



Enhanced Tumor Classification via Radiomics and Deep Feature Generation using Neural Network

Integrating Radiomics with CNN-Based Deep Features

Dilek Betül SARIDEDE

Biomedical Engineering Department,
Faculty of Engineering and Natural
Sciences, İstanbul Atlas University,
İstanbul, Türkiye
dilek.saridede@atlas.edu.tr

Rukiye Nur KAÇMAZ

Software Engineering Department
Faculty of Engineering, Erciyes
University, Kayseri, Türkiye
rukiyekacmaz@erciyes.edu.tr

Sevim CENGİZ

College of Technological Innovation,
Zayed University, Dubai, United
Arab Emirates
sevim.Cengiz@zu.ac.ae

Abstract—Glioblastoma (GBM) is the most aggressive and prevalent malignant brain tumor in adults. Accurate preoperative differentiation between GBM and other gliomas remains a critical challenge due to imaging overlaps and tumor heterogeneity. Radiomics has emerged as a powerful tool to extract quantitative imaging features; however, most existing approaches rely on traditional radiomic features derived from limited modalities or tumor regions. In this study, we have proposed a novel hybrid pipeline that integrates traditional radiomic features with deep features extracted using a convolutional neural network (CNN). We used multimodal magnetic resonance images and four pre-segmented tumor subregions (necrosis, edema, enhancing tumor, and whole tumor) from 185 brain tumor patients publicly available in The Cancer Imaging Archive (141 GBM, 44 non-GBM). A total of 1544 radiomic features were extracted using the PyRadiomics library for each patient. After applying FeatureWiz for feature selection, a Support Vector Machine classifier was trained using only selected traditional radiomic features and both selected traditional radiomic features and CNN-based deep features. Our results demonstrate that combining traditional radiomic and deep features significantly improves classification performance. Validation accuracy increased from 0.838 to 0.948, F1-score from 0.558 to 0.948, and recall from 0.436 to 0.950. These improvements indicate better generalization and robustness of the hybrid model. This study highlights the potential of integrating deep and traditional radiomic features for glioblastoma classification. Our hybrid pipeline addresses limitations of existing radiomics-only approaches and provides a promising direction for non-invasive tumor characterization in neuro-oncology.

Keywords: Radiomic Features; Machine Learning; Medical Imaging; Feature Engineering

I. INTRODUCTION

Gliomas are the most common primary tumors of the central nervous system, arising from glial cells in the brain or spinal cord. Among them, glioblastoma (GBM) is the most aggressive and prevalent malignant subtype in adults, classified as a World Health Organization (WHO) grade IV glioma [1]. Despite significant advances in neuro-oncology, accurate preoperative differentiation between glioblastoma and other gliomas such as lower-grade gliomas or astrocytomas remains a clinical challenge due to overlapping imaging features and tumor heterogeneity [2].

Radiomics has emerged as a promising approach in computational oncology, enabling the extraction of high-dimensional quantitative features from medical imaging data to detect subtle patterns of tumor phenotype and heterogeneity [3]. It has shown considerable potential in differentiating glioblastoma from lower-grade gliomas and in survival prediction tasks. Nonetheless, many prior studies have relied exclusively on traditional features extracted from either a single MRI modality or non-specific tumor masks [4–6]. Akbari et al. [5] addressed this limitation by incorporating both intratumoral and peritumoral regions in feature extraction, demonstrating that spatial context can enhance predictive performance. However, their study remained limited to classical radiomics-based approaches. More recently, Wan et al. [4] evaluated survival prediction in recurrent glioblastoma patients by comparing traditional radiomic signatures with an end-to-end CNN model, but without integrating traditional and deep features. Similarly, Tan et al. [2] explored the impact of different volume-of-interest fusion strategies for glioma grading, yet their approach was also restricted to traditional features. Hassan et al. [7] proposed a multi-modal integration strategy about traditional radiomic features and deep features generated via a CNN; however, their study focused on dementia subtyping using multi-omics data and structural neuroimaging. It did not address glioblastoma classification. Similarly, Kobayashi et al. [8] and Zhou et al. [9] investigated the integration of deep learning and radiomics for glioblastoma classification, but their CNN components were primarily designed either as full classification models or to enhance interpretability rather than as dedicated feature extractors. Therefore, advancing glioblastoma-specific classification models that combine traditional and learned features from multimodal, multi-region MRI inputs remains an important research direction.

To address this limitation, we propose a novel hybrid pipeline that integrates traditional radiomic features extracted from four magnetic resonance imaging (MRI) modalities and four tumor subregions (necrosis, edema, enhancing tumor and whole tumor) with deep features generated via a CNN. By combining both data-driven and domain-specific features, our

approach aims to improve the diagnostic performance in distinguishing glioblastoma from non-glioblastoma tumors.

