Radiogenomic Classification of Brain Tumor using MRI Sequences

Ganesh Sapkota*, Md. Hasibur Rahman*, Mizanur Rahman Jewel*, and Md. Sazedur Rahman *
Missouri University of Science and Technology
Rolla, MO, USA
{gsapkota, mrpk9, mj9vc, mrvfw}@mst.edu

Abstract

MRI sequences of brain tumor analysis are crucial for treatment planning and patient prognosis. In most cases, brain tumor analysis focuses on segmentation to identify the affected area of the brain, facilitating precise surgery. However, post-surgery, inspecting the genomic profile of the tumor through biopsy is necessary for further assessment and treatment, a process that is both time-consuming and costly. In this study, we tackle the issue of additional surgery by directly predicting the presence of the genomic property, known as O^6 -methylguanine DNA methyltransferase (MGMT) promoter methylation, by feeding MRI image sequences into our proposed model. Our methodology encompasses the segmentation and classification of radiology imaging of brain tumors. By applying advanced machine learning techniques, our framework extracts imaging features from Multiparametric mpMRI scans: T1w, T1wCE, T2w, and FLAIR to enhance classification accuracy. By integrating DenseNet-169 and SAM models, we aim to boost the performance of MGMT status prediction. Validated on the RSNA-MICCAI dataset, our approach achieves a maximum AUC score of 67.54% and an accuracy of 70%. These results underscore the potential of radiogenomics for non-invasive, personalized brain tumor management, offering valuable insights for clinical decision-making. The code is available here.

1. Introduction

Brain cancer is characterized by aberrant cell proliferation inside the brain. Because it can cause serious brain damage and even death in extreme circumstances, it presents a serious hazard to human health. The complexity of brain cancer contributes to its severity, making it difficult to diagnose and treat. According to a recent GLOBOCAN survey conducted in 185 countries, there were 251,329 fatalities and 308,102 new cases of brain cancer in 185 nations in 2020 alone [1]. With relatively few viable treatment options, glioblastoma multiforme (GBM) is one of the most aggressive and recalcitrant forms of malignant brain cancer. For patients with GBM, surgery to remove the tumor is usually advised, followed by chemotherapy and radiation therapy. Generally, a patient having GBM undergoes a severe risky surgery to extract some chemical sample from the tumor. After conducting the chemical test, the doctor decides on further surgery or chemotherapy decisions. But there are two major challenges here: i) Any kind of brain surgery is risky and having surgery multiple times makes it more deadly for a patient. ii) The chemical result comes after a handful amount of time within which the life of a patient may be more severe. So, It is crucial to develop an automated system that allows doctors to make non-surgical decisions and quickly obtain results through advanced imaging technologies Fortunately, the detection of O^6-methylguanine DNA methyltransferase (MGMT) promoter methylation in tumors is a positive indicator of prognosis and a reliable predictor of chemotherapy response. With the introduction of a radiomic imaging method, diagnosing brain tumors will become less invasive, significantly improving outcomes and patient survival rates for brain cancer.

Contributions. In response to the challenges and limitations outlined above, this paper proposes a new framework to find the MGMT status of a brain tumor with better performance. In summary, this paper makes the following contributions:

- The Segment Anything Model (SAM) has been applied to produce masks for the given mpMRI sequences and to extract radiomic features [2] aimed at predicting the status of the MGMT biomarker.
- Two preprocessing methods have been utilized, and all the models are evaluated on each method.
- Extensive experimentation has been conducted on the outputs from the DenseNet169 and SAM models by combining them into an ensemble.

This paper is organized into five Sections. Section 1 introduces the research problem, and motivation then

*All authors have contributed equally. Names are listed in alphabetical order.

highlights our contributions to solving the problem. Section 2 discusses Related Works. Section 3 describes our Proposed Method and explains the system overview in detail. Section 4 discusses the experimental setup, dataset, evaluation metrics, performance results, and comparison of different approaches. Section 5 concludes the paper with the application and future direction of research.

