POLITECNICO DI TORINO
Repository ISTITUZIONALE

Preoperative Spatial Risk Mapping of Glioblastoma Recurrence: A Radiomics-Based Framework for Surgical Planning

Original
Preoperative Spatial Risk Mapping of Glioblastoma Recurrence: A Radiomics-Based Framework for Surgical Planning / Seoni, S., La Paglia, F., Salvi, M., Morello, A., Bergoglio, G., Garbossa, D., Salvi, M., Cofano, F.. - In: APPLIED SCIENCES. - ISSN 2076-3417. - 16:14(2026). [10.3390/app16147344]

Availability:
This version is available at: 11583/3013412 since: 2026-07-22T12:22:43Z

Publisher:
MDPI

Published
DOI:10.3390/app16147344

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

applied sciences

MDPI

Article

Preoperative Spatial Risk Mapping of Glioblastoma Recurrence: A Radiomics-Based Framework for Surgical Planning

Silvia Seoni [1], Federica La Paglia [1], Matteo Salvi [1], Alberto Morello [2], Greta Bergoglio [2], Diego Garbossa [2], Massimo Salvi [1,*] and Fabio Cofano [2]

[1] InVisioning Lab, PoliToBIOMed Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy; silvia.seoni@polito.it (S.S.); fede.lapaglia@gmail.com (F.L.P.); matteo.salvi@polito.it (M.S.)

[2] Neurosurgery Unit, Department of Neuroscience "Rita Levi Montalcini", AOU Città della Salute e della Scienza di Torino, University Hospital, University of Turin, 10126 Turin, Italy; greta.bergoglio@edu.unito.it (G.B.); diego.garbossa@unito.it (D.G.); fabio.cofano@unito.it (F.C.)

* Correspondence: massimo.salvi@polito.it

Abstract

Glioblastoma recurrence remains nearly inevitable despite maximal resection and adjuvant therapy, with most relapses occurring within or adjacent to the original tumor site. Conventional MRI underestimates tumor infiltration beyond contrast-enhancing margins, limiting preoperative identification of peritumoral regions at higher risk of recurrence. We developed a radiomics-based machine-learning framework to generate preoperative spatial recurrence risk maps from routine MRI. Preoperative T1-weighted contrast-enhanced and FLAIR images from 79 patients with glioblastoma were analyzed. The cohort was divided at the patient level into a training set (n = 63) and an independent test set (n = 16). Using a balanced spatial ROI sampling strategy, local radiomic features were extracted from tumor and peritumoral regions and linked to recurrence sites identified on follow-up MRI acquired after at least 12 months after surgery. A CatBoost classifier achieved an AUC of 0.743 and a recall of 0.856 on an independent test set. The framework also generated preoperative probability maps that showed qualitative spatial correspondence with observed recurrence locations. These findings indicate that radiomic patterns from standard preoperative MRI may contain spatially localized information associated with future relapse. The proposed approach supports the feasibility of preoperative spatial risk stratification and may provide useful information for surgical planning and subsequent treatment strategies.

Keywords: glioblastoma; recurrence prediction; radiomics; machine learning; preoperative MRI; spatial risk mapping; texture analysis; surgical planning

Check for updates

Academic Editor: Nikolaos Kourkoumelis

Received: 10 June 2026
Revised: 16 July 2026
Accepted: 19 July 2026
Published: 22 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Glioblastoma (GBM) is the most common and aggressive primary malignant brain tumor in adults and remains associated with a poor prognosis despite advances in surgical techniques and adjuvant therapies [1,2]. Maximal safe resection followed by radiotherapy and chemotherapy represents the current standard of care; however, local recurrence remains almost inevitable, with the majority of relapses occurring within or adjacent to the original tumor site [3–5]. This behavior reflects the highly infiltrative nature of GBM, characterized by diffuse tumor cell migration into surrounding brain tissue, which severely limits long-term disease control [6].

Preoperative surgical planning is primarily guided by contrast-enhanced T1-weighted magnetic resonance imaging (MRI), which delineates regions of blood–brain barrier disruption and represents the current clinical standard for tumor visualization. However, conventional MRI has limited sensitivity for detecting the full biological extent of infiltrative tumor beyond the contrast-enhancing margins [7]. Peritumoral abnormalities visible on FLAIR sequences are commonly interpreted as vasogenic edema, although histopathological evidence demonstrates that these regions often harbor infiltrative tumor cells that contribute to recurrence [8]. Therefore, surgeons are required to balance maximal resection against the risk of neurological deficits in the absence of reliable imaging markers capable of identifying peritumoral regions at higher risk of future recurrence [1].

Given the limited sensitivity of conventional MRI in detecting microscopic tumor infiltration beyond contrast-enhancing margins, increasing attention has been directed toward quantitative imaging approaches capable of capturing subtle tissue heterogeneity. In this context, radiomics has emerged as a promising framework, enabling the extraction of high-dimensional quantitative features from standard medical images to characterize tumor and peritumoral tissue properties beyond visual assessment [9,10].

The high dimensionality of radiomic data, together with the complex relationships among imaging features, has made machine learning (ML) a natural choice for modeling associations with clinical outcomes. In GBM, several radiomics studies based on preoperative MRI have explored patient-level prediction, although with different methodological objectives. For example, Ko et al. [11] focused on recurrence/progression prediction using conventional radiomic features and SVM classification, whereas Gao et al. [12] investigated overall survival prediction by combining intratumoral and peritumoral radiomics with Random Forest models. Other studies have addressed molecular characterization or glioma grading through radiomics and deep learning approaches [13–15]. Despite these methodological differences, all of these approaches generate patient-level predictions and therefore provide limited support for surgical decisions requiring spatially localized information.

More recently, attention has shifted toward spatial modeling of recurrence risk. Most of these studies, however, are based on postoperative MRI acquired shortly after resection and are designed to generate voxel-wise or regional probability maps for radiotherapy planning [16–18]. These approaches typically exploit the postoperative anatomy and the presence of residual tumor tissue to estimate local recurrence risk, making them particularly suitable for radiotherapy target definition and postoperative management. While these methods are valuable for characterizing residual disease, they are not directly applicable to preoperative surgical planning, since resection and cavity formation substantially modify both anatomy and tissue context [19].

Only a limited number of studies have attempted to characterize recurrence or infiltration patterns directly from preoperative imaging. Some works rely on advanced MRI modalities, such as diffusion tensor imaging or perfusion imaging, to capture microstructural and physiological characteristics associated with tumor infiltration [20,21]. Although these modalities may improve tissue characterization, they are not routinely acquired in standard clinical protocols, limiting reproducibility and broader clinical applicability. Other studies have introduced spatial information through coarse representations, such as quadrant-based prediction or conceptual models of tumor growth [22,23]. While these approaches provide useful spatial context, they do not generate dense, patient-specific probability maps capable of capturing the highly focal and heterogeneous nature of glioblastoma infiltration. Consequently, there remains a need for methods capable of estimating localized recurrence risk directly from routine preoperative MRI to support surgical decision-making.

A schematic comparison of the most relevant studies is reported in Table 1, highlighting differences in imaging timing, predictive output, methodological approach, and

reported performance. Notably, all of these studies rely on private single-center or multi-center datasets, reflecting the limited availability of publicly accessible imaging cohorts with spatially resolved recurrence annotations.

Table 1. Overview of existing approaches for glioblastoma recurrence prediction.

Study	Imaging (n. of Patients)	Timing	Output	Approach
Metz et al. [20], 2020	DTI + FLAIR + T1ce (n = 35)	Pre-op	Spatial (patch-based)	Statistics + ML
Ko et al. [11], 2021	T1ce, T2 (n = 128)	Pre-op	Patient-level	ML (SVM)
Chougule et al. [21], 2022	T1ce, FLAIR, ADC (n = 29)	Pre-op	Spatial (voxel-wise)	ML (RF)
Cepeda et al. [16], 2023	mpMRI (n = 55)	Post-op	Spatial (voxel-wise)	ML (CatBoost)
Gao et al. [12], 2024	T1, T2, FLAIR (n = 163)	Pre-op	Patient-level (OS)	ML (RF)
Li et al. [22], 2025	FLAIR, T1ce (n = 149)	Pre-op	Spatial (region-based)	ML (LightGBM)
Cepeda et al. [17], 2025	mpMRI (n = 114 + 25 prospective)	Post-op	Spatial (voxel-wise)	ML (CatBoost)
This work	T1ce + FLAIR (n = 79)	Pre-op	Spatial (ROI-based probability maps)	ML (CatBoost)

DTI: diffusion tensor imaging; FLAIR: fluid-attenuated inversion recovery; T1ce: T1-weighted contrast-enhanced; T1/T2: T1- and T2-weighted imaging; ADC: apparent diffusion coefficient; mpMRI: multiparametric MRI; ML: machine learning; SVM: support vector machine; RF: Random Forest; OS: overall survival. Spatial indicates voxel-, patch-, or region-level prediction; patient level indicates global outcome prediction.

In this context, the present work introduces a radiomics-based ML framework for estimating preoperative, spatially resolved glioblastoma recurrence risk from standard MRI. The framework was developed and evaluated on a cohort of 79 patients with histopathologically confirmed glioblastoma. The proposed approach generates probabilistic maps of local recurrence based on preoperative imaging, aiming to support surgical planning through spatial risk stratification, optimize the extent of tumor resection, and ultimately improve progression-free and overall survival in future clinical applications. While previous studies have focused either on patient-level prediction or on spatial recurrence modeling from postoperative MRI, our framework addresses spatial recurrence risk directly from routine preoperative MRI, with the aim of supporting surgical planning. The main contributions of this study are as follows:

- We demonstrate the feasibility of estimating spatially localized glioblastoma recurrence risk from routine preoperative MRI, moving beyond patient-level recurrence prediction toward region-specific risk assessment.
- We show that local radiomic patterns extracted from co-registered T1ce and FLAIR images contain information associated with future recurrence locations, supporting the relevance of peritumoral tissue characterization before surgery.
- We provide an exploratory qualitative assessment of the generated spatial recurrence risk maps by comparing predicted high-risk regions with recurrence areas identified on follow-up imaging.
- We investigate the radiomic signatures underlying recurrence-risk prediction using SHAP-based explainability, highlighting the contribution of texture- and intensity-related features.