II. MATERIALS AND METHODS

A. Subjects and MRI Acquisition

Clinical information and MR images of 185 brain tumor patients, comprising 141 confirmed glioblastoma and 44 non-glioblastoma tumors were obtained from The Cancer Imaging Archive [https://www.cancerimagingarchive.net/collection/mu-glioma-post/] [10].

Clinical information used in this study included age and sex information of the patients (Table 1). T1-weighted, T1-weighted post-contrast, T2-weighted and Fluid-attenuated inversion recovery (FLAIR) MR images were used. The dataset provides manually pre-segmented tumor subregions, including necrosis, edema, and enhancing tumor regions.

Table 1. Demographic information of the patients.

Patient Group	Age	Gender (F:M)
Patient with GBM tumors	63.19±12.14	55:68
Patient with non-glioblastoma tumors	41.27±15.95	23:21

B. Radiomic Feature Extraction

Radiomic features were extracted using a custom Python script utilizing the PyRadiomics 2.2.0 library [https://github.com/Radiomics/pyradiomics]. Feature categories included: 3D shape features (n = 17), first-order statistics (n = 19), and higher-order texture features (n = 75). Due to high collinearity among certain shape and first-order features, a reduced set of features was used in the analysis. Specifically, compactness, compactness 2, and spherical disproportion were removed from the shape group, and standard deviation was removed from the first-order group due to its strong correlation with variance. Ultimately, 14 shape features and 18 first-order statistics features were retained. Shape features, being contrast-invariant, were extracted once per segmented compartment per patient. All other features were extracted across 4 MRI modalities (T1, T1-contrast, T2, and FLAIR) and 4 tumor subregions (necrosis, edema, enhancing tumor and whole tumor). Thus, the total feature number of each patient equals 1544. Fig. 1 shows that multimodal MR images with manually segmented tumor subregions (edema, enhancing tumor, necrosis/non-enhancing tumor) were used to extract shape, first-order, and texture features. The extracted features were combined into a structured tabular format, where each row corresponds to a patient and each column to a specific radiomic feature. This table served as the input for the classification.

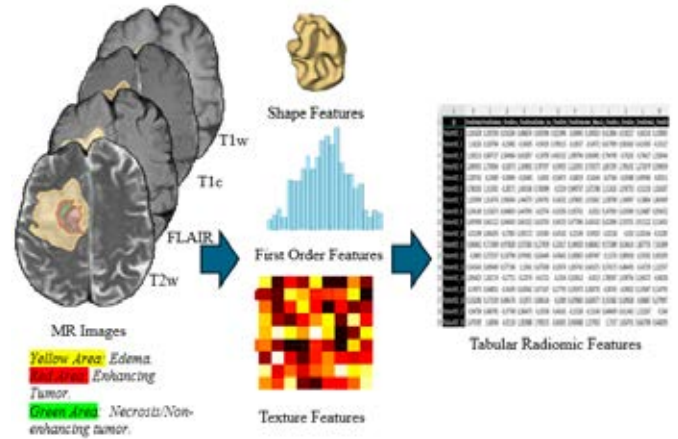


Fig. 1. The pipeline of the Radiomic Feature Extraction.

C. Classification

Fig. 2 shows the overview of the classification pipeline. Selected traditional radiomic features were directly used as input for machine learning algorithms (a). In the proposed model (b), the same features were passed through a CNN consisting of convolution, pooling, and flattening layers to extract deep features. These CNN-based deep features were then concatenated with the original radiomic features to form a combined feature set. After applying feature selection, the refined feature sets were used to train machine learning model for tumor classification.

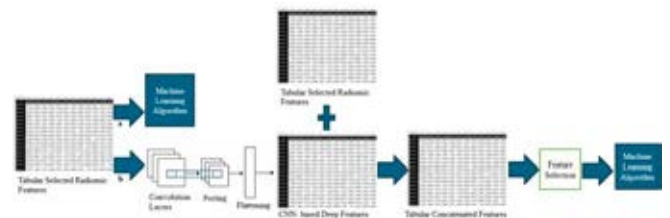


Fig. 2. The pipeline for the classification.