2. Related Works

The Radiological Society of North America held a competitive Kaggle competition in the summer of 2021 to support new research on MGMT promoter methylation detection through the application of data science and machine learning methods. More than 1,500 submissions were made to the challenge by researchers worldwide. Sveinn Palsson et al. in [3] have proposed deep neural networks to partition the tumor from the available MR images extract radiomic features and form features that are trained by a variational auto-encoder. To get predictions, they used a typical machine learning process that involved selecting features and then training a random forest classification model. used the RSNA-ASNR-MICCAI BraTS 2021 challenge dataset [4] is used to train and assess their approach. Suman das in [5] introduces a novel ML model called the Intermediate State Generator (IS-Gen) to normalize MRI scan slice thicknesses, leading to improved model performance and a 6% increase in average cross-validation accuracy compared to the best Kaggle model. The goal of the research is to increase the precision of MGMT promoter methylation prediction utilizing noisy and representative medical data by training and evaluating four alternative models. For object detection, Abdusalomov et al in [6] have proposed a deep learning method utilizing YOLOv7 that has already been trained. Furthermore. They have employed the Channel and Spatial Attention Module (CBAM) for enhanced attention mechanisms and the Bi-directional Feature Pyramid Network (BiFPN) for feature extraction. 99.5% accuracy was achieved but it depends on specific MRI scanners. Using FLAIR, T1w, T2, and T1Gd mpMRIs, S. A. Qureshi et al. [7] have developed a unique two-stage MGMT Promoter Methylation Prediction (MGMT-PMP) approach to assess the picture structure of glioblastoma (GB) in patients. To improve classification outcomes, it presents a Deep Learning Radiomic Feature Extraction (DLRFE) module that extracts dynamic features based on the problem structure. Recently, several deep learning architectures have been investigated by Amr Mohamed et al. [8]. Among them were Xception, ResNet50, and 3D Vision Transformer (ViT3D), the latter of which demonstrated encouraging outcomes in correctly predicting MGMT promoter status from mpMRI data. Predictions from various models trained on each type of mpMRI scan have been combined using ensemble approaches, such as averaging or stacking ensembles with logistic regression, to improve the overall predictive performance. The performance of models such as EfficientNet-B3 is moderate, suggesting diversity in model effectiveness across different architectures, whereas deep learning architectures like Xception has shown promising results.

3. Proposed Method

In this section, we provide an overview of our proposed approach, illustrated in Fig.1. It is designed to predict the MGMT status from a sequence of MRI images. Below, we provide a detailed description of how the framework operates:

1. At the bottom of Fig.1, the MRI image is passed to a pre-processing step where DICOM format is converted to NIfTI format which is well organized and easily manageable.
2. The converted NIfTI images are passed to the Segment Anything Model (SAM) [9] to generate radiomic features.
3. These radiomic features are passed to a machine learning model, Support Vector Machine(SVM) for classification.
4. On the top of Fig.1, the same MRI images are passed to another pre-processing block where 3D volumetric inputs are created using regular interval and central image-based sequence selection approaches.
5. Both the output images from the central and regular interval of dimension $256 * 256 * 64$ are passed to the Densenet169 model to extract image features.
6. Extracted image features are passed to a fully connected neural network for classification.
7. Lastly, the Bagging ensemble method is applied to combine both classification results from **Phase 3** and **Phase 6** to generate better accuracy.

Each phase is executed by a separate process, which will be discussed in detail in the following sections.

3.1. Data Pre-processing

Medical image is challenging because of its non-uniform nature and heterogeneous source. Thus, several data mining techniques have been applied to pre-process the MRI images.

3.1.1 DICOM to NIfTI Conversion

The conversion of DICOM (Digital Imaging and Communications in Medicine) format to NIfTI (Neuroimaging Informatics Technology Initiative) format is a critical step in preparing medical imaging data for advanced analysis and research, providing standardization, compatibility, and enhanced data handling that are essential for effective and accurate scientific outcomes. The conversion process is illus-