This paper is organized as follows. Section 2 describes the proposed methodology, Section 3 presents the experimental results, and Sections 4 and 5 discuss the findings and conclude the paper.

2. Materials and Methods

The proposed pipeline consists of two main stages: a training stage for ROI-level recurrence modeling and an inference stage for spatial risk map generation (Figure 1). During training, preoperative MRI volumes are preprocessed and segmented to identify binary masks corresponding to tumor-related regions. Local regions of interest (ROIs) are then extracted using the proposed sampling strategy, and radiomic features are computed from co-registered T1ce and FLAIR images. These features are used to train a supervised machine-learning classifier that estimates the probability of future recurrence at the local level.

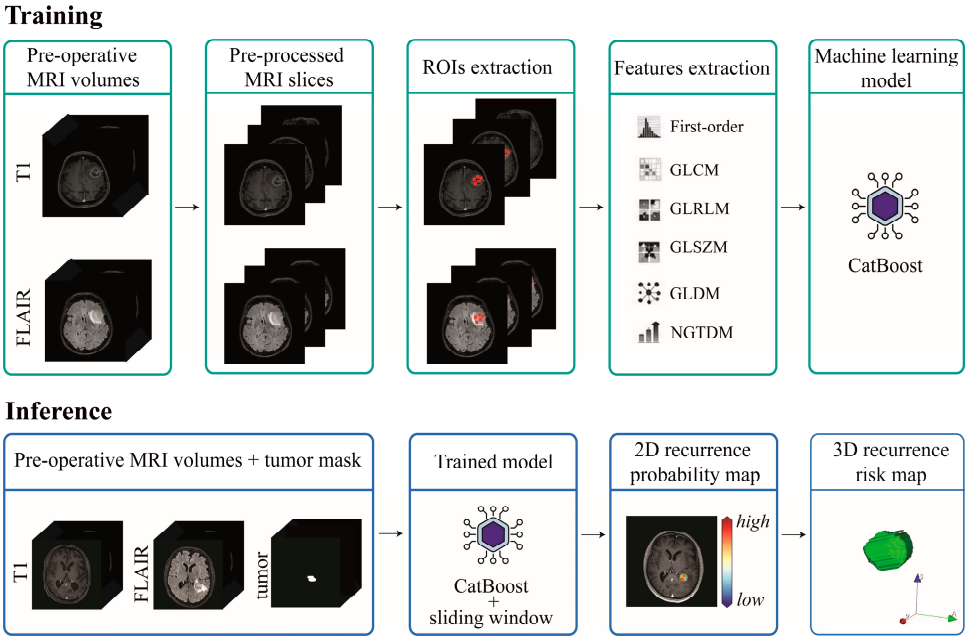


Figure 1. Overview of the proposed framework. Preoperative MRI (T1 and FLAIR) is preprocessed and segmented, local ROIs are sampled for radiomic feature extraction, and a CatBoost classifier is trained to predict local recurrence risk. During inference, dense ROI sampling is used to reconstruct spatially resolved preoperative recurrence probability maps.

During inference, the trained classifier is applied to unseen preoperative MRI scans. A dense sliding-window strategy with unit stride is used to extract ROIs across the preoperative lesion on co-registered T1ce and FLAIR images, and the same radiomic features are computed for each modality-specific ROI. The corresponding recurrence probabilities are then averaged for each ROI and mapped back onto the associated image region; overlapping ROI contributions are subsequently averaged to reconstruct slice-wise preoperative recurrence risk maps.

2.1. Dataset and MRI Acquisition

The dataset analyzed in this study consisted of preoperative and follow-up MRI scans from 79 adult patients with histopathologically confirmed glioblastoma. All patients were treated at a single institution (A.O.U. Città della Salute e della Scienza di Torino, Turin, Italy) and underwent surgical resection followed by routine clinical follow-up.

Patients were selected based on the following criteria: (1) tissue-based diagnosis of a previously untreated high-grade glioma according to the WHO 2016/2021 classification [24]; (2) age $\geq$ 18 years; (3) availability of preoperative MRI, including contrast-enhanced T1-weighted and T2/FLAIR sequences, together with postoperative/follow-up imaging available for review; and (4) follow-up of at least 12 months after histopathological diagnosis. Before inclusion, all examinations were reviewed by A.M. to confirm

the availability of the required MRI sequences, adequate follow-up, and that the images were of sufficient quality for the planned radiological and radiomic analyses. Patients with low-grade gliomas, infratentorial tumors, or multicentric lesions were excluded. Among the eligible cases, all patients (n = 79) were included regardless of recurrence size, and no additional exclusions were applied based on image processing or preprocessing results.

For each patient, preoperative MRI included contrast-enhanced T1-weighted imaging (T1ce) acquired after intravenous administration of a gadolinium-based contrast agent, as well as fluid-attenuated inversion recovery (FLAIR) sequences. Follow-up MRI scans were available for all patients and included T1ce and FLAIR sequences used to identify tumor recurrence. All examinations were performed using the same 3T MRI scanner (Ingenia, Philips Healthcare, Best, The Netherlands) to minimize inter-scanner variability and ensure imaging consistency across the cohort.

At the time of hospitalization, written consent for personal, biological, and radiological data processing for scientific purposes was explicitly asked. The data were anonymized before processing, as instructed by the EU General Data Protection Regulation, using the specific function available in the HOROS©DICOM image viewer. The present study was approved by the local Institutional Review Board (n. 00162/2022).

Preoperative images were used for radiomic feature extraction and model development, while follow-up scans served solely for spatial localization of recurrence regions during target definition. Tumor and recurrence regions were manually segmented using 3D Slicer (version 5.6.2) by an experienced neurosurgeon (F.C.). Primary lesions and the first recurrence event, defined according to the RANO criteria [25], were manually segmented on contrast-enhanced T1-weighted (T1ce) and FLAIR MRI sequences. Information from both modalities was integrated to improve lesion margin delineation, and all segmentations were radiologically confirmed. The contrast-enhancing (CE) tumor volume was measured on T1ce images, whereas the non-contrast-enhancing (non-CE) tumor component, defined as signal abnormalities extending beyond the enhancing tumor margins, was assessed on FLAIR images.

To ensure proper generalization, dataset partitioning was performed at the patient level, ensuring that no data from the same patient were shared across subsets. The cohort was randomly divided into a training set (80%, n = 63) and an independent test set (20%, n = 16). Hyperparameter optimization was conducted exclusively within the training set using k-fold cross-validation, while the held-out test set was reserved for final performance evaluation.

2.2. Image Pre-Processing

Before radiomic feature extraction, all MRI volumes underwent the same preprocessing steps to ensure anatomical alignment between sequences and reduce inter-subject variability in voxel geometry and intensity scale. Raw DICOM images were first converted to NIfTI format. For each patient, the preoperative T1ce scan was used as the reference volume. The corresponding preoperative FLAIR and follow-up T1ce scans were rigidly co-registered to the preoperative T1ce image in 3D Slicer (version 5.10.0). Registration consisted of an initial rigid alignment followed by manual refinement, performed by simultaneously inspecting the axial, coronal, and sagittal views to maximize the correspondence of anatomical landmarks and lesion boundaries. This semi-manual workflow was adopted because multimodal registration between preoperative and follow-up MRI is complicated by surgical cavity formation and tumor progression, making visual refinement necessary to ensure the best possible anatomical correspondence. Final alignment was verified slice by slice before radiomic feature extraction and ROI labeling.

After registration, all images were resampled to an isotropic voxel spacing of $1 \times 1 \times 1$ mm^3 using B-spline interpolation. To ensure a uniform input size across subjects, the resampled volumes were then adjusted to a matrix size of $240 \times 240 \times 170$ voxels, corresponding to the most frequent image dimensions in the dataset.

To reduce inter-scan intensity variability, a fixed intensity window (window width = 1860; window level = 1070) was empirically selected after preliminary inspection of the dataset and applied consistently to all images. This approach was adopted to standardize intensity values while preserving the intensity range containing the relevant tumor-related information. Compared with adaptive histogram-based normalization methods, which may produce variable intensity scaling depending on tumor extent and image intensity distribution, the fixed-window approach provided more consistent normalization across patients. Values outside the resulting range [140, 2000] were clipped, and the remaining intensities were linearly normalized to the range [0, 1] using min–max scaling applied independently to each volume. An example of the preprocessed and co-registered T1ce, FLAIR, and follow-up T1ce images is shown in Figure 2.

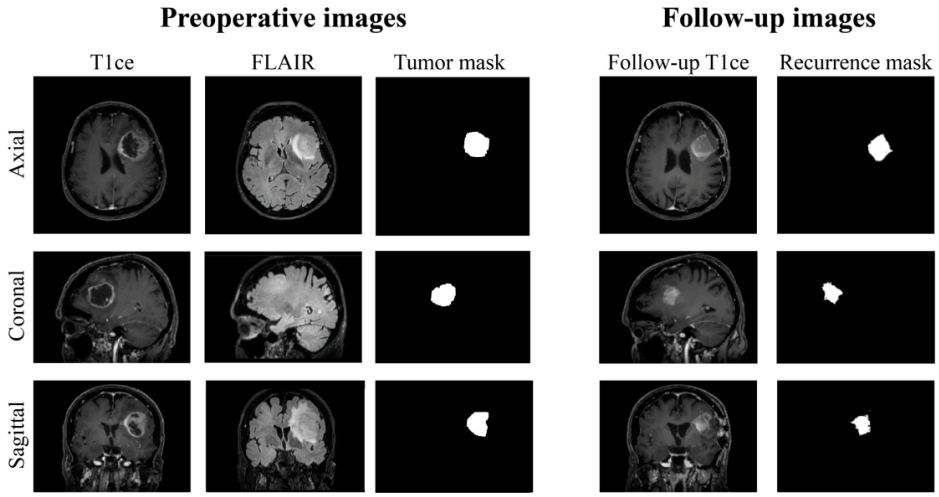


Figure 2. Example of preoperative and follow-up MRI after preprocessing and co-registration. From left to right: preoperative T1ce image, FLAIR image, tumor mask, follow-up T1ce image, and manually annotated recurrence mask. Rows show axial, coronal, and sagittal views.

