Research Article

Automatic Prediction of MGMT Status in Glioblastoma via Deep Learning-Based MR Image Analysis

Xin Chen [iD],[1] **Min Zeng,**[1] **Yichen Tong,**[2] **Tianjing Zhang,**[3] **Yan Fu,**[4] **Haixia Li,**[2] **Zhongping Zhang,**[3] **Zixuan Cheng,**[1] **Xiangdong Xu,**[1] **Ruimeng Yang,**[1] **Zaiyi Liu** [iD],[5] **Xinhua Wei** [iD],[1] **and Xinqing Jiang** [iD][1]

[1]*Department of Radiology, Guangzhou First People's Hospital, Guangzhou Medical University, Guangzhou 510180, China*
[2]*Sun Yat-sen University, Guangzhou, China*
[3]*Philips Healthcare, Guangzhou, China*
[4]*EPFL, Lausanne, Switzerland*
[5]*Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China*

Correspondence should be addressed to Zaiyi Liu; zyliu@163.com, Xinhua Wei; weixinhua@aliyun.com, and Xinqing Jiang; gzcmmcjxq@163.com

Methylation of the O^6-methylguanine methyltransferase (MGMT) gene promoter is correlated with the effectiveness of the current standard of care in glioblastoma patients. In this study, a deep learning pipeline is designed for automatic prediction of MGMT status in 87 glioblastoma patients with contrast-enhanced T1W images and 66 with fluid-attenuated inversion recovery(FLAIR) images. The end-to-end pipeline completes both tumor segmentation and status classification. The better tumor segmentation performance comes from FLAIR images (Dice score, 0.897 ± 0.007) compared to contrast-enhanced T1WI (Dice score, 0.828 ± 0.108), and the better status prediction is also from the FLAIR images (accuracy, 0.827 ± 0.056; recall, 0.852 ± 0.080; precision, 0.821 ± 0.022; and F_1 score, 0.836 ± 0.072). This proposed pipeline not only saves the time in tumor annotation and avoids interrater variability in glioma segmentation but also achieves good prediction of MGMT methylation status. It would help find molecular biomarkers from routine medical images and further facilitate treatment planning.

1. Introduction

Glioblastoma multiforme (GBM) is the most common and aggressive type of primary brain tumor in adults. It accounts for 45% of primary central nervous system tumors, and the 5-year survival rate is around 5.1% [1, 2]. The standard treatment for GBM is surgical resection followed by radiation therapy and temozolomide (TMZ) chemotherapy, which improves median survival by 3 months compared to radiotherapy alone [3]. Several studies indicated that O^6-methylguanine-DNA methyltransferase (MGMT) gene promoter methylation reported in 30-60% of glioblastomas [4] can enhance the response to TMZ, which has been proven to be a prognostic biomarker in GBM patients [3, 5]. Thus, deter-mination of MGMT promoter methylation status is important to medical decision-making.

Genetic analysis based on surgical specimens is the reference standard to assess the MGMT methylation status, while a large tissue sample is required for testing MGMT methylation status using methylation-specific polymerase chain reaction [6]. In particular, the major limitations are the possibility of incomplete biopsy samples due to tumor spatial heterogeneity and high cost [7]. Besides, it cannot be used for real-time monitoring of the methylation status.

Magnetic resonance imaging (MRI) is a standard conventional examination in diagnosis, preoperative planning, and therapy evaluation of GBM [8, 9]. Recently, radiomics, extracting massive quantitative features from medical

TABLE 1: Dataset distribution of each experiment.

	Phase	Cases (methylation/unmethylation)	CE-T1WI slices (methylation/unmethylation)	FLAIR slices (methylation/unmethylation)
FLAIR	Training	51 (25/26)	676 (288/388)	—
	Testing	15 (7/8)	167 (62/105)	
CE-T1WI	Training	70 (36/34)	—	1208 (609/599)
	Testing	17 (10/7)		220 (109/111)

Note: FLAIR: fluid-attenuated inversion recovery; CE-T1WI; contrast-enhanced T1-weighted imaging.

images, has been proposed to explore the correlation between image features and underlying genetic traits [10–12]. There is growing evidence that radiomics can be used in predicting the status of MGMT promoter methylation [13–15]. However, most previous works utilized handcrafted features. This procedure includes tumor segmentation, feature extraction, and informatics analysis [16–19]. In particular, tumor segmentation is a challenging and important step because most works depend on manual delineation. This step is burdensome and time consuming, and inter- or intraobserver disagreement is unavoidable. Deep learning which can extract features automatically has been emerging as an innovative technology in many fields [20]. The convolutional neural network (CNN) is proven to be effective in image segmentation, disease diagnosis, and other medical image analysis tasks [21–25]. Compared to traditional methods with handcrafted features, deep learning shows several advantages of being robust to distortions such as changes in shape and lower computational cost. A few studies have shown that deep learning can be used to segment tumors and predict MGMT methylation status for glioma [26]. However, to the best of our knowledge, there is no previous report regarding building a pipeline for both glioma tumor segmentation and MGMT methylation status prediction in an end-to-end manner. Therefore, we investigate the feasibility of integrating the tumor segmentation and status prediction of GBM patients into a deep learning pipeline in this study.