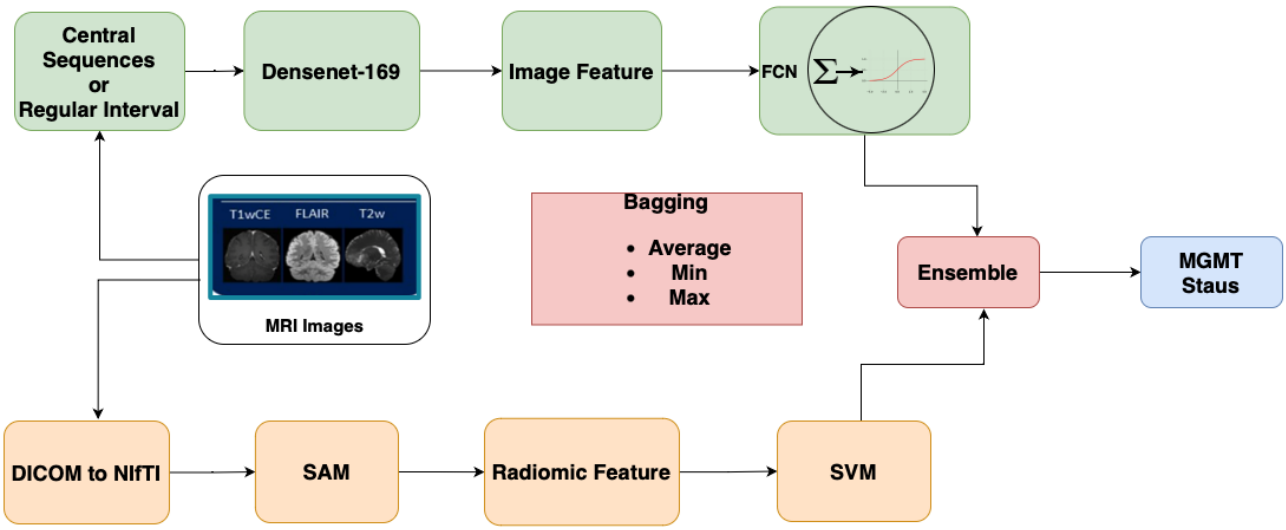

Figure 1. Proposed Method.

trated in Fig.2. The original MRI images are in DICOM format with several samples from individual patients. Each sample is comprised of different sequences. Thus the original DICOM format is converted to NIfTI format. It is to be noted that, only the flair sequences are considered because of resource constraints. Each flair sequence is processed by affine transformation to generate uniform sequences. The converted input image ends up with the dimension of 256 * 256 * 64 for image feature extraction which represents the stack of 64 images with a length and width of 256px. While the input image dimension for radiomic feature extraction becomes 512*512*512.

Every individual case presents a varying number of MRI

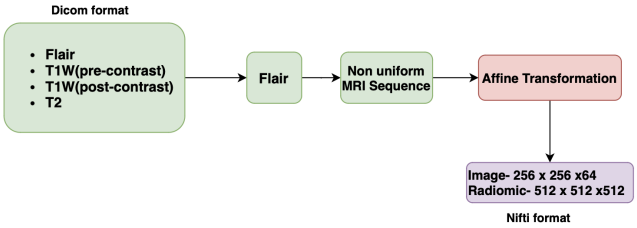

Figure 2. DICOM to NIfTI.

scan images. Utilizing all scans can lead to confusion for the model in understanding the spatial relationships between brain pixels. So we applied two different approaches to select some scan sequences of the brain to obtain a 3D volumetric stack as an input.

3.1.2 Sequence Selection at Regular Interval

One solution is to select the sequence at equal intervals and stack them to obtain the 3D volumetric image as input as shown in Algorithm 1. For instance, 64 sequences can be obtained at an equal interval of INT(100/64) for 100 Fluid Attenuated inversion recovery (FLAIR) sequences. However, stacking sequences from equal intervals doesn't ensure consistent brain portion representation due to the unequal number of sequences in each case. Also, skipping the sequences that represent most of the parts of the brain may result in losing important features. So, the Model struggles to learn crucial tumor patterns and may learn irrelevant spatial patterns.