2.3. Radiomics-Based Spatial Modeling Framework

The proposed framework uses radiomic analysis to characterize local tissue patterns on preoperative MRI that may be associated with future tumor recurrence. It consists of four main components: (1) ROI extraction, sampling, and labeling using preoperative MRI and spatially matched follow-up recurrence information; (2) radiomic feature extraction from co-registered T1ce and FLAIR images; and (3) supervised machine-learning classification at the ROI level. Together, these steps enable spatially resolved prediction of recurrence risk from preoperative imaging. The individual components are described in the following subsections.

2.3.1. ROI Extraction, Sampling and Labeling

Radiomic features were computed from two-dimensional ROIs extracted from axial slices of the preoperative MRI volumes. ROIs were defined as square patches of fixed size 5×5 pixels and were sampled within the segmented preoperative tumor lesion. A sparse sampling strategy was adopted, with up to five ROIs extracted from each slice depending on the spatial extent of the lesion. Consequently, slices containing smaller tumor regions yielded fewer valid ROIs. Each sampled ROI was extracted from the co-registered preoperative T1ce and FLAIR images.

After ROI extraction, each sampled ROI was assigned a binary label for supervised learning using follow-up recurrence information. Tumor recurrence was defined on follow-up T1ce MRI based on manual segmentation of the enhancing recurrent tumor, as described in Section 2.1. These recurrence masks were used exclusively to generate the ground-truth labels and were not provided as input to the predictive model.

For each sampled ROI, the corresponding spatial location was matched to the co-registered follow-up recurrence mask. An ROI was labeled as positive (recurrence = 1) if at least one pixel overlapped the recurrence mask; otherwise, it was labeled as negative (recurrence = 0). This permissive criterion was adopted to maximize sensitivity in identifying local regions associated with future recurrence. Accordingly, ROIs overlapping the recurrence mask were considered recurrence ROIs, whereas ROIs outside the recurrence mask were considered non-recurrence ROIs (Figure 3).

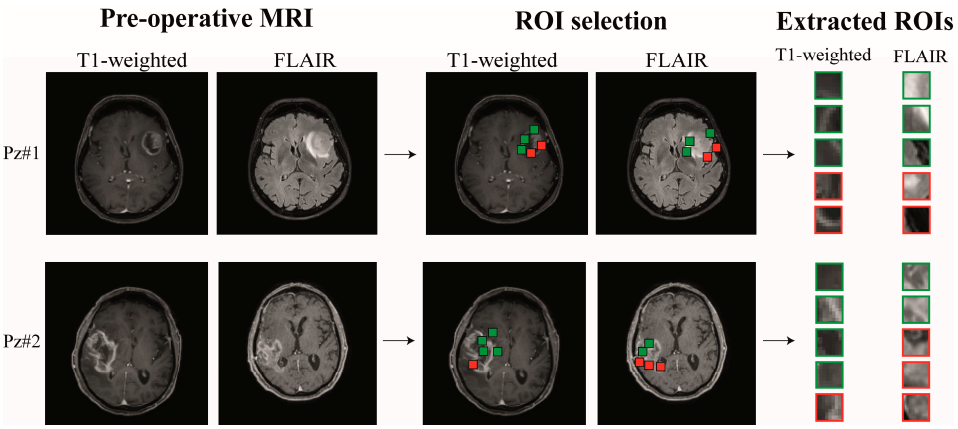


Figure 3. Illustration of the proposed ROI extraction and labeling strategy in two representative cases. Square ROIs are sampled within the segmented preoperative tumor lesion on co-registered preoperative T1ce and FLAIR images. The dashed contour indicates the recurrence mask identified on follow-up T1ce MRI after co-registration. ROIs overlapping the recurrence mask are shown in red and assigned a positive label (recurrence = 1), whereas ROIs outside the recurrence mask are shown in green and assigned a negative label (recurrence = 0).

This sampling strategy was designed to reduce class imbalance by restricting ROI extraction to the segmented preoperative tumor region, thereby avoiding the large predominance of negative samples that would result from conventional random or grid-based sampling over the entire image volume. However, the final proportion of positive and negative ROIs depends on the subsequent labeling process, which is determined by the spatial overlap between each ROI and the recurrence mask, as well as by the extent and location of the recurrence itself. Consequently, although the resulting dataset is not perfectly balanced, the proposed strategy substantially mitigates class imbalance while preserving the spatial relationship between the preoperative lesion and the observed recurrence, as further supported by the ablation study presented in Section 3.3. The overall number of patients, MR slices, and extracted ROIs included in the analysis is summarized in Table 2.

Table 2. Summary of the dataset used for radiomic analysis, including the number of patients, MR slices, and extracted ROIs.

Subset	Number of Patients	MR Slices	ROIs (Positive vs. Negative)
Training	63	4303	17,791/12,151
Test	16	1054	4319/2957
Total	79	5357	22,110/15,108

2.3.2. Radiomic Feature Extraction

Radiomic features were extracted separately from the preoperative T1ce and FLAIR ROIs to characterize local intensity and texture patterns. Because ROIs were defined on individual axial slices, all features were computed in 2D.

First-order statistical features were extracted to describe the intensity distribution within each ROI, including descriptors of central tendency, dispersion, and distribution shape. To capture local spatial heterogeneity, texture features were computed using the PyRadiomics library (version 3.0.1) [26]. These included features derived from the Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM), Gray-Level Dependence Matrix (GLDM), and Neighboring Gray Tone Difference Matrix (NGTDM) [27,28]. These feature families were selected because they represent the standard radiomic descriptors most commonly adopted for quantitative analysis of medical images and are fully supported by the PyRadiomics library following the recommendations of the Imaging Biomarker Standardization Initiative (IBSI) [29].

Radiomic features extracted from T1ce and FLAIR ROIs were used to build the input dataset for the machine-learning model. A summary of the extracted radiomic feature families is provided in Table 3.

Table 3. Radiomic feature families extracted from each modality-specific ROI. All texture features were computed in 2D. Directionality refers to whether features depend on spatial orientation. For GLCM and GLRLM, texture features were computed across multiple in-plane directions and averaged to obtain rotation-independent descriptors.

Feature Family	No. of Features	Directionality
First-order statistics	10	Not directional
GLCM	12	Averaged across directions
GLRLM	8	Averaged across directions
GLSZM	8	Direction-independent
GLDM	7	Direction-independent
NGTDM	5	Direction-independent
Total per modality	50	

Gray-Level Co-occurrence Matrix; GLRLM: Gray-Level Run Length Matrix; GLSZM: Gray-Level Size Zone Matrix; GLDM: Gray-Level Dependence Matrix; NGTDM: Neighboring Gray Tone Difference Matrix.

2.3.3. Machine Learning Classifier

The predictive framework was based on the CatBoost gradient boosting algorithm. CatBoost was selected for its robustness in heterogeneous tabular feature spaces, its ability to capture complex non-linear relationships, and its strong generalization performance on medium-sized medical imaging datasets [30,31]. The recurrence prediction task was formulated as a binary classification problem at the ROI level, where each sampled ROI was labeled as recurrence-positive or recurrence-negative. Radiomic features extracted from preoperative MRI were used as input to the classifier.

Model hyperparameters were optimized exclusively on the training set using k-fold cross-validation. The hyperparameter search included the number of trees, learning rate, tree depth, and L2 regularization coefficient. Further details on hyperparameter tuning are reported in Section 3.1.

Model performance was evaluated on the independent test set using the area under the receiver operating characteristic curve (AUC), accuracy, precision, recall, and F1-score. The test set was strictly held out during both training and hyperparameter tuning to ensure an unbiased estimate of performance. For comparison, additional machine-learning

models, including Random Forest and XGBoost, were trained and evaluated under the same experimental conditions.

2.4. Generation of Preoperative Recurrence Risk Maps

After model inference, recurrence probabilities were estimated from ROI-based radiomic features extracted from the preoperative MRI volumes. For each sampled window, radiomic features were computed separately from the preoperative T1ce and FLAIR images using the same preprocessing and feature extraction pipeline described above. The trained CatBoost model was then applied to each modality-specific feature set, yielding two probability estimates for that spatial location. The final recurrence probability assigned to the window was obtained by averaging the two modality-specific predictions.

To generate spatially resolved recurrence risk maps, a dense sliding-window inference procedure was applied within and around the tumor mask on the preoperative MRI volumes. Specifically, square windows of size 5×5 pixel were extracted with a stride of one pixel, and recurrence probabilities were computed for all valid window locations. The resulting probabilities were projected back onto their corresponding spatial coordinates to generate slice-wise probability maps.

In regions with overlapping windows, probability values were locally averaged to improve spatial continuity. Slice-wise maps were then combined across axial slices to obtain a three-dimensional recurrence risk volume for each patient. The resulting maps provide a spatially resolved representation of recurrence likelihood and can be visualized as color-coded overlays on anatomical MRI sequences.