2. Methods

2.1. Data Collection. A total of 106 GBM patients were analyzed in our study. MR images, including presurgical axial contrast-enhanced T1-weighted images (CE-T1WI) and T2-weighted fluid-attenuated inversion recovery (FLAIR) images, were collected from The Cancer Imaging Archive (http://www.cancerimagingarchive.net). The images were originated from four centers (Henry Ford Hospital, University of California San Francisco, Anderson Cancer Center, and Emory University). Clinical and molecular data were also obtained from the open-access data tier of the TCGA website.

Genomic data were from the TCGA data portal. MGMT methylation status analysis was performed on Illumina HumanMethylation27 and HumanMethylation450 BeadChip platforms. A median cutoff using the level 3 beta-value present in the TCGA was utilized for categorizing methylation status. Illumina Human Methylation probes

(cg12434587 and cg12981137) were selected in this study [27].

Of 106 GBM cases, 87 cases were with CE-T1W images, and 66 cases with FLAIR images. We randomly split the cases into training and testing sets with the ratio of 8 : 2 and applied 10-fold cross-validation to the training set with scikit-learn library (https://scikit-learn.org/stable/). The dataset distribution is listed in Table 1.

2.2. Image Preprocessing. For general images, the pixel values contain reliable image information. However, MR images do not have a standard intensity scale. In Figure 1(a), we show the density plot of two raw MR images. In each plot, there are two peaks, the peak around 0 refers to background pixels, and the other peak refers to white matter. The white matter peaks of the two images are far away. Thus, MR images normalization is needed to guarantee that the grey values of the same tissue among different MR images are close to each other [28].

The piece-wise linear histogram matching was used to normalize the intensity distribution of MR images [29]. Firstly, we studied standard histogram distribution via averaging the 1st to 99th percentile of all images. Then, we linearly mapped the intensities of each image to this standard histogram. In Figure 1(b), we can see that the white matter peaks of two images coincide with each other after normalization. Secondly, the images were normalized to zero mean and unit standard deviation only on valued voxels. At last, data augmentation was used to increase the dataset size to avoid overfitting. We rotated images for every 5 degrees from -20 to +20 degrees, resulting in a 9-fold increment in the number of MRI scans.

2.3. Segmentation. As for tumor segmentation, one state-of-the-art model [30] in BraTS 2018 challenge (Multimodal Brain Tumor Segmentation 2018 Challenge http://braintumorsegmentation.org/) was adapted. The whole network architecture is shown in Figure 2.

In short, the deep learning model added a variational autoencoder (VAE) branch to a fully convolutional network model. The decoder part was shared for both segmentation and VAE tasks. The prior distribution taken for the KL divergence in the VAE part is $N(0, 1)$. ResNet blocks used in the architecture [31] included two 3×3 convolutions with normalization and ReLU as well as skip connections. In the encoder part, the image dimension was downsampled using stride convolution by 2 and increased channel size by 2. For the decoder part, the structure was similar to that of the
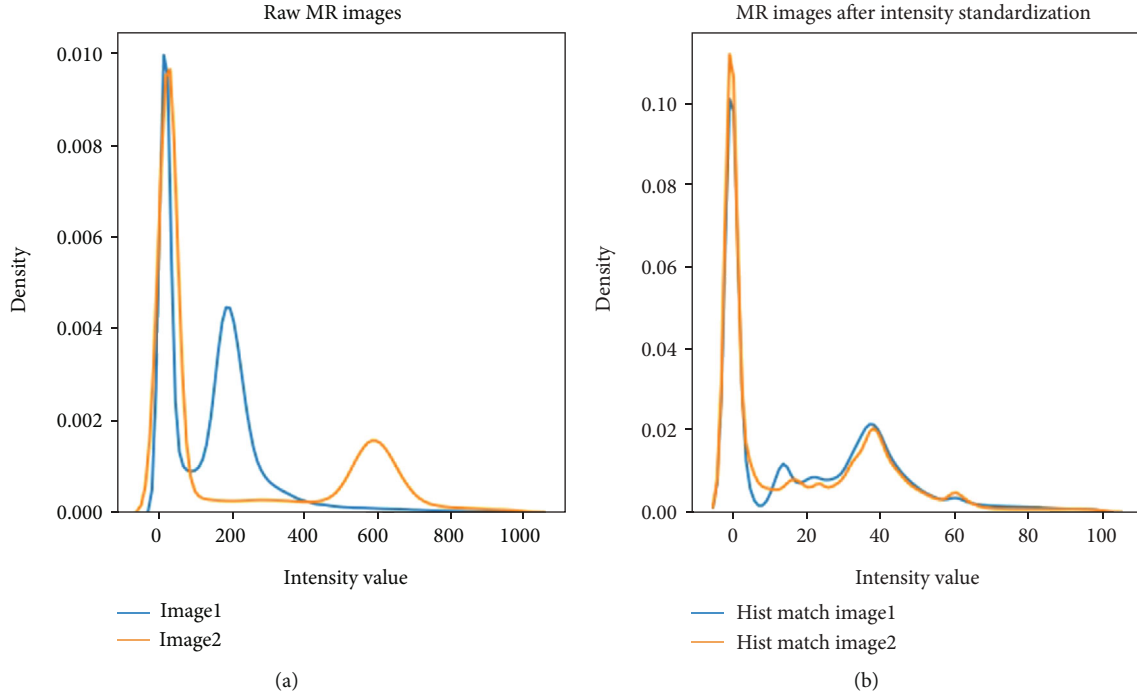
FIGURE 1: Density plot of two different MR images (a) before and (b) after piece-wise linear histogram matching.
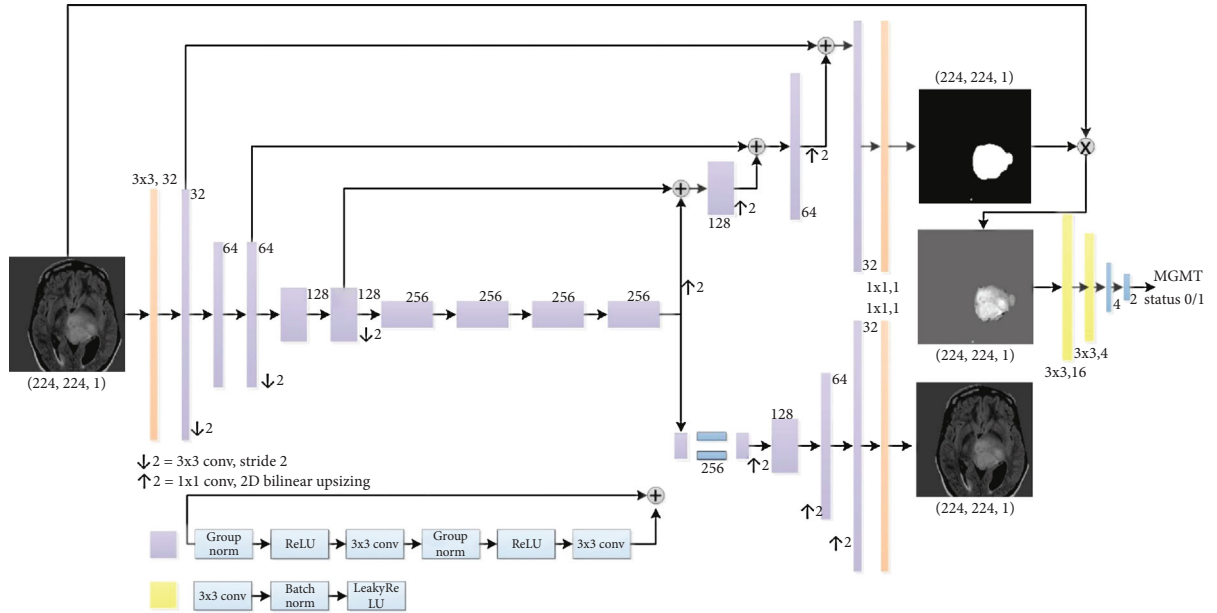