Algorithm 1 Create 3D Stack using Regular Interval

Input: $data\_directory$, $split$, $scan\_id$, mri_type, num_imgs
Output: $img3d$ (3D image stack)
$files \leftarrow$ natural_sort$(path)$
$every\_nth \leftarrow$ length$(files)/num_imgs$
$indexes \leftarrow [\min((i \times every\_nth), \text{length}(files) - 1))$ for i in range $(0, num\_imgs)]$
$files\_selected \leftarrow [files[index]$ for $index$ in $indexes]$
$img3d \leftarrow$ stack([load_image(f) for f in $files\_selected]$)

3.1.3 Sequence Selection based on Central Image

We observed improved results in training and validation when we considered the largest visible brain cutaway in the scans as the central image and created the 3D volumetric stack based on this central image as shown in Algorithm 2.

Consider Scan 1 with 100 Fluid Attenuated inversion recovery (FLAIR) scans and Scan 2 with 400 FLAIR scans, where the model inputs 64 images. Our solution selects the image with the largest visible brain section (image 50 for Scan 1 and image 150 for Scan 2) and constructs 3D images accordingly. For Scan 1, this spans from image 18 to 82, while for Scan 2, it extends from image 168 to 232. This approach ensures consistent brain portion representation and thus captures tumor features from the same portion of the brain for every scan case.

Algorithm 2 Create 3D Volumetric input based on Central Sequence

Input: $files$ (array of file paths), num_imgs (desired number of images)
Output: $image\_stack$ (3D volumetric stack)
$middle \leftarrow \text{length}(files)//2$
$num\_imgs2 \leftarrow num\_imgs//2$
$start \leftarrow \max(0, middle - num\_imgs2)$
$end \leftarrow \min(\text{length}(files), middle + num\_imgs2)$
$image\_stack \leftarrow$ empty list
for f in $files[start : end]$ **do**
    $image \leftarrow \text{load\_image}(f)$
    Append $image$ to $image\_stack$
end for
return $image\_stack$

3.1.4 Train Test Split

There are a total of 585 train samples, 3 of them are corrupted and discarded. The total dataset is split into 80% train set and 20% test set. Then stratification is used to handle the class imbalance problem.

3.2. Mask Generation

Fig. 3 illustrates the process of mask generation using the segment anything model (SAM). At first, from 3D images, masks are generated by slicing the sequences. Then converted 2D images are passed to the image encoder portion of SAM that generates the mask of each 2D image. And finally, to generate a 3D NIfTI mask, all the 2D masks are concatenated.

3.3. Radiomic Feature

Among different radiomic features, we have considered extracting shape-based and intensity-based features such as volume and energy. Extraction of those two types of radiomic features of each 3D NIfTI mask is shown in Fig.4. First, we calculate binary 3D NIfTI mask {0, 1} from the 3D nifit mask {0, 255} shown in section 3.2. Then we calculate the volume and energy of the tumor object in the mask.

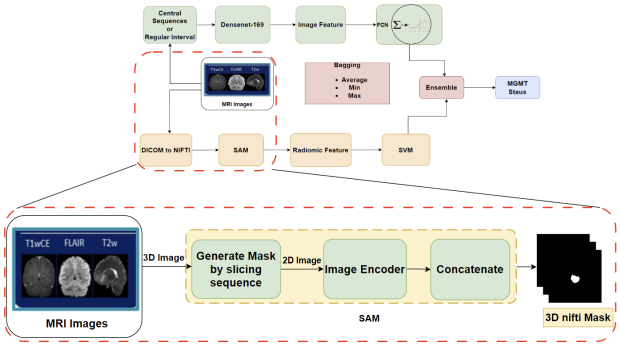

Figure 3. Mask Generation

$$\text{volume}, V = \sum_{i \in H} \sum_{j \in W} \sum_{k \in D} M_{i,j,k} \tag{1}$$

where H, W, and D is the height, weight, and depth of the 3D NIfTI binary mask M, and each $M_{i,j,k}$ has value $\{0, 1\}$.

$$\text{Intensity}, I = A \odot M \tag{2}$$

where A is the Original 3D nifit image and BM is 3D NIfTI binary mask.

$$\text{Energy}, E = \sum_{i \in H} \sum_{j \in W} \sum_{k \in D} I^2_{i,j,k} \tag{3}$$

An SVM (support vector machine) model is trained on the extracted radiomic features with a linear kernel.