2.5. Model Explainability and Radiomic Pattern Interpretation

To better understand the radiomic patterns driving model predictions, SHAP (SHapley Additive exPlanations) was applied to the trained CatBoost model using the TreeExplainer implementation [32]. All explainability analyses were performed on the independent test set.

First, feature-family importance was assessed by computing the mean absolute SHAP value for each radiomic feature across all ROIs in the test set. Features were then grouped into their respective families, including first-order statistics and texture features derived from GLCM, GLRLM, GLSZM, GLDM, and NGTDM. Family-level importance was obtained by summing the contributions of the features belonging to each group.

Second, individual feature importance was evaluated by ranking all radiomic features according to their mean absolute SHAP value across the test ROIs. This analysis was used to identify the descriptors that contributed most strongly to the model predictions.

Third, SHAP summary plots (beeswarm plots) were generated for the highest-ranking features to visualize both the magnitude and direction of their contribution across ROIs. These plots also illustrate the variability of feature effects across samples.

3. Results

3.1. Hyperparameter Tuning

Before hyperparameter optimization, the quality of the preoperative-to-follow-up image registration was quantitatively assessed across the study cohort. Since the preoperative T1-weighted contrast-enhanced (T1ce) image was used as the reference volume during registration, the evaluation was performed between the preoperative T1ce image and the corresponding registered follow-up T1ce image. Registration quality was assessed using the Structural Similarity Index Measure (SSIM) and Normalized Mutual Information (NMI), yielding mean values of 0.86 ± 0.05 and 1.67 ± 0.09, respectively.

Hyperparameter optimization for the CatBoost classifier was performed exclusively on the training set using stratified 5-fold cross-validation, so as to preserve the class distribution across folds. Bayesian optimization was implemented through BayesSearchCV (scikit-optimize 0.10.2), using ROC-AUC as the optimization objective.

A total of 50 hyperparameter combinations were explored within predefined search ranges, including the learning rate, number of estimators, tree depth, and loss function. For each configuration, performance was averaged across the five validation folds, and the set of hyperparameters yielding the highest mean ROC-AUC was selected. The final CatBoost model was then retrained on the full training set using the optimal configuration. The selected hyperparameters are reported in Table 4.

Table 4. Hyperparameter search space and selected configuration for the proposed CatBoost model.

Hyperparameter	Search Range	Selected Value
Learning rate	[0.01, 0.3] (log-uniform prior)	0.0488
Number of estimators	[10, 500]	500
Tree depth	[1, 10]	7
Loss function	{LogLoss, CrossEntropy}	CrossEntropy

3.2. ROI-Based Performance of the Proposed Model

Model performance was first examined within the training set to assess stability across the stratified 5-fold cross-validation procedure. For each fold, the CatBoost model was trained on four folds and evaluated on the corresponding validation fold. Across the five validation folds, the model achieved a mean AUC of 0.7583 $\pm$ 0.0051 and a mean F1-score of 0.7569 $\pm$ 0.0047.

As an additional assessment of robustness, the five-fold-specific models obtained during cross-validation were also applied to the independent test set. In this setting, the mean test AUC was 0.7428 $\pm$ 0.0059 and the mean F1-score was 0.7708 $\pm$ 0.0047, with a mean recall of 0.8561 $\pm$ 0.0055. The close agreement between validation and independent test performance suggests that the selected hyperparameters generalized well to the held-out test cohort.

Following this robustness assessment, the final CatBoost model was retrained on the complete training set using the selected hyperparameters to maximize the amount of training data available for model development and to obtain a single optimized model for all subsequent analyses. When evaluated on the independent test set, the final CatBoost model achieved an overall ROI-level AUC of 0.743, an accuracy of 0.698, a precision of 0.701, a recall of 0.856, and an F1-score of 0.771. To account for inter-patient variability, we additionally performed a patient-wise ROI evaluation by computing the performance metrics independently for each of the 16 test patients and reporting the corresponding mean $\pm$ standard deviation. Since all test patients contributed both recurrence-positive and recurrence-negative ROIs, with relatively balanced class proportions (ranging approximately from 40:60 to 60:40), all metrics, including AUC, could be computed for every patient. The patient-wise analysis yielded a mean AUC of 0.748 $\pm$ 0.118, a mean accuracy of 0.702 $\pm$ 0.097, a mean precision of 0.708 $\pm$ 0.104, a mean recall of 0.847 $\pm$ 0.129 and a mean F1-score of 0.774 $\pm$ 0.081, closely matching the pooled ROI-level estimates. The results of both analyses are summarized in Table 5.

Table 5. Final ROI-level performance of the proposed CatBoost model on the independent test set.

Metric	ROI-Level	Patient-Wise ROI (Mean ± SD)
AUC	0.743	0.748 ± 0.118
Accuracy	0.698	0.702 ± 0.097
Precision	0.701	0.708 ± 0.104
Recall	0.856	0.847 ± 0.129
F1-score	0.771	0.774 ± 0.081

3.3. Ablation Study

An ablation study was performed to evaluate three key components of the proposed pipeline: (i) ROI size, (ii) ROI sampling strategy, and (iii) classifier selection. For each configuration, hyperparameter optimization was carried out on the training set using stratified 5-fold cross-validation, followed by retraining on the full training set and final evaluation on the independent test set.

As a preliminary analysis, different ROI sizes (3×3, 5×5 and 7×7 pixels) were evaluated to identify the most suitable spatial support for radiomic feature extraction (Table 6). Smaller ROIs (3×3) generated a larger number of training samples but captured only a limited local tissue context, resulting in lower predictive performance. Conversely, larger ROIs (7×7) included a broader anatomical context and potentially richer radiomic information, but substantially reduced the number of available training samples and showed poorer generalization on the independent test set, suggesting increased susceptibility to overfitting. Overall, 5×5 ROIs provided the best compromise between local tissue representation, sample availability, and predictive performance, and were therefore adopted in all subsequent experiments.

Table 6. Preliminary evaluation of different ROI sizes.

ROI Size	AUC	Accuracy	Precision	Recall	F1-Score
3×3	0.689	0.645	0.667	0.718	0.680
5×5	0.743	0.698	0.701	0.856	0.771
7×7	0.701	0.664	0.690	0.747	0.713

To assess the effect of ROI sampling, the proposed balanced spatial sampling strategy was compared with two alternatives. The first was a grid-based strategy, in which non-overlapping ROIs were extracted within the preoperative tumor region according to a regular spatial grid. The second was a random tumor-only strategy, in which non-overlapping ROIs were randomly sampled from the tumor region alone. In contrast, the proposed approach adopted the ROI sampling and labeling strategy described in Section 2.3.1, which was specifically designed to reduce class imbalance while preserving the spatial relationship between the preoperative lesion and the observed recurrence.

The results are reported in Table 7. The ablation study was conducted using the same ROI-level evaluation protocol adopted throughout model development to ensure a fair comparison among the alternative sampling strategies. The proposed balanced spatial sampling strategy achieved the best overall performance, with an AUC of 0.743, an accuracy of 0.698, a precision of 0.701, a recall of 0.856, and an F1-score of 0.771. Performance was markedly lower with both alternative strategies. The grid-based approach reached an AUC of 0.548 and an F1-score of 0.501, whereas the random tumor-only strategy yielded the weakest performance overall.

Table 7. Final ROI-level performance on the independent test set for different ROI sampling strategies.

ROI Sampling Strategy	AUC	Accuracy	Precision	Recall	F1-Score
Grid-based sampling	0.548	0.502	0.625	0.419	0.501
Random tumor-only sampling	0.491	0.299	0.733	0.195	0.308
Balanced spatial sampling (proposed)	0.743	0.698	0.701	0.856	0.771

Using the proposed balanced spatial sampling strategy, we then compared CatBoost with Random Forest and XGBoost under the same preprocessing, feature extraction, and training procedure (Table 8). CatBoost achieved the highest AUC (0.743) and accuracy (0.698), followed closely by XGBoost (AUC = 0.739; accuracy = 0.694), while Random Forest achieved the highest recall (0.882) despite a slightly lower AUC (0.725). Overall, the three classifiers showed broadly comparable performance, suggesting that the choice of classifier had a relatively limited impact under the adopted experimental setting. In contrast, the ROI sampling strategy had a substantially greater influence on predictive performance, as demonstrated by the marked differences observed in Table 7.

Table 8. ROI-level performance on the independent test set for different machine learning classifiers, using the proposed balanced spatial sampling strategy.

Model	AUC	Accuracy	Precision	Recall	F1-Score
Random Forest	0.725	0.694	0.689	0.882	0.774
XGBoost	0.739	0.694	0.701	0.845	0.766
CatBoost (proposed)	0.743	0.698	0.701	0.856	0.771

Finally, we evaluated the contribution of each MRI modality by training the proposed CatBoost classifier using radiomic features extracted from T1ce alone, FLAIR alone, and the combination of both modalities (Table 9). The same preprocessing, ROI sampling strategy, and training protocol were adopted in all cases to ensure a fair comparison. The combined T1ce + FLAIR model achieved the best overall performance, improving the AUC from 0.720 (FLAIR alone) and 0.704 (T1ce alone) to 0.743. It also provided the highest recall (0.856), suggesting that the two modalities provide complementary information for recurrence prediction.

Table 9. Final ROI-level performance on the independent test set for different MRI input modalities.

Input Modality	AUC	Accuracy	Precision	Recall	F1-Score
T1ce	0.704	0.663	0.681	0.766	0.721
FLAIR	0.720	0.676	0.692	0.802	0.743
T1ce + FLAIR	0.743	0.698	0.701	0.856	0.771

3.4. Preoperative Recurrence Risk Maps

To assess the spatial behavior of the proposed framework, the trained classifier was applied to the preoperative MRI volumes to generate patient-specific recurrence risk maps. For each case in the independent test set, dense sliding-window inference was performed on the preoperative T1ce and FLAIR images within the annotated tumor region and the surrounding peritumoral area. For each window location, radiomic features were extracted separately from the two modalities and processed by the classifier to obtain two recurrence probability estimates, which were then averaged to assign a single risk value to that spatial location.