FIGURE 2: An end-to-end deep learning pipeline for both tumor segmentation and status classification.

encoder part but using upsampled. The decoder endpoint had the same size as the input image followed by sigmoid activation, and its output was for tumor segmentation. As for the VAE part, the encoder output was reduced to 256, and the input image was reconstructed by using a similar structure as the decoder without skip connection. The segmentation part output the tumor segmentation and the VAE branch attempted to reconstruct the input image. Except for the input and output layers, all blocks in

Figure 2 utilized the ResNet block with different channel numbers (depicted aside each layer). For the input layer, a 3×3 convolution was with 3 channels; and for both output layers, a 3×3 convolution with a dropout rate of 0.2 and L_2 regularization with weight $1e - 3$ were used to avoid overfitting. The loss function consists of 3 terms as shown in

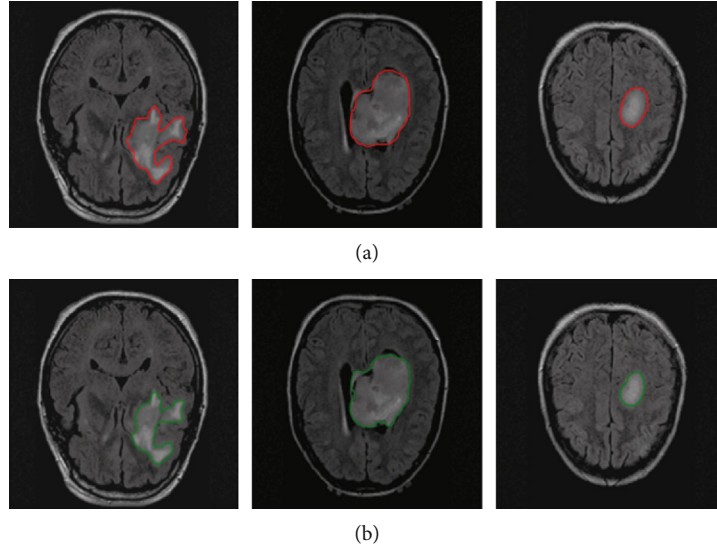$$L = L_{\text{Dice}} + 0.1 \times L_{L2} + 0.1 \times L_{KL}, \tag{1}$$

(a)

(b)