Firstly, the extracted features were preprocessed by replacing missing values using median imputation within each class. Featurewiz was applied for automated feature selection (correlation threshold = 0.85). Class imbalance was addressed with Synthetic Minority Over-sampling Technique (SMOTE) oversampling in the training dataset. In small datasets, particularly common in medical research, applying SMOTE effectively improves model performance by mitigating class imbalance [11]. Support Vector Machine (SVM) Learning Model was employed. SVM is frequently highlighted as a preferred classifier in SMOTE-augmented, small-sample applications due to their robustness and generalization capabilities [12]. Performance metrics included train accuracy, validation accuracy, F1-score, Receiver-operating characteristic curve (ROC) Area under the curve (AUC), precision, recall (sensitivity) and specificity were evaluated. In

the proposed model, to enhance the feature space, Deep Feature Generation (DFG) was proposed. DFG employed CNN as a feature extractor. After the application of Featurewizz, the selected traditional radiomic features were scaled and reshaped to fit a CNN which was constructed with two convolutional layers followed by global average pooling. A dense (fully connected) layer with 64 units was used to represent deep features: Input $\rightarrow$ Conv1D(32) $\rightarrow$ Dropout $\rightarrow$ Conv1D(16) $\rightarrow$ GlobalAvgPooling $\rightarrow$ Dense(64). These CNN-based deep features were then concatenated with the traditional radiomic features. The combined feature set was reprocessed using featurewiz and the same ML classifier was re-applied. Then, performance metrics were evaluated. As with the classical pipeline, oversampling (SMOTE), normalization, and inclusion of demographic variables (sex) were carefully handled during training.

III. RESULTS

There was a clear increase in performance metrics including validation accuracy, F1 score, ROC AUC, precision, and recall when concatenated features (Selected Traditional Radiomic Features + CNN-based Deep Features) were used, compared with using either Selected Traditional Radiomic Features or CNN-based Deep Features alone (Table 2). Specifically, when Concatenated Features (Selected Traditional Radiomic Features + CNN-based Deep Features) were compared with Selected Traditional Radiomic Features alone, F1 score improved from 0.558 to 0.948, recall from 0.436 to 0.950, and ROC AUC from 0.860 to 0.981. Although train accuracy rose only slightly (from 0.992 to 0.998), validation accuracy showed a more meaningful increase (from 0.838 to 0.948). On the other hand, using CNN-based Deep Features alone resulted in much lower performance (validation accuracy 0.503, F1 0.505, ROC AUC 0.379).

Table 2. Performance Comparison of SVM Classifier Using Selected Traditional Features vs. Concatenated Features (Selected Traditional Features + CNN-based Deep Features).

Performance Metrics	Selected Traditional Features	CNN-based Deep Features	Concatenated Features (Selected Traditional Features + CNN-based Deep Features)
Train Accuracy	0.992	0.527	0.998
Validation Accuracy	0.838	0.503	0.948
F1 Score	0.558	0.505	0.948
ROC AUC	0.860	0.379	0.981
Precision	0.797	0.584	0.948
Recall (Sensitivity)	0.436	0.364	0.950
Specificity	0.965	0.646	0.946

Fig. 3 shows that the learning curve shows the differences in training and validation accuracy across all 5 folds when using the concatenated feature set. In each fold, both training and validation accuracy remain consistently high.



Fig. 3. Differences in accuracy of training and testing applied by SVM with Concatenated Features (Selected Features + DFG).

IV. DISCUSSION

In this study, we propose a hybrid classification pipeline that integrates traditional radiomic features with deep features generated via a CNN to distinguish glioblastoma from non-glioblastoma brain tumors. Our results demonstrate that combining traditional radiomic features with data-driven deep features significantly enhances classification performance across multiple evaluation metrics.

Our results demonstrate that combining radiomic features with deep generation features significantly enhances classification performance across multiple evaluation metrics. While training accuracy remained consistently high (0.992 vs. 0.998), the validation accuracy improved markedly from 0.838 to 0.948 when using the concatenated feature set, suggesting enhanced generalizability and reduced overfitting. This finding aligns with prior work highlighting the complementary value of integrating deep learning and traditional radiomics for tumor classification. This finding aligns with prior work highlighting the complementary value of integrating deep learning and radiomics for tumor classification [8,9]. However, unlike these studies, where CNNs were primarily used either as the main classification model or for interpretability purposes, our approach employs a CNN as a dedicated feature extractor operating on multi-region, multi-modality traditional radiomic features. The resulting deep features are systematically fused with pre-selected traditional radiomic features, followed by an additional feature selection step to enhance generalizability and reduce overfitting. The result of this study shows that the relatively small gap between training and validation curves, which suggests that the model generalizes well to unseen data. This stability indicates that integrating CNN-based deep features with the selected features contributes positively to model robustness and predictive performance. Furthermore, our findings resonate with the conceptual basis of radiomics defined by Lambin et al. [3], who underlined the value of quantitative imaging biomarkers, and with Akbari et al. [5], who highlighted the importance of spatial context in radiomic feature design. Similarly, Tan et al. [2] and Wan et al. [4] showed that combining multi-region or multi-modal features

improves grading and survival prediction tasks, though their studies lacked deep learning integration. By incorporating a CNN-based deep feature generator into the radiomic pipeline, our model bridges this gap and complements previous traditional-only studies like Chaddad et al. [6].