The resulting probabilities were projected back onto the image grid and locally averaged across overlapping windows, yielding slice-wise recurrence risk maps on preoperative MRI. These maps were visually compared with the recurrence regions identified on follow-up imaging after spatial co-registration. In representative cases, high-probability areas were mainly observed within the tumor and adjacent peritumoral tissue and showed qualitative spatial correspondence with the regions that later developed recurrence.

Figure 4 shows a representative example of this spatial reconstruction. Figure 4a presents the slice-wise comparison between preoperative MRI, the predicted recurrence risk map, and the ground-truth recurrence region. Figure 4b shows the corresponding three-dimensional visualization, including the recurrence region identified on follow-up imaging, the predicted recurrence region derived from the risk map, the corresponding binary prediction mask, and their spatial overlap. These visual results support the ability of the proposed framework to capture spatial patterns associated with future tumor regrowth from preoperative imaging, although the present evaluation remains qualitative.

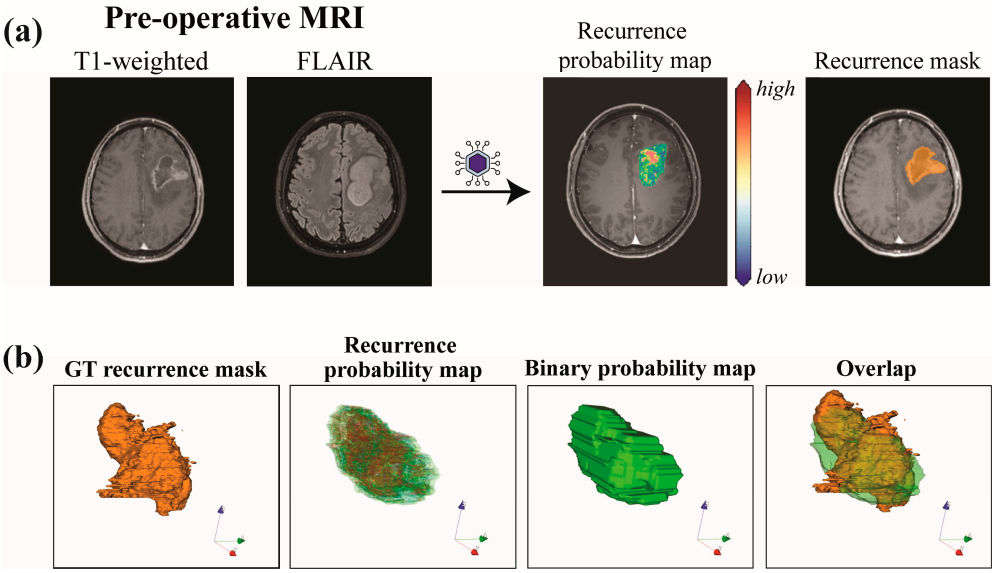


Figure 4. Representative best-performing case of preoperative recurrence risk visualization in 2D and 3D. (**a**) Slice-wise comparison on preoperative MRI showing the annotated tumor region, the predicted recurrence risk map, and the ground-truth recurrence region identified on follow-up MRI after co-registration. (**b**) Three-dimensional rendering showing the recurrence region identified on follow-up imaging, the predicted recurrence region derived from the risk map, the corresponding binary prediction mask, and their spatial overlap. In the overlap view, predicted recurrence is shown in green and observed recurrence in orange.

On the workstation used in this study (NVIDIA RTX A6000 GPU), the complete inference pipeline required approximately 11 min per patient (range: 6–18 min), with processing time mainly depending on the tumor volume and the corresponding number of sampled ROIs.

3.5. Analysis of Radiomic Features Relevance

At the individual feature level (Figure 5a), the ranking based on mean absolute SHAP values identified the radiomic descriptors with the greatest influence on model predictions. Among the top 10 features, GLCM descriptors were the most represented, accounting for four of the ten highest-ranking variables, followed by four first-order features with slightly lower importance values. The most influential features included GLCM Correlation, GLRLM Run Length Non-Uniformity, and GLCM Inverse Difference Normalized (IDN), together with first-order features such as Minimum Intensity, Energy, 10th Percentile

Intensity, and Total Energy. Additional relevant contributors included GLCM Cluster Shade, GLCM Joint Average, and GLDM Dependence Non-Uniformity. Overall, these findings indicate that both local texture organization and intensity-based image characteristics play an important role in recurrence prediction.

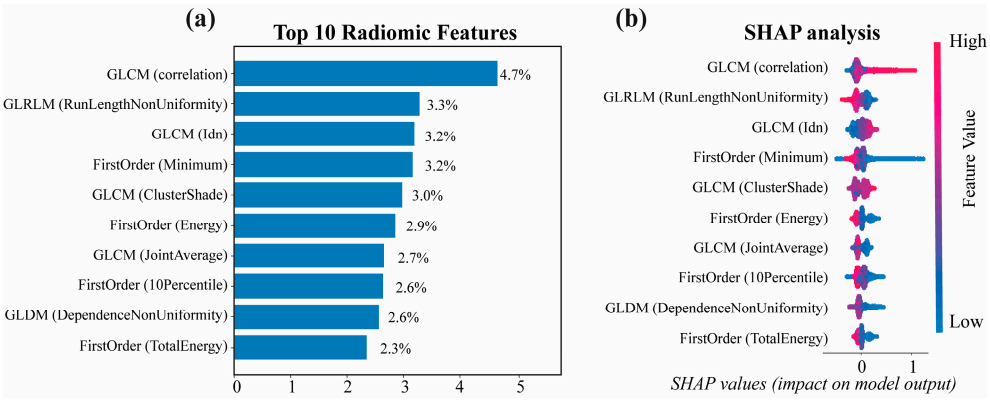


Figure 5. SHAP-based explainability analysis of the radiomic model. (**a**) Ranking of the top 10 radiomic features based on mean absolute SHAP values. (**b**) SHAP summary plot illustrating the magnitude and direction of feature effects across ROIs in the test set.

At the feature-family level, GLCM and first-order features provided the largest overall contribution to the model, accounting for 30.6% and 28.8% of the total importance, respectively. GLRLM (13.3%), GLDM (12.8%), and GLSZM (10.6%) showed more moderate contributions, whereas NGTDM features had only a limited impact (4.0%). The SHAP summary plot (Figure 5b) further illustrates the direction and magnitude of feature effects across ROIs. Positive SHAP values indicate a contribution toward recurrence prediction, whereas negative values indicate a contribution toward non-recurrence. Higher values of GLCM Correlation were generally associated with recurrence. In contrast, lower values of First-order Minimum Intensity were more often linked to recurrence, while lower values of First-order Maximum Intensity tended to be associated with non-recurrence. GLRLM Run Length Non-Uniformity showed a similar directional pattern, with lower values contributing more frequently to recurrence and higher values to non-recurrence.

4. Discussion

A central challenge in glioblastoma surgery is that preoperative MRI does not fully reflect the true extent of infiltrative disease [1]. While contrast-enhanced T1-weighted imaging delineates the enhancing core, tumor infiltration often extends beyond these margins, and FLAIR abnormalities remain difficult to interpret because they may reflect both edema and infiltrative tumor [33]. This complicates the identification of peritumoral regions at higher risk of recurrence during surgical planning. In this context, the proposed framework was developed to extract spatially resolved information on recurrence risk from routine preoperative MRI and translate it into maps potentially relevant for preoperative decision-making.

4.1. Main Findings and Interpretation

This study presents a radiomics-based framework for preoperative spatial prediction of glioblastoma recurrence using standard MRI. By combining balanced spatial ROI sampling, radiomic feature extraction, and gradient boosting classification, the approach generates localized recurrence risk maps directly from preoperative imaging. On the independent test set, the model achieved an AUC of 0.743, an F1-score of 0.771, and a recall of 0.856, indicating good sensitivity in identifying regions associated with future recurrence (Table 5).

The complementary patient-wise evaluation yielded comparable mean performance but also revealed a relatively large inter-patient variability (AUC = 0.748 ± 0.118), indicating that predictive performance was not uniform across individual cases. Given the marked biological and radiological heterogeneity of glioblastoma, this variability is not unexpected and highlights that the proposed framework currently provides more reliable predictions for some patients than for others.

The ablation study showed that both ROI sampling strategy and model selection materially affected predictive performance. In particular, the proposed balanced spatial sampling approach outperformed both grid-based and random tumor-only sampling. While the tumor-only strategy randomly samples ROIs across the preoperative tumor region, it does not explicitly ensure adequate representation of tissue that will eventually recur. As a result, recurrence-associated imaging patterns are sparsely sampled relative to non-recurrent regions, yielding a weakly informative training set and limiting the model's ability to learn discriminative radiomic signatures. In contrast, the proposed balanced spatial sampling strategy explicitly includes ROIs from both recurrence and non-recurrence regions, providing a more representative characterization of local tissue heterogeneity. These findings suggest that recurrence prediction depends not only on the discriminative value of the extracted radiomic features, but also on how spatial information is sampled and represented during training. The modality ablation study further showed that combining T1ce and FLAIR provided higher predictive performance than either modality alone. This finding is consistent with the complementary roles of the two MRI sequences, with T1ce primarily characterizing the contrast-enhancing tumor core and FLAIR providing additional information from the surrounding peritumoral tissue, where infiltrative tumor cells are frequently present.

The qualitative analysis of the preoperative recurrence risk maps further supports the spatial relevance of the proposed framework (Figure 4). In representative test cases, high-probability regions on the preoperative maps showed spatial correspondence with recurrence areas identified on follow-up MRI. Although this observation remains qualitative, it suggests that radiomic patterns derived from standard preoperative imaging may retain local information associated with subsequent tumor regrowth.