FIGURE 3: Automatic segmentation results of brain tumors with FLAIR images. (a) The ground truth of tumor boundaries in FLAIR images and (b) automatic segmentation results using the proposed network with FLAIR images.
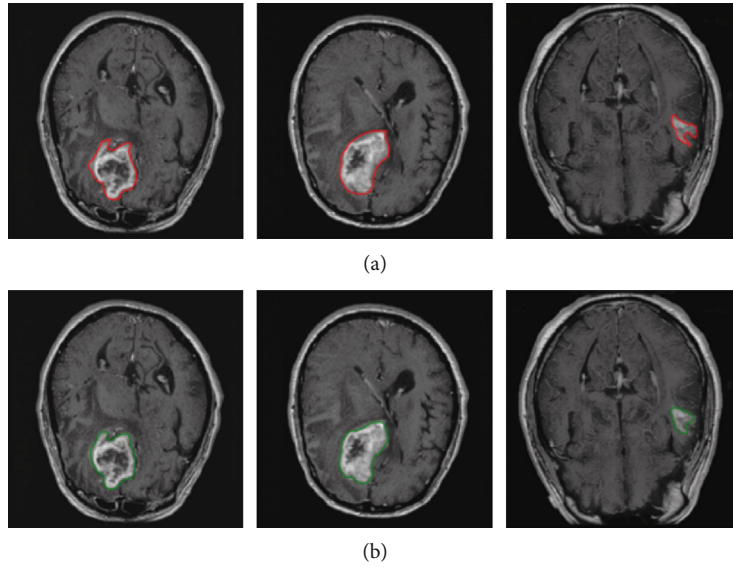

(a)

(b)

FIGURE 4: Three representative cases of brain tumor manual annotation and automatic segmentation with CE-T1WI images. (a) The manual annotation and (b) the automatic segmentation results with our proposed network.

where L_{Dice} is the soft Dice loss between the predicted segmentation and the ground truth labels. The ground truth labels were manually annotated with ImageJ (https://imagej.nih.gov) by one neuroradiologist with 10 years' experience specialized in brain disease diagnosis. L_{L2} is the L_2 loss on the VAE branch output image and the input image, and L_{KL} is the standard VAE penalty term [32, 33]. Then, the Dice coefficient as defined in function (2) was calculated to assess the performance of segmentation:

$$\mathrm{Dice} = \sum_i \frac{2 \cdot p_i \cdot \widehat{p}_i}{\|p_i\|^2 + \|p_i \wedge\|^2 + \mathrm{epsilon}}, \tag{2}$$

where p_i is the ground truth, $\widehat{p}_i$ is the prediction for pixel i, and epsilon $= 1e - 8$.

2.4. Status Classification.

Meanwhile, for the classification of MGMT methylation status, a 4-layer CNN was designed. Further, the classification model was cascaded with the tumor segmentation model. At the stage of the tumor segmentation model design, the classification network was tried with different numbers of convolutional layers [2–5], and we found that 2 convolutional layers with 2 fully connected (FC) layers performed the best for this task. The first convolutional layer had 16 filters, and the second one had 4 filters. All the convolutional layers had a kernel size of 3×3 and stride of

TABLE 2: Dice scores of the deep network on tumor segmentation using MR images.

Modality	Training	Validation	Testing
CE-T1WI	0.832 ± 0.009	0.831 ± 0.012	0.828 ± 0.108
FLAIR	0.893 ± 0.004	0.892 ± 0.008	0.897 ± 0.007

Note: the number in the table referred to the mean ± standard deviation values of 10 cross-validation experiments. CE-T1WI: contrast-enhanced T1-weighted imaging; FLAIR: fluid-attenuated inversion recovery.

TABLE 3: Inference time (seconds) of one MR slice for glioma segmentation.

Modality	Manual annotation	Deep model
CE-T1WI	50 s	0.11 s
FLAIR	60 s	0.07 s

Note: CE-T1WI: contrast-enhanced T1-weighted imaging; FLAIR: fluid-attenuated inversion recovery.

1 followed by LeakyReLU, batch normalization, and max pooling. LeakyReLU was an advanced ReLU activation that avoids dead neurons by setting a negative half-axis slope 0.3 instead of 0. Its advantages include good performance in eliminating gradient saturation, low computational cost, and faster convergence. Batch normalization was used to normalize features by the mean and variance within a small batch. It helped to solve the covariance shift issue and ease optimization. Max pooling with a 4×4 filter was used to downsample image features extracted through convolutional layers and then fed into 2 FC layers. ReLU and softmax were adapted as activation functions for the first and second FC layers, respectively. The weight initialization of all layers was done by He-normal [34].

2.5. Parameter Settings and Software. All experiments were conducted under the open-source framework Keras (https://keras.io/) on one GeForce RTX 2080Ti GPU. The numbers of parameters of the segmentation and classification model are, respectively, 6,014,721 and 3,498. In tumor segmentation, Adam optimizer was adapted with a self-designed learning rate scheduler which was initialized with a learning rate $1e - 4$; then, the learning rate was divided by 2 when the validation loss did not reduce in the past 5 epochs. The epoch was set at 50 and batch size at 8. Every epoch took around 50 seconds. In tumor classification, 4-CNN was trained for 50 epochs which utilized Adam with learning rate $2e - 4$, and the batch size was 32. If the validation accuracy was observed stable for over 10 epochs, the training process would be ended. The averaged elapsed time for each epoch was 5 seconds.

2.6. Statistical Analysis. The Dice coefficient was calculated for evaluating the performance of tumor segmentation. For the MGMT methylation status classification, the accuracy rate, recall, precision, and F_1 score were calculated according to equations listed below. In addition, the receiver operating characteristic (ROC) curve was plotted, and the area under the ROC curve (AUC) was reported to measure the classification accuracy. All the parameters were calculated in PyCharm

with the programming language of Python (version 3.6.8; Wilmington, DE, USA; http://www.python.org/):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$
$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{3}$$

where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative.