The proposed model used FeatureWiz-selected radiomic features (about 300–400) instead of raw 3D MRI images. Because CNN operated only on this small, pre-filtered table derived from the traditional radiomic features, the computational cost remained very low: it required only a few hundred thousand parameters, trained within minutes on a standard GPU, and needed far less memory than image-based CNNs. The CNN-based deep features were simply new hidden representations learned from these selected radiomics, not a second round of image processing. Our results showed that using only Selected Traditional Features yielded about 84% validation accuracy, using CNN-based deep features alone yielded about 62%, while combining both raised validation accuracy to about 95%. Similar findings in other study have reported that direct 3D CNNs often perform worse on small medical datasets [7]. Therefore, radiomic features were enriched with deep learning rather than extracting deep features directly from images.

To the best of our knowledge, dataset used in this study was most recently updated on March 21, 2025 and has not yet been cited or analyzed in any published studies. Therefore, no directly comparable classification results are currently available. [10]

Despite these promising results, there are several limitations including small dataset size and the use of neural network architecture which may restrict generalizability and highlight the need for further validation across diverse cohorts.

V. CONCLUSION

This study presents a novel classification framework that combines traditional radiomic features and CNN-based deep features, achieving notable improvements in diagnostic performance for distinguishing glioblastoma from non-glioblastoma tumors. The results support our hypothesis that integrating domain-driven and data-driven features enhances the representational power of the model and improves generalization, particularly in small, imbalanced medical datasets. Future work may explore external validation cohorts and the extension of this framework to other tumor types or prognostic tasks, further advancing precision diagnostics in neuro-oncology.

REFERENCES

- [1] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 World Health Organization classification of tumors of the central nervous system: A summary," *Acta*

Neuropathologica, vol. 131, no. 6, pp. 803–820, 2016, doi: 10.1007/s00401-016-1545-1.

- [2] R. Tan, C. Sui, C. Wang, and T. Zhu, "MRI-based intratumoral and peritumoral radiomics for preoperative prediction of glioma grade: A multicenter study," *Frontiers in Oncology*, vol. 14, p. 1401977, 2024, doi: 10.3389/fonc.2024.1401977.
- [3] P. Lambin et al., "Radiomics: Extracting more information from medical images using advanced feature analysis," *European Journal of Cancer*, vol. 48, no. 4, pp. 441–446, 2012, doi: 10.1016/j.ejca.2011.11.036.
- [4] Q. Wan et al., "Comparative analysis of deep learning and radiomic signatures for overall survival prediction in recurrent high-grade glioma treated with immunotherapy," *Cancer Imaging*, vol. 25, no. 5, 2025, doi: 10.1186/s40644-024-00818-0.
- [5] H. Akbari et al., "Imaging surrogates of infiltration obtained via multiparametric imaging pattern analysis predict subsequent location of recurrence of glioblastoma," *Neurosurgery*, vol. 78, no. 4, pp. 572–580, 2016, doi: 10.1227/NEU.0000000000001102.
- [6] A. Chaddad, C. Desrosiers, and M. Toews, "Radiomic analysis of multi-contrast brain MRI for the prediction of survival in patients with glioblastoma," *IEEE Access*, vol. 6, pp. 73031–73041, 2018, doi: 10.1109/ACCESS.2018.2880542.
- [7] S. Hassan et al., "MINDSETS: Multi-omics integration with neuroimaging for dementia subtyping and effective temporal study," *Scientific Reports*, vol. 15, no. 15835, 2025, doi: 10.1038/s41598-025-97674-0.
- [8] Z. Hu et al., "A deep learning model with radiomics analysis integration for glioblastoma post-resection survival prediction," *arXiv preprint, arXiv:2203.05891*, 2022.
- [9] Z. A. Shboul, M. Alam, L. Vidyaratne, L. Pei, M. I. Elbakary, and K. M. Iftekharuddin, "Feature-Guided Deep Radiomics for Glioblastoma Patient Survival Prediction," *Frontiers in Neuroscience*, vol. 13, p. 966, 2019, doi: 10.3389/fnins.2019.00966.
- [10] D. Yaseen et al., "University of Missouri post-operative glioma dataset (MU-Glioma-Post) (Version 1) [Data set]," *The Cancer Imaging Archive*, 2025, doi: 10.7937/7K9K-3C83.
- [11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
- [12] Q. Wang, Z. Luo, J. Huang, Y. Feng, and Z. Liu, "A novel ensemble method for imbalanced data learning: Bagging of extrapolation-SMOTE SVM," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 1827016, 2017, doi: 10.1155/2017/1827016.