochans, A. M. Cybulska, D. Simińska, J. Korbecki, K. Kojder, D. Chlubek, and I. Baranowska-Bosiacka, "Epidemiology of glioblastoma multiforme–literature review," *Cancers*, vol. 14, no. 10, p. 2412, 2022.

[2] H. B. Newton, *Handbook of neuro-oncology neuroimaging*. Academic Press, 2016.

[3] L. E. Contreras, "Epidemiología de tumores cerebrales," *Revista Médica Clínica Las Condes*, vol. 28, no. 3, pp. 332–338, 2017.

[4] A. E. Pegg, "Repair of o6-alkylguanine by alkyltransferases," *Mutation Research/Reviews in Mutation Research*, vol. 462, no. 2-3, pp. 83–100, 2000.

[5] V. Quillien, A. Lavenu, L. Karayan-Tapon, C. Carpentier, M. Labussière, T. Lesimple, O. Chinot, M. Wager, J. Honnorat, S. Saikali *et al.*, "Comparative assessment of 5 methods (methylation-specific polymerase chain reaction, methylight, pyrosequencing, methylation-sensitive high-resolution melting, and immunohistochemistry) to analyze o6-methylguanine-dna-methyltranferase in a series of 100 glioblastoma patients," *Cancer*, vol. 118, no. 17, pp. 4201–4211, 2012.

[6] H. Xie, R. Tubbs, and B. Yang, "Detection of mgmt promoter methylation in glioblastoma using pyrosequencing," *International journal of clinical and experimental pathology*, vol. 8, no. 1, p. 636, 2015.

[7] C. Suh, H. Kim, S. Jung, C. Choi, and S. Kim, "Clinically relevant imaging features for mgmt promoter methylation in multiple glioblastoma studies: a systematic review and meta-analysis," *American Journal of Neuroradiology*, vol. 39, no. 8, pp. 1439–1445, 2018.

[8] G. Shukla, G. S. Alexander, S. Bakas, R. Nikam, K. Talekar, J. D. Palmer, and W. Shi, "Advanced magnetic resonance imaging in glioblastoma: a review," *Chin Clin Oncol*, vol. 6, no. 4, p. 40, 2017.

[9] S. Chabert, R. Salas, E. Cantor, A. Veloz, A. Cancino, M. González, F. Torres, and C. Bennett, "Hemodynamic response function description in patients with glioma," *Journal of Neuroradiology*, 2023.

[10] S. Pálsson, S. Cerri, and K. Van Leemput, "Prediction of mgmt methylation status of glioblastoma using radiomics and latent space shape features," in *International MICCAI Brainlesion Workshop*. Springer, 2021, pp. 222–231.

[11] E. Cantor, R. Salas, H. Rosas, and S. Guauque-Olarte, "Biological knowledge-slanted random forest approach for the classification of calcified aortic valve stenosis," *BioData mining*, vol. 14, pp. 1–11, 2021.

[12] P. Franco, J. Sotelo, A. Guala, L. Dux-Santoy, A. Evangelista, J. Rodríguez-Palomares, D. Mery, R. Salas, and S. Uribe, "Identification of hemodynamic biomarkers for bicuspid aortic valve induced aortic dilation using machine learning," *Computers in biology and medicine*, vol. 141, p. 105147, 2022.

[13] M. Zhou, J. Scott, B. Chaudhury, L. Hall, D. Goldgof, K. W. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel *et al.*, "Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches," *American Journal of Neuroradiology*, vol. 39, no. 2, pp. 208–216, 2018.

[14] Y. Liang, Y. Wei, F. Xu, and X. Wei, "MRI-based radiomic models for the preoperative prediction of extramural venous invasion in rectal cancer: A systematic review and meta-analysis," *Clinical Imaging*, vol. 110, Jun. 2024, publisher: Elsevier. [Online]. Available: https://www.clinicalimaging.org/article/S0899-7071(24)00076-7/abstract

[15] M. Tabassum, A. A. Suman, E. Suero Molina, E. Pan, A. Di Ieva, and S. Liu, "Radiomics and Machine Learning in Brain Tumors and Their Habitat: A Systematic Review," *Cancers*, vol. 15, no. 15, p. 3845, Jan. 2023, number: 15 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2072-6694/15/15/3845

[16] G. Tulum, "Novel radiomic features versus deep learning: differentiating brain metastases from pathological lung cancer types in small datasets," *British Journal of Radiology*, vol. 96, no. 1146, p. 20220841, Jun. 2023. [Online]. Available: https://doi.org/10.1259/bjr.20220841