3. Results

3.1. Tumor Segmentation

3.1.1. Qualitative Observation. Tumors could be accurately delineated by the proposed pipeline. Figure 3 shows the annotated ground truth (the first row) and corresponding segmentation results (the second row) of GBM in FLAIR images. It is observed that tumor boundaries could be accurately localized by using the deep learning network, and the major hyperintense regions are delineated. The three cases show that automatic segmentation is quite close to the ground truth.

Figure 4 shows the GBM in CE-T1WI images, and the ground truth (the first row) and the segmentation results (the second row) are presented. Tumor boundaries are localized, and it seems that there is no obvious difference between the manual annotation and its corresponding segmentation results obtained from our proposed network, and the suspicious regions are mainly contoured. The three cases show that segmentation results from the deep network approximate the manual delineation.

3.1.2. Quantitative Evaluation. The quantitative performance of automatic tumor segmentation is summarized in Table 2. The deep network obtained good testing performance on tumor segmentation using CE-T1WI (Dice score, 0.828 ± 0.108) and FLAIR (Dice score, 0.897 ± 0.007). And the Dice scores from FLAIR were slightly higher than those from CE-T1WI across training, validation, and testing sets. The maximum difference of the Dice score between average Dice scores from CE-T1W images in training and validation sets was 0.026, indicating that the model was not overfitting.

3.1.3. Computational Performance. Time consumption between manual annotation and automatic prediction per MR slice is compared as shown in Table 3. For the evaluation of time consumption, we recorded the total time and divided it by the number of slices. So, the time listed in Table 3 was the average segmentation time per slice. It was observed that the deep network was more efficient, and it took less than 0.2 seconds to complete the segmentation of an MR slice, while manual annotation required more than 30 seconds.

TABLE 4: Results of MGMT methylation status classification.

Modality	Phase	Classification			
		Accuracy	Recall	Precision	F_1 score
CE-T1WI	Training	0.894 ± 0.012	0.906 ± 0.007	0.886 ± 0.018	0.896 ± 0.010
	Validation	0.839 ± 0.046	0.866 ± 0.044	0.823 ± 0.051	0.845 ± 0.045
	Testing	0.804 ± 0.011	0.818 ± 0.033	0.798 ± 0.014	0.808 ± 0.015
FLAIR	Training	0.941 ± 0.056	0.943 ± 0.104	0.947 ± 0.026	0.945 ± 0.081
	Validation	0.885 ± 0.090	0.941 ± 0.105	0.857 ± 0.028	0.889 ± 0.101
	Testing	0.827 ± 0.056	0.852 ± 0.080	0.821 ± 0.022	0.836 ± 0.072

Note: the number in the table referred to the mean ± standard deviation values of 10 cross-validation experiments. CE-T1WI: contrast-enhanced T1-weighted imaging; FLAIR: fluid-attenuated inversion recovery.

3.2. Classification of MGMT Promoter Methylation Status. Table 4 shows the prediction performance of MGMT promoter methylation status which is evaluated from four classification metrics (accuracy, recall, precision, and F_1 score) on three stages (training, validation, and testing) when using different MR images (CE-T1WI, FLAIR). In general, the model trained with FLAIR achieves better results for all metrics across three stages, followed by the model trained with CE-T1WI images. Specifically, the accuracy, recall, precision, and F_1 score of the deep model trained with FLAIR images reach 0.827, 0.852, 0.821, and 0.836 in the testing stage, respectively.

ROC curves of the prediction results are demonstrated in Figures 5 and 6. Figure 5 shows the best status classification when using FLAIR images for a deep model, which achieves an AUC of 0.985 (yellow curve), 0.968 (green curve), and 0.905 (red curve) on the training, validation, and testing datasets, respectively.

The best status classification when using CE-T1WI images for deep model training is shown in Figure 6. The well-trained deep model obtains AUC up to 0.973 (yellow curve), 0.942 (green curve), and 0.887 (red curve) on the training, validation, and testing datasets, respectively.

4. Discussion

This study presents an MR-based deep learning pipeline for automatic tumor segmentation and MGMT methylation status classification in an end-to-end manner for GBM patients. Experimental results demonstrate promising performance on accurate glioma delineation (Dice score, 0.897) and MGMT status prediction (accuracy, 0.827; recall, 0.852; precision, 0.821; and F_1 score, 0.836) coming from the model trained with FLAIR images. In addition, the proposed pipeline dramatically shortens the inference time on glioma segmentation.

For glioma segmentation, one state-of-the-art deep model is utilized and obtains impressive performance on the involved MGMT dataset for GBM segmentation. Its performance is close to these deep network-based tumor segmentation studies. Hussain et al. [35] reported a CNN approach for glioma MRI segmentation, and the model achieved a Dice score of 0.87 on the BRATS 2013 and 2015 datasets. Cui et al. [36] proposed an automatic semantic segmentation model on the BRATS 2013 dataset, and the Dice score was near 0.80 on the combined high- and low-grade glioma datasets. Kaldera et al. [37] proposed a faster RCNN method and achieved a Dice score of 0.91 on 233 patients' data. These studies suggest that deep networks are full of potential for accurate tumor segmentation in MR images.