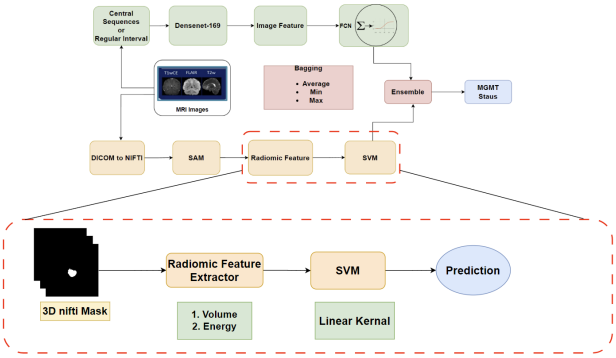

Figure 4. Radiomic Feature.

3.4. Image Feature Generation using Densenet169

On the top side of the proposed method shown in Fig.5, Densenet169 is used to generate image features from the MRI images. Densenet architecture was originally proposed by Huang. et. al. in [10]. We have utilized Densenet169 for our task mainly of two reasons: i)Features can be reused in this model. ii) No vanishing gradient issues. Besides that, focal loss shown in equation 4 is used

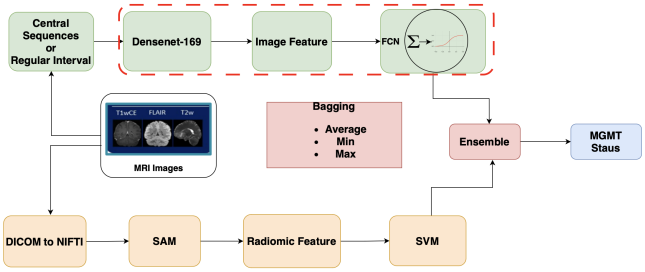

Figure 5. Image Feature using Densenet169.

instead of cross-entropy loss. The focal loss places more emphasis on hard-to-classify examples.

$$\text{FL}(p_t) = -(1-p_t)^{\gamma} \cdot \log(p_t) \tag{4}$$

By incorporating the modulating factor $(1-p_t)^{\gamma}$, where p_t is the predicted probability of the correct class, it effectively increases the loss for misclassified examples (p_t close to 0) while reducing the loss for well-classified examples (p_t close to 1). That's why focal loss was chosen. The parameter settings for training the Densenet169 model are :

- γ=2
- Image Size = 256
- Num_Img= 64
- Batch_Size = 1
- Learning Rate = 1e-6
- Lr_Decay = 0.9
- Num_Epochs = 150

3.5. Ensemble using Bagging

After generating predictions from both Densenet169 and the SAM-SVM model, the bagging ensemble technique is employed to get the best performance. Three types of bagging methods are employed in the proposed method:

I. Average Bagging: The prediction probabilities from the Densenet169 applied in FLAIR, T2W, provide us with two predictions. We ensembled them by averaging the probabilities. The ensembling between the SAM and densenet169 is also performed.

II. Min Bagging: The final prediction is obtained by taking the minimum value of the prediction probabilities made by different models.

III. Max Bagging: The final prediction is obtained by taking the maximum value of the prediction probabilities made by different models.

4. Experiments

4.1. Experimental Setup

Experiments are done in Almalinux v8.9 setup with 8 CPU cores, NVIDIA A100 80GB GPU, CUDA version 12.4, and PyTorch version 2.2.0.

4.2. Dataset

We utilize the RSNA-MICCAI Brain Tumor Radiomic Classification dataset [4] to evaluate our approach. Multiparametric MRI scans, including T1-weighted (T1w), T1-weighted with contrast enhancement (T1wCE), T2-weighted (T2w), and Fluid Attenuated Inversion Recovery (FLAIR) sequences, are commonly included in the dataset. Accurate tumor segmentation depends on the complementary information these modalities provide about the brain and its abnormalities. There are a total of 672 patient samples among which 585 samples are used for training and 87 samples are used for testing. For our experiment, we have only used the training sample and split that into a train and test set as we were unable to get the original test sample labels.