The explainability analysis adds another layer of interpretation. The most influential features were mainly drawn from GLCM and first-order families, indicating that the model relies on both texture-related descriptors and intensity-based statistics. This pattern suggests that recurrence prediction is informed not only by signal intensity distribution, but also by the spatial organization of local image patterns. Among individual features, GLCM Correlation and GLRLM Run Length Non-Uniformity were among the strongest contributors, supporting the importance of local textural organization and tissue heterogeneity in the model's predictions.

The explainability analysis also helps interpret the imaging patterns learned by the proposed model. The predominance of GLCM and first-order features suggests that recurrence prediction depends on both local texture organization and signal intensity, rather than on a single imaging characteristic. This is consistent with the heterogeneous nature of glioblastoma infiltration, where subtle local differences in tissue architecture may already be present before recurrence becomes visible on conventional MRI.

Among the individual features, GLCM Correlation and GLRLM Run Length Non-Uniformity indicate that the spatial organization of neighboring pixels contributes substantially to the model predictions, suggesting that local texture carries information beyond simple intensity differences. Similarly, the association between lower First-order Minimum Intensity values and recurrence indicates that small low-intensity components within a local region may also contribute to recurrence prediction. Although these intensity varia-

tions cannot be directly linked to specific pathological substrates, they may reflect subtle differences in tissue composition that are not appreciable by visual inspection alone. These findings suggest that the proposed model identifies localized radiomic patterns associated with future recurrence by combining information on tissue heterogeneity and signal intensity. Understanding the pathological basis of these radiomic signatures will require future studies integrating radiomic, histopathological, and molecular data. It should also be noted that these radiomic descriptors were computed from small local ROIs (5×5 pixels), rather than from conventional whole-tumor regions. Accordingly, their biological interpretation may differ from that of texture features extracted from larger tumor volumes, despite the superior predictive performance observed for this ROI size in the ablation study.

From a modeling perspective, the choice of a conventional machine-learning approach was deliberate. Given the limited cohort size and the importance of maintaining interpretability, CatBoost offered a reasonable balance between predictive performance, robustness, and explainability. In this setting, the use of an explainable tabular model also made it possible to characterize the radiomic signature through SHAP in a way that would have been less straightforward with a fully deep learning–based framework.

4.2. Clinical Implications for Surgical Planning

From a clinical perspective, the ability to generate preoperative spatial recurrence maps may be highly relevant for surgical planning in glioblastoma. Current surgical decision-making still relies primarily on contrast-enhanced T1-weighted MRI, which delineates areas of blood–brain barrier disruption but does not fully capture the true biological extent of the disease [1,2]. In this setting, our framework aims to identify peritumoral regions with a higher likelihood of future recurrence, potentially providing spatial information that may complement conventional imaging during preoperative planning. More broadly, a preoperative map of localized recurrence risk could support surgical planning through spatial risk stratification, helping surgeons identify peritumoral regions that may deserve closer consideration during resection. In particular, such maps may facilitate a more biologically informed approach to supramaximal resection by highlighting FLAIR abnormalities that are more likely to harbor infiltrative tumor while avoiding unnecessary extension into regions predominantly reflecting vasogenic edema.

The high recall observed in our model is particularly relevant in this context. In spatial recurrence prediction, false negatives correspond to regions at risk that are assigned a low probability and may therefore be overlooked. By favoring sensitivity, the proposed approach reduces the likelihood of underestimating areas prone to recurrence, which is consistent with the clinical objective of maximizing tumor removal while preserving neurological function.

These considerations are particularly relevant in light of the growing interest in supramaximal resection [34], defined as surgical removal extending beyond the contrast-enhancing tumor margins, which has emerged as a promising strategy for improving local tumor control and survival. However, FLAIR hyperintensity represents a heterogeneous imaging substrate that encompasses both infiltrative tumor and vasogenic edema. Consequently, FLAIR-guided supramaximal resection may expose patients to unnecessary functional risk, particularly when lesions are located near eloquent brain regions. Incorporating radiomic recurrence-risk maps into intraoperative neuronavigation may therefore enable a more targeted, biologically informed, and function-preserving supramaximal resection strategy that is more specific than conventional FLAIR-based approaches. By identifying spatially localized regions at increased risk of recurrence, such maps could potentially guide surgical extension toward biologically relevant tissue while limiting unnecessary resection of areas less likely to harbor infiltrative disease.

Ultimately, this radiomic framework may contribute to optimizing the extent of tumor resection and, consequently, has the potential to improve progression-free and overall survival. Nevertheless, these potential applications remain exploratory and require prospective validation, as well as integration with functional brain mapping and other anatomical information, and careful assessment of safety, particularly for tumors involving eloquent brain regions, before clinical implementation can be considered.

4.3. Positioning with Respect to Previous Studies

The present findings should be interpreted in the context of previous radiomics and machine-learning studies on glioblastoma recurrence. Earlier MRI-based radiomics approaches have shown that quantitative imaging features can provide prognostic information for outcomes such as survival, progression, and molecular characterization [11,12]. However, most of these studies generate patient-level predictions and therefore do not provide the spatially resolved information that would be needed for surgical planning. In contrast, the aim of our framework is to generate local recurrence risk maps, moving from global prognostic stratification toward spatially explicit preoperative assessment.

Recent studies by Cepeda et al. [16,17] reported comparable performance in spatial recurrence modeling, with AUC values around 0.81. However, those approaches were based on postoperative MRI and evaluated on substantially larger cohorts, including a retrospective cohort of 114 patients and an additional prospective validation cohort of 25 patients. In contrast, the proposed framework focuses on preoperative, spatially resolved recurrence-risk estimation, addressing a different stage of the clinical workflow in which surgical planning decisions are made. Moreover, the preoperative setting requires complete preoperative imaging, longitudinal follow-up, and accurate spatial correspondence between preoperative and recurrence annotations, resulting in a more demanding patient selection process and a consequently smaller eligible cohort. Therefore, direct comparison of the reported performance metrics should be interpreted in light of these methodological and cohort-related differences.

Another relevant aspect is that the proposed method relies only on standard MRI sequences routinely acquired in clinical practice, namely T1ce and FLAIR. This differs from preoperative studies that incorporate more specialized imaging modalities, including diffusion-based acquisitions, to better characterize infiltrative tissue [20]. Although such approaches may improve predictive performance, they are less consistently available across centers and may therefore be harder to integrate into routine workflows. In this respect, our results suggest that spatial prediction of recurrence may also be feasible using conventional imaging, provided that the sampling strategy and feature modeling are carefully designed.

Compared with earlier preoperative spatial approaches, our framework also seeks to provide a more detailed spatial representation of recurrence risk. Some previous methods relied on relatively coarse spatial representations [18] or addressed recurrence in ways that were less directly oriented toward surgical support [21]. By contrast, the present approach generates dense preoperative probability maps with the aim of capturing localized risk patterns that may be relevant during resection planning.

4.4. Limitations and Future Perspectives

Before discussing the general limitations of the proposed framework, it is useful to examine some representative challenging cases from the independent test set (Figure 6). These cases were selected from patients with comparatively lower patient-wise prediction performance and illustrate the inter-patient variability reflected by the standard deviation reported in Table 5. They highlight situations in which the current model provides less accurate spatial predictions and therefore help identify directions for future methodological improvements.

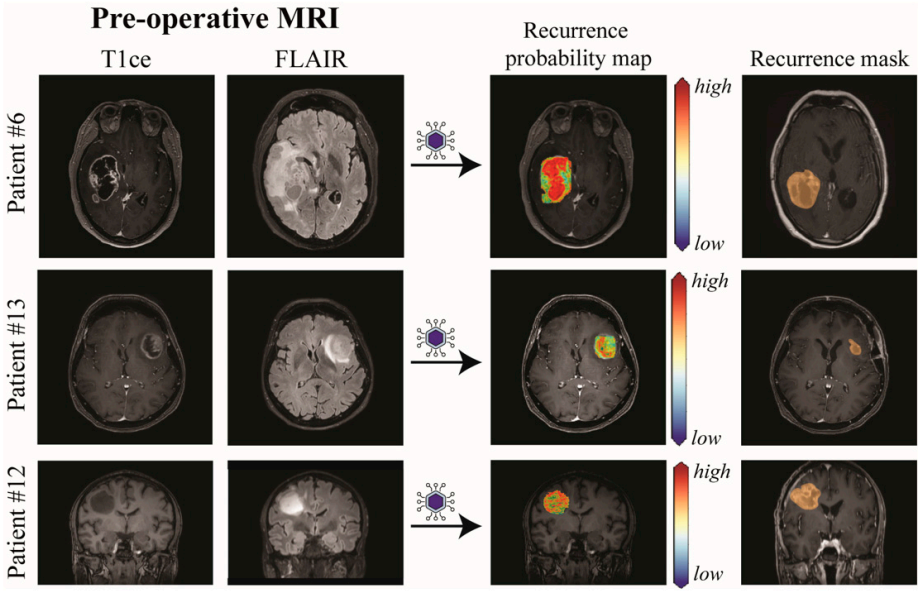


Figure 6. Representative challenging cases from the independent test set. For each patient, the figure shows the preoperative MRI, the predicted recurrence-risk map, and the corresponding recurrence region identified on follow-up MRI after co-registration.

In the first case, the predicted recurrence-risk map correctly identifies the main recurrence region but also assigns high probability to an adjacent area that does not subsequently recur, reflecting reduced spatial specificity. In the second case, the recurrence is relatively small and, although the highest probabilities are correctly concentrated around the recurrent lesion, the predicted map remains more spatially diffuse. Finally, in the third case, the model successfully identifies the overall recurrence region but produces a heterogeneous probability distribution within the lesion, with alternating areas of higher and lower predicted risk. These examples illustrate the current challenges of spatial recurrence prediction from routine preoperative MRI and motivate future developments toward more accurate and spatially consistent risk estimation.