Several deep models have been designed for the classification of MGMT methylation status in GBM patients. Chang et al. [38] proposed a deep neural network which achieved a classification accuracy of 83% for 259 gliomas patients with T1W, T2W, and FLAIR images. Korfiatis et al. [26] compared different sizes of the ResNet baseline model and reached the highest accuracy of 94.9% in 155 GBM patients with T2W images. Han et al. [39] proposed a bidirectional convolutional recurrent neural network architecture for MGMT methylation classification, while the accuracy was around 62% for 262 GBM patients with T1W, T2W, and FLAIR images. In this study, a shallow CNN is used, and the classification performance is promising. The best performance comes from the model trained with FLAIR images, and we achieved a satisfactory result with the highest accuracy of 0.827 and recall of 0.852 in consideration of the relatively small dataset.

In the previous studies, Drabycz et al. [40] analyzed handcrafted features to distinguish methylated from unmethylated GBM and figured out that texture features from T2-weighted images were important for the prediction of MGMT methylation status. Han et al. [41] found that MGMT promoter-methylated GBM was prone to more tumor necrosis, while T2-weighted FLAIR sequence may be more sensitive to necrosis than T1-weighted images. Interestingly, we also find that better performances of both GBM segmentation and molecular classification are achieved on FLAIR images in our study although the images of CE-T1W and FLAIR did not come from the same patients.

The strengths of this study lie in the fully automatic glioma segmentation and predicting the MGMT methylation status based on a small dataset. Generally, it takes a radiologist about one minute per slice in tumor annotation, while the inference time of the deep learning model is about 0.1 seconds which is around 1/600 times used in manual annotation. Additionally, manual annotation is burdensome and prone to introduce inter- and intraobserver variability. While once well trained, a deep learning model can continuously
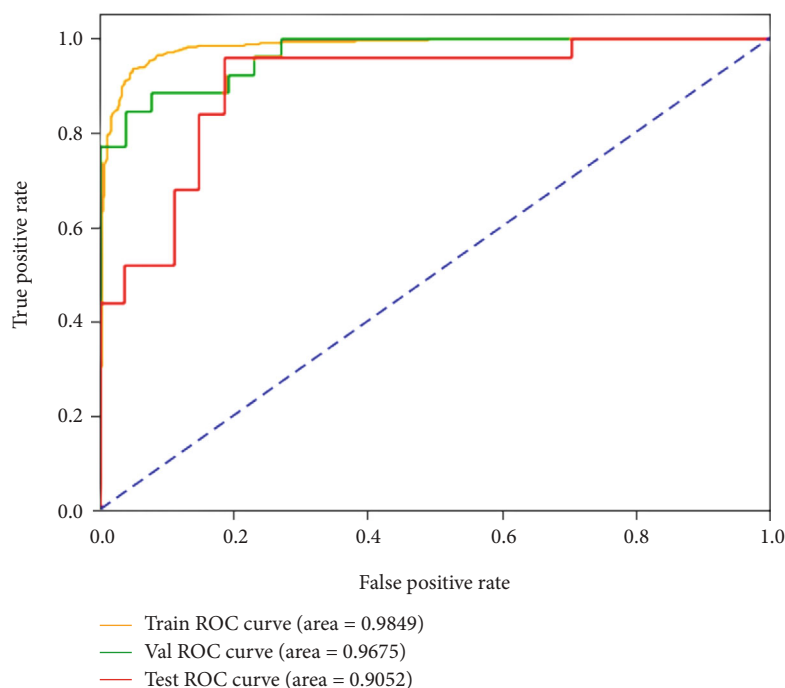
FIGURE 5: ROC curves of the best result on the FLAIR images for MGMT promoter methylation status classification on the training, validation, and testing datasets.
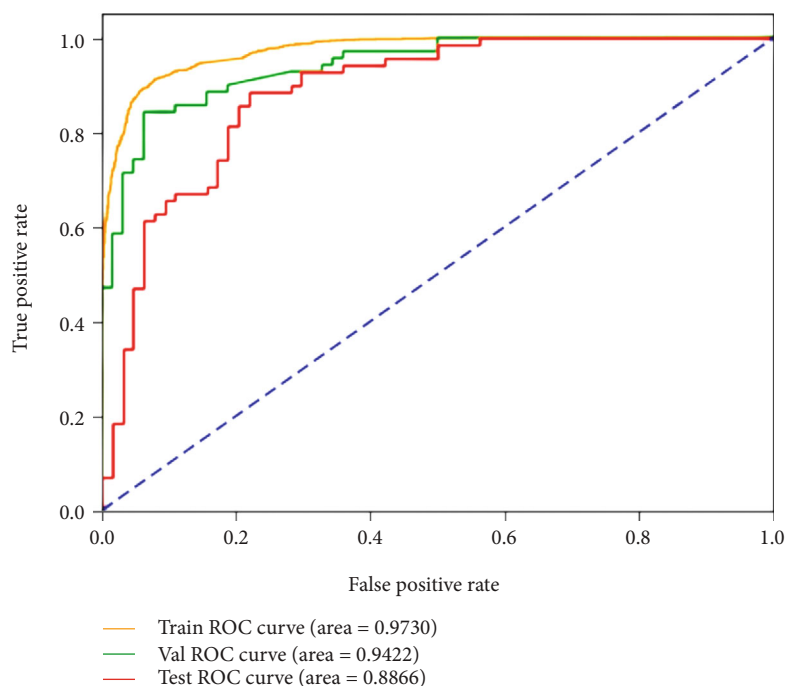