4.3. Evaluation Metrics

We have employed several standard evaluation metrics to justify the performance of the proposed method. These are: **ROC**: The Receiver Operating Characteristic (ROC) curve is a graphical plot created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The true positive rate (TPR), also known as sensitivity, is defined as:

$$TPR = \frac{TP}{TP+FN} \tag{5}$$

The false positive rate (FPR), also known as the fall-out, is defined as:

$$FPR = \frac{FP}{FP+TN} \tag{6}$$

Accuracy: Accuracy represents the proportion of correctly classified instances out of the total number of instances evaluated.The accuracy of a classification model is calculated using the following formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

AUC Score: The Area Under the ROC Curve (AUC) represents the probability that the model will rank a randomly chosen positive instance higher than a randomly chosen negative instance. Given the ROC curve with m points (FPR_i, TPR_i), where $i = 1, 2, ..., m$, the AUC score can be computed as:

$$\text{AUC} = \sum_{i=1}^{m-1} \frac{(FPR_{i+1} - FPR_i) \cdot (TPR_i + TPR_{i+1})}{2} \tag{8}$$

The AUC score ranges from 0 to 1.

Where:

- TP is the number of true positive predictions.
- TN is the number of true negative predictions.

- FP is the number of false positive predictions.
- FN is the number of false negative predictions.

4.4. Densenet169 (Regular Interval Approach)

In regular interval image sequencing, we evaluate our model using all of the given types of MRI sequences, amongst them T2w provided the best AUC score and accuracy (depicted in Fig.6). Accuracy is measured based on the 50% threshold which means while the probability score of a record is greater than or equal to 0.5, it is defined as 1 otherwise 0. Fig.7 depicts the ROC curve for the ensembling of

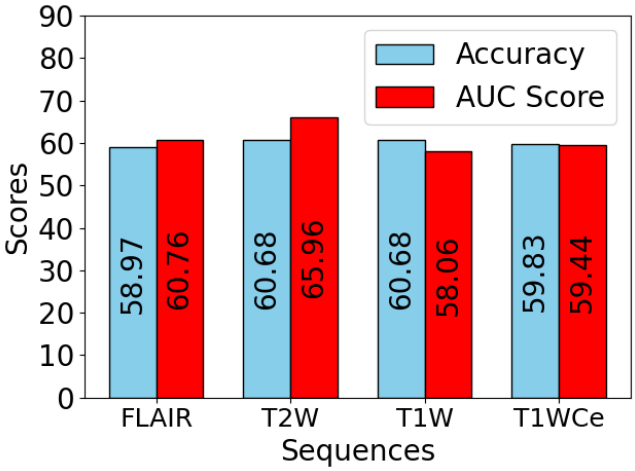

Figure 6. Densenet169 (Regular Interval).

the densenet169 model's output on FLAIR, T2w, and T1w MRI sequences. Max ensemble is providing the best AUC score of 0.66 which is similar to the T2w which is evident that bagging ensembling is not performing well in this scenario.

4.5. Densenet169 (Central Sequence Approach)

Fig.8 illustrates the accuracy and AUC values utilizing the central sequence approach by Deensenet169. When combining 64 image sequences from the middle, instead of at regular intervals, we observed better performance in terms of FLAIR and T2w. We only considered FLAIR and T2w in this processing, as these MRIs contain the liquid profile, and the MGMT profile is based on the liquid profile. However, ensembling the results using the minimum, maximum, or average of the prediction values on different MRIs does not show an improvement from the FALIR only.

4.6. Densenet19 and SAM (Regular Interval)

We have also performed the SAM model to generate masks and radiomic features. It appears that without training, the SAM model achieves a good AUC score, although it lags behind the densenet169. Training the SAM model on