Some limitations should be acknowledged. First, the study was conducted on a relatively small single-center cohort, reflecting the challenges of collecting homogeneous datasets including histopathologically confirmed glioblastoma, standardized preoperative MRI, and longitudinal follow-up. Although this design reduced technical variability, it may limit the generalizability of the proposed framework across institutions, scanners, and clinical protocols [2]. Accordingly, larger prospective multicenter studies will be essential to validate the robustness and clinical applicability of the proposed approach. Tumor delineation in the preoperative setting was performed manually by a single experienced neurosurgeon, while co-registration between preoperative and follow-up MRI relied on a semi-manual workflow. Although this approach ensured methodological consistency across the entire cohort, it does not allow assessment of inter-observer variability and may introduce operator-dependent variability, as well as a limited degree of label noise. Future multicenter studies incorporating multi-rater annotations and dedicated inter-observer variability analyses will be needed to further assess the robustness and reproducibility of the proposed framework.

Second, molecular biomarkers, including IDH mutation status and MGMT promoter methylation, were not incorporated into the present framework. Although these factors are known to influence glioblastoma biology and recurrence patterns, they were beyond the scope of the present study and may therefore represent potential confounding factors. Future studies should investigate the integration of radiomic features with molecular and histopathological biomarkers to develop multimodal predictive frameworks that

may improve recurrence-risk estimation, enable patient-specific stratification, and provide further insight into the biological mechanisms underlying glioblastoma recurrence.

Third, although recurrence risk maps were reconstructed in three dimensions, their evaluation was mainly qualitative. The current comparison involves a continuous recurrence-risk probability map and a binary recurrence mask, making quantitative spatial validation inherently challenging. More generally, the present spatial modeling strategy is still constrained by local patch-based sampling and 2D slice-wise feature extraction, so the full three-dimensional continuity of infiltrative spread may not be completely captured. Future work should therefore focus on developing dedicated quantitative validation strategies for probabilistic spatial risk maps, together with wider regions of analysis and fully three-dimensional modeling approaches.

Fourth, the pipeline is computationally demanding because of the stride-1 sampling strategy and the subsequent 3D reconstruction step. These choices were made to preserve fine spatial detail, but they reduce efficiency and may limit broader clinical applicability. In addition, patch-level predictions were aggregated at inference by simple averaging, which is straightforward and reproducible but may not fully reflect the heterogeneous contribution of different spatial regions. Future work should investigate more computationally efficient inference strategies, including end-to-end deep learning approaches capable of directly generating recurrence-risk maps without explicit patch-wise classification and reconstruction, together with alternative aggregation strategies.

Future work could also investigate label-efficient learning strategies to reduce the manual annotation burden associated with generating tumor and recurrence segmentations. Recent advances in deep learning [35,36], self-supervised learning, and foundation models [37] have shown considerable potential for medical image analysis and automated image interpretation. To reduce the burden of manual annotations, semi-supervised segmentation approaches [38] could facilitate the delineation of preoperative tumors and postoperative recurrence volumes in 3D MRI, while expert refinement and validation would ensure the reliability of the resulting annotations [36]. Such strategies could substantially improve scalability and facilitate the construction of larger multicenter datasets for future model development.

Finally, future versions of the framework should incorporate both patient-specific explainability and uncertainty quantification [39,40]. While the present SHAP analysis provides a global interpretation of the trained model, extending explainability to the individual-patient level may provide additional insight into the radiomic patterns driving each prediction. In parallel, providing not only recurrence probabilities but also an estimate of prediction reliability could improve interpretability and help distinguish more reliable high-risk regions from less certain ones.

5. Conclusions

We proposed a radiomics-based framework for generating preoperative spatial risk maps of glioblastoma recurrence from routine MRI. The results show that radiomic patterns extracted from T1ce and FLAIR images can provide spatially resolved information on future recurrence, enabling localized risk estimation before surgery. These findings provide preliminary evidence supporting the feasibility of the proposed approach for surgical planning, particularly in identifying peritumoral regions that may extend beyond conventional imaging boundaries. Further multicenter and quantitative validation will be necessary, but this framework represents a promising step toward preoperative, image-guided precision management of glioblastoma.

Author Contributions: Conceptualization, A.M. and F.C.; methodology, S.S. and F.L.P.; software, F.L.P. and M.S. (Matteo Salvi); validation, S.S. and F.L.P.; formal analysis, F.L.P.; investigation, S.S., A.M. and M.S. (Massimo Salvi); resources, A.M., G.B., D.G. and F.C.; data curation, F.L.P. and G.B.; writing—original draft preparation, S.S., M.S. (Massimo Salvi) and A.M.; writing—review and editing, F.L.P., M.S. (Matteo Salvi), G.B., D.G. and F.C.; visualization, S.S. and M.S. (Matteo Salvi); supervision, M.S. (Massimo Salvi), D.G. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the local Institutional Review Board of A.O.U. Città della Salute e della Scienza di Torino (protocol code 00162/2022, approval date: 15 September 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available from the corresponding author upon reasonable request.

Acknowledgments: The authors used Claude Sonnet 4.5 to improve the readability and language quality of the manuscript. The authors carefully reviewed and edited the output as needed and take full responsibility for the content of the publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADC	Apparent diffusion coefficient
AUC	Area under the receiver operating characteristic curve
CE	Contrast-enhancing
DTI	Diffusion tensor imaging
FLAIR	Fluid-attenuated inversion recovery
GBM	Glioblastoma
GLCM	Gray-Level Co-occurrence Matrix
GLDM	Gray-Level Dependence Matrix
GLRLM	Gray-Level Run Length Matrix
GLSZM	Gray-Level Size Zone Matrix
IBSI	Imaging Biomarker Standardization Initiative
IDN	Inverse Difference Normalized
ML	Machine learning
MRI	Magnetic resonance imaging
NGTDM	Neighboring Gray Tone Difference Matrix
non-CE	Non-contrast-enhancing
OS	Overall survival
RANO	Response Assessment in Neuro-Oncology
RF	Random Forest
ROI	Region of interest
SHAP	SHapley Additive exPlanations
SVM	Support vector machine
T1ce	T1-weighted contrast-enhanced
WHO	World Health Organization
XGBoost	eXtreme Gradient Boosting

References

1. Tan, A.C.; Ashley, D.M.; López, G.Y.; Malinzak, M.; Friedman, H.S.; Khasraw, M. Management of Glioblastoma: State of the Art and Future Directions. *CA Cancer J. Clin.* **2020**, *70*, 299–312. [CrossRef] [PubMed]