FIGURE 6: ROC curves of the best result on the CE-T1W images for MGMT promoter methylation status classification in the training, validation, and testing datasets.

and repeatedly perform tumor segmentation regardless of the observers. On the other hand, the training strategy in this study is beneficial for small dataset analysis. In general, a deep model requires a large number of training instances. However, it is challenging or impossible to provide massive high-quality images in medical imaging. Finally, although several studies tried to use deep networks for automatic glioma segmentation [35, 36, 42] or molecular classification [26, 38, 39], the proposed network in this study could integrate both glioma segmentation and classification in a

seamless connection pipeline. And the performance is competitive to the state-of-the-art studies in tumor segmentation and classification.

There are several limitations to our study. First, the sample size is small in the study; we will further confirm the findings in a study with larger samples. Second, a multicenter research trial is helpful to validate the capability of the proposed pipeline, while the variations of MR imaging sequences, equipment venders, and other factors could impose difficulties on model building. Third, we failed to investigate the value of combined CE-T1WI and FLAIR in tumor segmentation and classification considering the fewer samples. In the future, we will explore multiple MR sequences for MGMT methylation status prediction, such as amide-proton-transfer-weighted imaging and diffusion-weighted imaging. These may have great potential to improve the performance of MGMT methylation status prediction.

5. Conclusion

An MRI-based end-to-end deep learning pipeline is designed for tumor segmentation and MGMT methylation status prediction in GBM patients. It can save time and avoid interobserver variability in tumor segmentation and help discover molecular biomarkers from routine medical images to aid in diagnosis and treatment decision-making.

Data Availability

All MRI data are available in the cancer imaging archive (https://www.cancerimagingarchive.net/), and clinical and molecular data are obtained from the open-access data tier of the TCGA website.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Authors' Contributions

Xin Chen, Min Zeng, and Yichen Tong contributed equally.

Acknowledgments

References

[1] L. L. Morgan, "The epidemiology of glioma in adults: a "state of the science" review," *Neuro-Oncology*, vol. 17, no. 4, pp. 623-624, 2015.

[2] Q. T. Ostrom, H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, and J. S. Barnholtz-Sloan, "CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2011-2015," *Neuro-oncology*, vol. 20, supplement_4, pp. iv1–iv86, 2018.

[3] M. E. Hegi, A. C. Diserens, T. Gorlia et al., "*MGMT* Gene Silencing and Benefit from Temozolomide in Glioblastoma," *The New England Journal of Medicine*, vol. 352, no. 10, pp. 997–1003, 2005.

[4] M. Weller, R. Stupp, G. Reifenberger et al., "*MGMT* promoter methylation in malignant gliomas: ready for personalized medicine?," *Nature Reviews Neurology*, vol. 6, no. 1, pp. 39–51, 2010.

[5] A. A. Brandes, E. Franceschi, A. Tosoni et al., "*MGMT* Promoter Methylation Status Can Predict theIncidence and Outcome of Pseudoprogression After Concomitant Radiochemotherapy in Newly Diagnosed Glioblastoma Patients," *Journal of Clinical Oncology*, vol. 26, no. 13, pp. 2192–2197, 2008.

[6] L. Wang, Z. Li, C. Liu et al., "Comparative assessment of three methods to analyze MGMT methylation status in a series of 350 gliomas and gangliogliomas," *Pathology-Research and Practice*, vol. 213, no. 12, pp. 1489–1493, 2017.

[7] N. R. Parker, A. L. Hudson, P. Khong et al., "Intratumoral heterogeneity identified at the epigenetic, genetic and transcriptional level in glioblastoma," *Scientific Reports*, vol. 6, no. 1, article 22477, 2016.

[8] P. Y. Wen, D. R. Macdonald, D. A. Reardon et al., "Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group," *Journal of clinical oncology*, vol. 28, no. 11, pp. 1963–1972, 2010.

[9] M. J. van den Bent, J. S. Wefel, D. Schiff et al., "Response assessment in neuro-oncology (a report of the RANO group): assessment of outcome in trials of diffuse low-grade gliomas," *The lancet oncology*, vol. 12, no. 6, pp. 583–593, 2011.

[10] L. Macyszyn, H. Akbari, J. M. Pisapia et al., "Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques," *Neuro-Oncology*, vol. 18, no. 3, pp. 417–425, 2015.

[11] L. S. Hu, S. Ning, J. M. Eschbacher et al., "Radiogenomics to characterize regional genetic heterogeneity in glioblastoma," *Neuro-Oncology*, vol. 19, no. 1, pp. 128–137, 2017.

[12] E. K. Hong, S. H. Choi, D. J. Shin et al., "Radiogenomics correlation between MR imaging features and major genetic profiles in glioblastoma," *European radiology*, vol. 28, no. 10, pp. 4350–4361, 2018.

[13] V. G. Kanas, E. I. Zacharaki, G. A. Thomas, P. O. Zinn, V. Megalooikonomou, and R. R. Colen, "Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma," *Computer Methods and Programs in Biomedicine*, vol. 140, pp. 249–257, 2017.