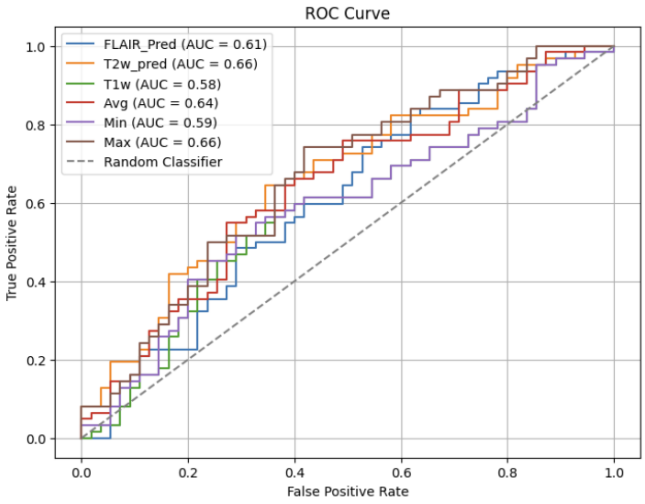

Figure 7. ROC curve of Densenet169 (Regular Interval and ensembling of different types of MRI sequences).

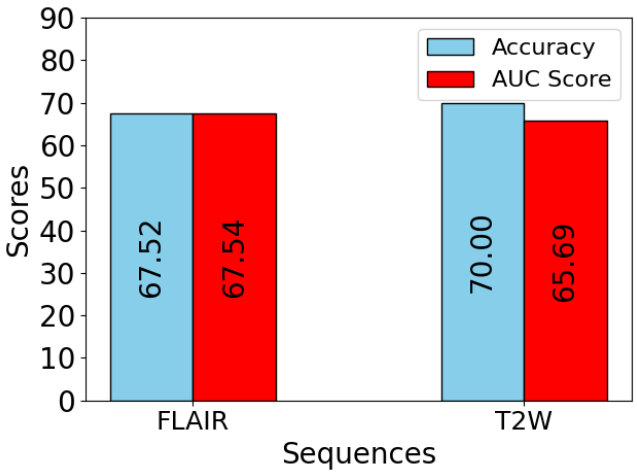

Figure 8. Densenet169 (Central Sequence).

brain tumor MRI images may potentially enhance its performance. However, ensembling SAM and densenet169 on FLAIR data has not yielded a better AUC score as depicted in Fig. 11.

5. Conclusion and Future Work

A new approach to brain tumor classification by predicting the presence of the genomic property, known as O6 -methylguanine DNA methyltransferase (MGMT) promoter methylation is proposed, which integrates radiological imaging with genomic research. Our strategy aims to enhance classification accuracy by utilizing advanced machine learning techniques to extract imaging features from MRI sequences including T1-weighted, T2-weighted, FLAIR, and T1wce. We employ a combination of SAM and Densenet-169 models to enhance efficiency and predict the

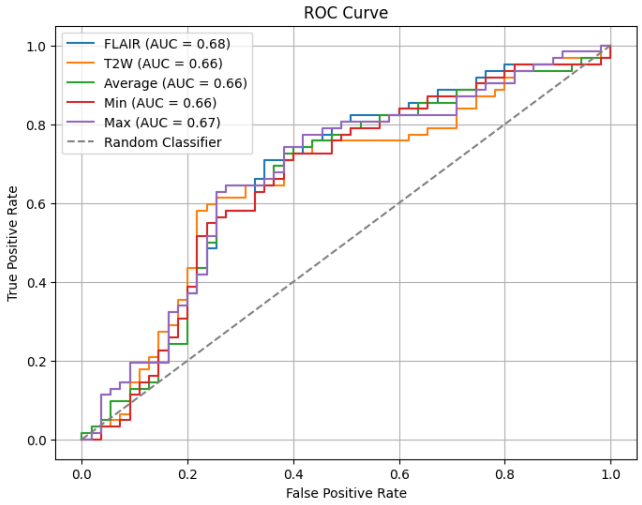

Figure 9. Densenet169 model on FLAIR and T2W with ensembling for Central Image Sequence.