2. McKinnon, C.; Nandhabalan, M.; Murray, S.A.; Plaha, P. Glioblastoma: Clinical Presentation, Diagnosis, and Management. *BMJ* **2021**, *374*, n1560. [CrossRef] [PubMed]
3. Scherer, H.J. The forms of growth in gliomas and their practical significance. *Brain* **1940**, *63*, 1–35. [CrossRef]
4. Chamberlain, M.C. Radiographic Patterns of Relapse in Glioblastoma. *J. Neurooncol.* **2011**, *101*, 319–323. [CrossRef] [PubMed]
5. Cofano, F.; Morello, A.; Rizzo, F.; De Marco, R.; Garbossa, D. Management of Recurrent Gliomas: An Update of Strategies and Treatment Options. In *Neurosurgery's Frontline Role in Gliomas Treatment: Trends, Technologies and Treatments*; Springer: Cham, Switzerland, 2025; pp. 255–266. [CrossRef]
6. Milano, M.T.; Okunieff, P.; Donatello, R.S.; Mohile, N.A.; Sul, J.; Walter, K.A.; Korones, D.N. Patterns and Timing of Recurrence after Temozolomide-Based Chemoradiation for Glioblastoma. *Int. J. Radiat. Oncol. Biol. Phys.* **2010**, *78*, 1147–1155. [CrossRef] [PubMed]
7. Obrador, E.; Moreno-Murciano, P.; Oriol-Caballo, M.; López-Blanch, R.; Pineda, B.; Gutiérrez-Arroyo, J.L.; Loras, A.; Gonzalez-Bonet, L.G.; Martinez-Cadenas, C.; Estrela, J.M.; et al. Glioblastoma Therapy: Past, Present and Future. *Int. J. Mol. Sci.* **2024**, *25*, 2529. [CrossRef] [PubMed]
8. D'Alessio, A.; Proietti, G.; Sica, G.; Scicchitano, B.M. Pathological and Molecular Features of Glioblastoma and Its Peritumoral Tissue. *Cancers* **2019**, *11*, 469. [CrossRef] [PubMed]
9. Kickingereder, P.; Burth, S.; Wick, A.; Götz, M.; Eidel, O.; Schlemmer, H.P.; Maier-Hein, K.H.; Wick, W.; Bendszus, M.; Radbruch, A.; et al. Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. *Radiology* **2016**, *280*, 880–889. [CrossRef] [PubMed]
10. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; De Jong, E.E.C.; Van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The Bridge between Medical Imaging and Personalized Medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [CrossRef] [PubMed]
11. Ko, C.C.; Zhang, Y.; Chen, J.H.; Chang, K.T.; Chen, T.Y.; Lim, S.W.; Wu, T.C.; Su, M.Y. Pre-Operative MRI Radiomics for the Prediction of Progression and Recurrence in Meningiomas. *Front. Neurol.* **2021**, *12*, 636235. [CrossRef] [PubMed]
12. Gao, M.; Cheng, J.; Qiu, A.; Zhao, D.; Wang, J.; Liu, J. Magnetic Resonance Imaging (MRI)-Based Intratumoral and Peritumoral Radiomics for Prognosis Prediction in Glioma Patients. *Clin. Radiol.* **2024**, *79*, e1383–e1393. [CrossRef] [PubMed]
13. Pease, M.; Gersey, Z.C.; Ak, M.; Elakkad, A.; Kotrotsou, A.; Zenkin, S.; Elshafeey, N.; Mamindla, P.; Kumar, V.A.; Kumar, A.J.; et al. Pre-Operative MRI Radiomics Model Non-Invasively Predicts Key Genomic Markers and Survival in Glioblastoma Patients. *J. Neurooncol.* **2022**, *160*, 253–263. [CrossRef] [PubMed]
14. Yin, S.; Ming, J.; Chen, H.; Sun, Y.; Jiang, C. Integrating Deep Learning and Radiomics for Preoperative Glioma Grading Using Multi-Center MRI Data. *Sci. Rep.* **2025**, *15*, 36756. [CrossRef] [PubMed]
15. Li, L.; Xiao, F.; Wang, S.; Kuang, S.; Li, Z.; Zhong, Y.; Xu, D.; Cai, Y.; Li, S.; Chen, J.; et al. Preoperative Prediction of MGMT Promoter Methylation in Glioblastoma Based on Multiregional and Multi-Sequence MRI Radiomics Analysis. *Sci. Rep.* **2024**, *14*, 16031. [CrossRef] [PubMed]
16. Cepeda, S.; Luppino, L.T.; Pérez-Núñez, A.; Solheim, O.; García-García, S.; Velasco-Casares, M.; Karlberg, A.; Eikenes, L.; Sarabia, R.; Arrese, I.; et al. Predicting Regions of Local Recurrence in Glioblastomas Using Voxel-Based Radiomic Features of Multiparametric Postoperative MRI. *Cancers* **2023**, *15*, 1894. [CrossRef] [PubMed]
17. Cepeda, S.; Esteban-Sinovas, O.; Luppino, L.T.; Kuttner, S.; Wodzinski, M.; Solheim, O.; Romero, R.; Pérez-Núñez, A.; Eikenes, L.; Karlberg, A.; et al. Radiomics-Based Quantification of Tumor Infiltration in the Non-Enhancing Peritumoral Region on Postoperative MRI Is Associated with Survival in Glioblastoma. *Sci. Rep.* **2025**, *15*, 43932. [CrossRef] [PubMed]
18. Li, Y.; Jiang, Z.; Tan, J.; Wang, Z.; Ma, Y.; Xiong, D.; Liu, Y.; Li, K.; Lui, S.; Wu, M. An Intelligent Postoperative Management System for Glioblastoma Integrating Automated Segmentation, Risk Stratification, and Recurrence Spatial Mapping. *Radiother. Oncol.* **2025**, *213*, 111180. [CrossRef] [PubMed]
19. Champ, C.E.; Siglin, J.; Mishra, M.V.; Shen, X.; Werner-Wasik, M.; Andrews, D.W.; Mayekar, S.U.; Liu, H.; Shi, W. Evaluating Changes in Radiation Treatment Volumes from Post-Operative to Same-Day Planning MRI in High-Grade Gliomas. *Radiat. Oncol.* **2012**, *7*, 220. [CrossRef] [PubMed]
20. Metz, M.C.; Molina-Romero, M.; Lipkova, J.; Gempt, J.; Liesche-Starnecker, F.; Eichinger, P.; Grundl, L.; Menze, B.; Combs, S.E.; Zimmer, C.; et al. Predicting Glioblastoma Recurrence from Preoperative MR Scans Using Fractional-Anisotropy Maps with Free-Water Suppression. *Cancers* **2020**, *12*, 728. [CrossRef] [PubMed]
21. Chougule, T.; Gupta, R.K.; Saini, J.; Agrawal, S.; Gupta, M.; Vakharia, N.; Singh, A.; Patir, R.; Vaishya, S.; Ingalhalikar, M. Radiomics Signature for Temporal Evolution and Recurrence Patterns of Glioblastoma Using Multimodal Magnetic Resonance Imaging. *NMR Biomed.* **2022**, *35*, e4647. [CrossRef] [PubMed]
22. Li, Q.; Xiang, C.; Zeng, X.; Liao, A.; Chen, K.; Yang, J.; Li, Y.; Jia, M.; Song, L.; Hu, X. Multimodal Radiomics in Glioma: Predicting Recurrence in the Peritumoural Brain Zone Using Integrated MRI. *BMC Med. Imaging* **2025**, *25*, 321. [CrossRef] [PubMed]
23. Luzzi, S.; Agosti, A. Radiomics Multifactorial in Silico Model for Spatial Prediction of Glioblastoma Progression and Recurrence: A Proof-of-Concept. *World Neurosurg.* **2024**, *183*, e677–e686. [CrossRef] [PubMed]

24. Louis, D.N.; Perry, A.; Wesseling, P.; Brat, D.J.; Cree, I.A.; Figarella-Branger, D.; Hawkins, C.; Ng, H.K.; Pfister, S.M.; Reifenberger, G.; et al. The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary. *Neuro. Oncol.* **2021**, *23*, 1231–1251. [CrossRef] [PubMed]
25. Vogelbaum, M.A.; Jost, S.; Aghi, M.K.; Heimberger, A.B.; Sampson, J.H.; Wen, P.Y.; MacDonald, D.R.; Van Den Bent, M.J.; Chang, S.M. Application of Novel Response/Progression Measures for Surgically Delivered Therapies for Gliomas: Response Assessment in Neuro-Oncology (RANO) Working Group. *Neurosurgery* **2012**, *70*, 234–244. [CrossRef] [PubMed]
26. Radiomic Features—Pyradiomics v3.1.0rc2.Post5+g6a761c4 Documentation. Available online: https://pyradiomics.readthedocs.io/en/latest/features.html (accessed on 20 March 2026).
27. Löfstedt, T.; Brynolfsson, P.; Asklund, T.; Nyholm, T.; Garpebring, A. Gray-Level Invariant Haralick Texture Features. *PLoS ONE* **2019**, *14*, e0212110. [CrossRef] [PubMed]
28. Haralick, R.M.; Dinstein, I.; Shanmugam, K. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
29. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. Image Biomarker Standardization Initiative. *Radiology* **2019**, *295*, 2. [CrossRef] [PubMed]
30. Chakraborty, P. CatBoost-Stacked Heterogeneous Deep Ensembles for Explainable Multi-Class Brain Tumor MRI Classification. *J. Med. Health Stud.* **2026**, *7*, 18–24. [CrossRef]
31. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased Boosting with Categorical Features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 3–8 December, 2018; pp. 6638–6648.
32. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, New York, NY, USA, 4–9 December 2017; pp. 4766–4775.
33. Merisaari, H.; Jambor, I.; Ettala, O.; Boström, P.J.; Montoya Perez, I.; Verho, J.; Kiviniemi, A.; Syvänen, K.; Kähkönen, E.; Eklund, L.; et al. IMPROD Biparametric MRI in Men with a Clinical Suspicion of Prostate Cancer (IMPROD Trial): Sensitivity for Prostate Cancer Detection in Correlation with Whole-Mount Prostatectomy Sections and Implications for Focal Therapy. *J. Magn. Reson. Imaging* **2019**, *50*, 1641–1650. [CrossRef] [PubMed]
34. Certo, F.; Pluchino, A.; Maugeri, A.; Ferranti, G.; Broggi, G.; Caltabiano, R.; Melcarne, A.; Rudà, R.; Della Pepa, G.M.; La Rocca, G.; et al. Is FLAIRectomy Directly Correlated with Prolonged Survival in Glioblastoma? A Prospective National Multicenter Study on Correlation Between Extent of Tumor Resection and Clinical Outcome. *Neurosurgery* **2025**, *97*, 489–500. [CrossRef] [PubMed]
35. Shahali, S.; Mortimer, D.; O'Bryan, M.K.; McLachlan, R.; Zander-Fox, D.; Ackermann, K.; Ahmad, G.; Neild, A.; Nosrati, R. An Artificial Intelligence-Based Computer Vision Model for Human Sperm Concentration, Motility, and Kinematics Analysis. *Smart Med.* **2026**, *5*, e70026. [CrossRef] [PubMed]
36. Wu, W.; Yang, X.; Yao, C.; Liu, O.; Zhao, Q.; Shuai, J. E-SegNet: E-Shaped Structure Networks for Accurate 2D and 3D Medical Image Segmentation. *Research* **2025**, *8*, 0869. [CrossRef] [PubMed]
37. Jiao, L.; Hao, J.; Li, R.; Li, L.; Liu, X.; Liu, F.; Ma, W.; Chen, P.; Huang, Z.; Yang, J.; et al. Foundation Models Meet Medical Image Interpretation. *Research* **2026**, *9*, 1024. [CrossRef] [PubMed]
38. Nitti, F.; Seoni, S.; Morello, A.; Dolci, L.; Piazza, A.; Esposito, V.; Rosito, L.; Garbossa, D.; Cofano, F.; Sengur, A.; et al. Uncertainty-Aware Semi-Supervised Learning for Neurosurgical Navigation. *Appl. Soft Comput.* **2026**, *197*, 115252. [CrossRef]
39. Seoni, S.; Jahmunah, V.; Salvi, M.; Barua, D.; Molinari, F.; Rajendra Acharya, U.; Seoni, S.; Jahmunah, V.; Salvi, M.; Barua, P.D.; et al. Application of Uncertainty Quantification to Artificial Intelligence in Healthcare: A Review of Last Decade (2013–2023). *Comput. Biol. Med.* **2023**, *165*, 107441. [CrossRef] [PubMed]
40. Salvi, M.; Seoni, S.; Campagner, A.; Gertych, A.; Acharya, U.R.; Molinari, F.; Cabitza, F. Explainability and Uncertainty: Two Sides of the Same Coin for Enhancing the Interpretability of Deep Learning Models in Healthcare. *Int. J. Med. Inform.* **2025**, *197*, 105846. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.