[14] Y. B. Xi, F. Guo, Z. L. Xu et al., "Radiomics signature: a potential biomarker for the prediction of MGMT promoter methylation in glioblastoma," *Journal of Magnetic Resonance Imaging*, vol. 47, no. 5, pp. 1380–1387, 2018.

[15] F. Tixier, H. Um, D. Bermudez et al., "Preoperative MRI-radiomics features improve prediction of survival in

glioblastoma patients over MGMT methylation status alone," *Oncotarget*, vol. 10, no. 6, pp. 660–672, 2019.

[16] P. Kickingereder, U. Neuberger, D. Bonekamp et al., "Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma," *Neuro-Oncology*, vol. 20, no. 6, pp. 848–857, 2018.

[17] Z. Liu, X. Y. Zhang, Y. J. Shi et al., "Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer," *Clinical Cancer Research*, vol. 23, no. 23, pp. 7253–7262, 2017.

[18] P. Lambin, E. Rios-Velazquez, R. Leijenaar 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.

[19] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar et al., "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature Communications*, vol. 5, no. 1, 2014.

[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[21] P. Rajpurkar, J. Irvin, R. L. Ball et al., "Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheX-NeXt algorithm to practicing radiologists," *PLoS medicine*, vol. 15, no. 11, article e1002686, 2018.

[22] Y. Yang, L. F. Yan, X. Zhang et al., "Glioma grading on conventional MR images: a deep learning study with transfer learning," *Frontiers in Neuroscience*, vol. 12, 2018.

[23] A. Akay and H. Hess, "Deep learning: current and emerging applications in medicine and technology," *EEE journal of biomedical and health informatics*, vol. 23, no. 3, pp. 906–920, 2019.

[24] G. S. Tandel, M. Biswas, O. G. Kakde et al., "A review on a deep learning perspective in brain cancer classification," *Cancers*, vol. 11, no. 1, 2019.

[25] L. Zou, S. Yu, T. Meng, Z. Zhang, X. Liang, and Y. Xie, "A technical review of convolutional neural network-based mammographic breast cancer diagnosis," *Computational and mathematical methods in medicine*, vol. 2019, Article ID 6509357, 16 pages, 2019.

[26] P. Korfiatis, T. L. Kline, D. H. Lachance, I. F. Parney, J. C. Buckner, and B. J. Erickson, "Residual deep convolutional neural network predicts MGMT methylation status," *Journal of Digital Imaging*, vol. 30, no. 5, pp. 622–628, 2017.

[27] P. Bady, D. Sciuscio, A.-C. Diserens et al., "*MGMT* methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status," *Acta Neuropathologica*, vol. 124, no. 4, pp. 547–560, 2012.

[28] J. C. Reinhold, B. E. Dewey, A. Carass, and J. L. Prince, "Evaluating the impact of intensity normalization on MR image synthesis," in *SPIE Medical Imaging*, vol. 10949, San Diego, California, United States, 2019.

[29] M. Shah, Y. Xiao, N. Subbanna et al., "Evaluating intensity normalization on MRIs of human brain with multiple sclerosis," *Medical Image Analysis*, vol. 15, no. 2, pp. 267–282, 2011.

[30] A. Myronenko, *3D MRI Brain Tumor Segmentation Using Autoencoder Regularization*, International MICCAI Brainlesion Workshop: Springer, 2019.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, 2016.

[32] C. Doersch, "Tutorial on variational autoencoders," 2016, https://arxiv.org/abs/1606.05908.

[33] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, https://arxiv.org/abs/1312.6114.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, Santiago, Chile, 2015.

[35] S. Hussain, S. M. Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, vol. 282, pp. 248–261, 2018.

[36] S. Cui, L. Mao, J. Jiang, C. Liu, and S. Xiong, "Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network," *Journal of healthcare engineering*, vol. 2018, Article ID 4940593, 14 pages, 2018.

[37] H. Kaldera, S. Gunasekara, and M. B. Dissanayake, "MRI based glioma segmentation using deep learning algorithms," in *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, pp. 51–56, Colombo, Sri Lanka, 2019.

[38] P. Chang, J. Grinband, B. D. Weinberg et al., "Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas," *AJNR. American Journal of Neuroradiology*, vol. 39, no. 7, pp. 1201–1207, 2018.

[39] L. Han and M. R. Kamdar, "MRI to MGMT: predicting methylation status in glioblastoma patients using convolutional recurrent neural networks," *Biocomputing*, vol. 23, pp. 331–342, 2018.

[40] S. Drabycza, G. Roldánbcde, P. Roblesbde et al., "An analysis of image texture, tumor location, and *MGMT* promoter methylation in glioblastoma using magnetic resonance imaging," *NeuroImage*, vol. 49, no. 2, pp. 1398–1405, 2010.

[41] Y. Han, L. F. Yan, X. B. Wang et al., "Structural and advanced imaging in predicting MGMT promoter methylation of primary glioblastoma: a region of interest based analysis," *BMC Cancer*, vol. 18, no. 1, 2018.

[42] S. Wu, H. Li, D. Quang, and Y. Guan, "Three-plane-assembled deep learning segmentation of gliomas," *Radiology: Artificial Intelligence*, vol. 2, no. 1, article e190011, 2020.