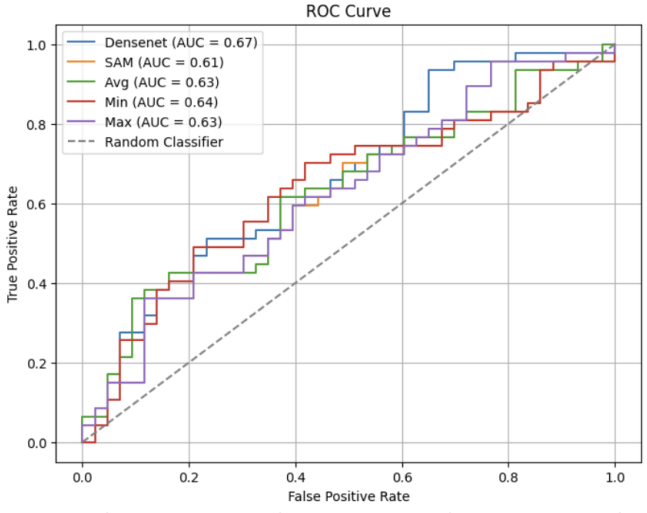

Figure 10. ROC curve of Densenet19 and SAM (regular Interval).

MGMT status. Validated on the RSNA-MICCAI dataset, our method achieves an AUC score of 61% using SAM and an AUC score of 67.54% using Densenet169. These findings highlight the potential of radiogenomics to inform clinical decision-making and tailored non-invasive brain tumor treatment approaches.

Despite achieving these good results, we anticipate that the proposed method still has potential areas for enhancement. In our future work, we plan to train the Medical Segment Anything Model (MedSAM) developed on a large-scale medical image dataset containing 1,570,263 image-mask pairs. We intend to generate more radiomic features like surface area, gray-level co-occurrence matrix (GLCM), run length matrix (RLM), size zone matrix (SZM) using all types of mpMRI sequences (T1w, T1wcE, T2w) other than

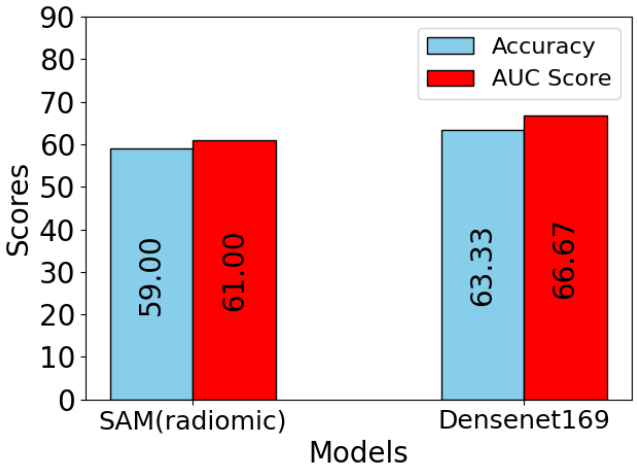

Figure 11. Densenet19 and SAM (Regular Interval).

the FLAIR in the current pipeline. We intend to apply other ensemble methods like Adaboost to justify the performance of individual models.

References

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209–249, 2021. 1

[2] M. E. Mayerhoefer, A. Materka, G. Langs, I. Häggström, P. Szczypiński, P. Gibbs, and G. Cook, "Introduction to radiomics," *Journal of Nuclear Medicine*, vol. 61, no. 4, pp. 488–495, 2020. 1

[3] S. Pálsson, S. Cerri, and K. V. Leemput, "Prediction of mgmt methylation status of glioblastoma using radiomics and latent space shape features," 2021. 2

[4] Available online: https://www.kaggle.com/c/rsna-miccai-brain-tumor-radiogenomic-classification. 2, 5

[5] S. Das, "Optimizing prediction of mgmt promoter methylation from mri scans using adversarial learning," in *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 1047–1054, IEEE, 2022. 2

[6] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, p. 4172, 2023. 2

[7] S. A. Qureshi, L. Hussain, U. Ibrar, E. Alabdulkreem, M. K. Nour, M. S. Alqahtani, F. M. Nafie, A. Mohamed, G. P. Mohammed, and T. Q. Duong, "Radiogenomic classification for mgmt promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpmri scans," *Scientific reports*, vol. 13, no. 1, p. 3291, 2023. 2

[8] A. Mohamed, M. Rabea, A. Sameh, and E. Kamal, "Brain tumor radiogenomic classification," 2024. 2

[9] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4015–4026, 2023. 2

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017. 4