[17] P. Papadimitroulas, L. Brocki, N. C. Chung, W. Marchadour, F. Vermet, L. Gaubert, V. Eleftheriadis, D. Plachouris, D. Visvikis, G. C. Kagadis, and M. Hatt, "Artificial intelligence: Deep learning in oncological radiomics and challenges of interpretability and data harmonization," *Physica Medica: European Journal of Medical Physics*, vol. 83, pp. 108–121, Mar. 2021, publisher: Elsevier. [Online]. Available: https://www.physicamedica.com/article/S1120-1797(21)00125-3/fulltext

[18] R. Zhang, M. Hong, H. Cai, Y. Liang, X. Chen, Z. Liu, M. Wu, C. Zhou, C. Bao, H. Wang, S. Yang, and Q. Hu, "Predicting the pathological invasiveness in patients with a solitary pulmonary nodule via Shapley additive explanations interpretation of a tree-based machine learning radiomics model: a multicenter study," *Quantitative Imaging in Medicine and Surgery*, vol. 13, no. 12, pp. 7828–7841, Dec. 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10722047/

[19] L. Brocki and N. C. Chung, "Integration of Radiomics and Tumor Biomarkers in Interpretable Machine Learning Models," *Cancers*, vol. 15, no. 9, p. 2459, Jan. 2023, number: 9 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2072-6694/15/9/2459

[20] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F. C. Kitamura, S. Pati *et al.*, "The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification," *arXiv preprint arXiv:2107.02314*, 2021.

[21] E. Cavieres, C. Tejos, R. Salas, and J. Sotelo, "Automatic segmentation of brain tumor in multi-contrast magnetic resonance using deep neural network," in *18th International Symposium on Medical Information Processing and Analysis*, vol. 12567. SPIE, 2023, pp. 81–89.

[22] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[23] F. Hu, A. A. Chen, H. Horng, V. Bashyam, C. Davatzikos, A. Alexander-Bloch, M. Li, H. Shou, T. D. Satterthwaite, M. Yu *et al.*, "Image harmonization: A review of statistical and deep learning methods for removing batch effects and evaluation metrics for effective harmonization," *NeuroImage*, p. 120125, 2023.

[24] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka *et al.*, "3d slicer as an image computing platform for the quantitative imaging network," *Magnetic resonance imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[25] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J.-C. Fillon-Robin, S. Pieper, and H. J. W. L. Aerts, "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, 2017. [Online]. Available: https://doi.org/10.1158/0008-5472.CAN-17-0339

[26] S. M. Lundberg *et al.*, "Shap: Shapley additive explanations," https://github.com/slundberg/shap, 2021.

[27] S.-H. Lee, H. Park, and E. S. Ko, "Radiomics in breast imaging from techniques to clinical applications: a review," *Korean journal of radiology*, vol. 21, no. 7, p. 779, 2020.

[28] S. Bakas, U. Baid, E. Calabrese, C. Carr, E. Colak, K. Farahani, A. Flanders, J. Kalpathy-Cramer, F. Kitamura, B. Menze, J. Rudie, J. Mongan, J. Elliott, L. Prevedello, M. Riopel, and R. T. Shinohara, "Rsna-miccai brain tumor radiogenomic classification," 2021. [Online]. Available: https://kaggle.com/competitions/rsna-miccai-brain-tumor-radiogenomic-classification

[29] N. Q. K. Le, D. T. Do, F.-Y. Chiu, E. K. Y. Yapp, H.-Y. Yeh, and C.-Y. Chen, "XGBoost improves classification of mgmt promoter methylation status in idh1 wildtype glioblastoma," *Journal of personalized medicine*, vol. 10, no. 3, p. 128, 2020.

[30] D. Shin, "The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable ai," *International Journal of Human-Computer Studies*, vol. 146, p. 102551, 2021.

[31] S. Pszczolkowski, J. P. Manzano-Patrón, Z. K. Law, K. Krishnan, A. Ali, P. M. Bath, N. Sprigg, and R. A. Dineen, "Quantitative CT radiomics-based models for prediction of haematoma expansion and poor functional outcome in primary intracerebral haemorrhage," *European Radiology*, vol. 31, no. 10, pp. 7945–7959, Oct. 2021.

[32] V. K. Ladenhauf, M. Galijasevic, J. Kerschbaumer, C. F. Freyschlag, M. Nowosielski, A. M. Birkl-Toeglhofer, J. Haybaeck, E. R. Gizewski, S. Mangesius, and A. E. Grams, "Peritumoral ADC Values Correlate with the MGMT Methylation Status in Patients with Glioblastoma," *Cancers*, vol. 15, no. 5, 2023. [Online]. Available: https://www.mdpi.com/2072-6694/15/5/1384