Proceedings of CONF-CDS 2026 Symposium: Machine Learning and Neural Network Applications in Engineering
DOI: 10.54254/2755-2721/2026.AST37396

Review of Multimodal MRI Brain Tumor Segmentation Methods in the Absence of Modality

Zixuan Zeng

School of Life Sciences, Nantong University, Nanjing, China
zzz040126@outlook.com

Abstract. Nowadays, the global prevalence of brain tumors has been increasing steadily. Magnetic resonance imaging (MRI) is widely used in clinical practice to detect brain abnormalities owing to its noninvasive nature. By acquiring multiple pulse sequences, MRI enables tissue characterization with complementary contrasts from a single examination and provides imaging data for subsequent diagnosis. However, prolonged scan time, patient motion and hardware limitations may result in the absence of one or more MRI sequences in practice thus affecting the diagnostic accuracy. Accurate brain tumor segmentation under missing modality condition is stillamazingly challenging in medical image analysis. The review focuses on multimodal brain MRI: T1-weighted, T2-weighted andfluid attenuated inversion recovery (FLAIR) sequences as well as automated brainlesion segmentation. We organize and compare thestate-of-the-art methods for missing modality in a systematic way includingmodality synthesis, shared latent-space learning, knowledge distillation,mutual-information-based learning, Transformer-based fusion and Markov chainbased state-space models. The paper finds that complete-modality inputs provide an important foundation for stable segmentation; the absence of T1-weighted imaging may impair anatomical localization while the absence of T2-weighted or FLAIR imaging may impair the delineation of edema and the whole-tumor region.

Keywords: Multimodal MRI, Tumor Segmentation, Missing-Modality, Deep Learning

1. Introduction

Brain imaging is a noninvasive approach for visualizing the brain, and X-ray, ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) are widely used in clinically imaging modalities. For detailed brain assessment, X-ray imaging is not very often used. The reason is that its soft-tissue contrast is limited, and the structural information is not complete. CT is a more advanced modality, and it is based on X-ray. It acquires projections from multiple angles. Then, computer-assisted reconstruction is used to generate three-dimensional representations from two-dimensional cross-sectional images. However, CT depends on ionizing radiation. This may cause tissues to be exposed to a relatively high level of radiation. MRI is different. It is very suitable for visualizing and segmenting brain-tissue structures in the axial, sagittal, and coronal planes. Also, it provides superior soft-tissue contrast. MRI generates two- or three-dimensional images by using magnetic fields and radiofrequency waves, not by using ionizing radiation. Therefore, it is safer than

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.AST37396

CT. The MRI sequences that are commonly used for brain analysis include T1-weighted imaging, T2-weighted imaging, and fluid-attenuated inversion recovery (FLAIR). These sequences can provide complementary information. This information includes anatomical structures, edema extent, and tissue abnormalities [1].

With the development of deep learning, brain-tumor segmentation has progressively shifted from traditional methods, such as thresholding and region growing, to U-shaped networks (U-Net), convolutional neural networks (CNNs), and Transformer-based models [2]. Most multimodal datasets used for model training and most deep-learning methods assume that complete MRI sequences are available. In clinical practice, however, hardware limitations, patient motion, and prolonged scan times may result in the absence of one or more modalities [3]. For example, the absence of T2-weighted or FLAIR images may impair the delineation of edema and the whole-tumor region [4]. Therefore, developing robust tumor-segmentation methods for incomplete-modality settings is considerably important. Deep-learning approaches for handling missing modalities can generally be categorized into three groups: (1) synthesizing missing modalities using generative adversarial networks (GANs), hypernetworks, or diffusion models [5]; (2) transferring knowledge from complete modalities through shared latent-space learning, feature disentanglement, knowledge distillation, and mutual-information-based learning; and (3) directly fusing available modalities using modality masking, Transformers, feature interaction, or Mamba-based models [6]. Although these methods have achieved promising results on the Multimodal Brain Tumor Image Segmentation (BraTS) dataset, they remain limited by structural distortions in synthesized images, unstable performance under severe modality loss, insufficient cross-center generalization, high computational costs, and limited clinical validation.

The study of multimodal MRI brain tumor segmentation in the absence of modality has important clinical and scientific value. On the one hand, when the sequence is incomplete due to limited scanning time, patient movement, equipment conditions or contraindications to contrast agents, the stable segmentation results can provide support for brain tumor auxiliary diagnosis, treatment planning and efficacy evaluation, reduce repeated examinations and improve the efficiency of medical resource utilization; On the other hand, a systematic review of modal generation, knowledge transfer and adaptive fusion methods can provide references for subsequent researchers in data set selection, model design and evaluation index setting, and promote the relevant methods from ideal complete modal experiments to real clinical scenarios and multicenter applications.

Based on the above, this review mainly focuses on multimodal brain MRI. This kind of MRI contains T1-weighted, T2-weighted, and FLAIR sequences. It also focuses on automated brain-tumor segmentation. In this review, the influence of missing modalities on tumor-segmentation performance is examined. After that, the feasibility of three kinds of methods is compared. These methods are generative compensation, knowledge transfer, and modality-adaptive fusion. In the end, this review summarizes their limitations and also discusses future research directions.

2. Multimodal MRI-based brain tumor segmentation: principles and methods

2.1. Imaging characteristics and complementary roles of multimodal brain MRI

The multimodal nature of MRI arises from the use of different pulse sequences; each of them provides information about specific tissue characteristics. T1-weighted imaging emphasizes fat signal and primarily depicts the anatomical structure of the brain; T2-weighted imaging is sensitive to water and fat, with tissues containing higher water content generally appearing hyperintense; therefore it can reveal tumors, necrosis, and edema; Fluid-attenuated inversion recovery (FLAIR)

DOI: 10.54254/2755-2721/2026.AST37396

suppresses the high signal of cerebrospinal fluid (CSF), thereby improving the visualization of abnormalities in the brain parenchyma and peritumoral edema. The combined use of these three sequences enables clearer delineation of the whole tumor and tumor core, supporting the diagnosis of brain tumors [4].

Figure 1 presents visualizations of the three sequences in a representative brain lesion; their complementary information can provide clinicians with a more comprehensive basis for assessment.

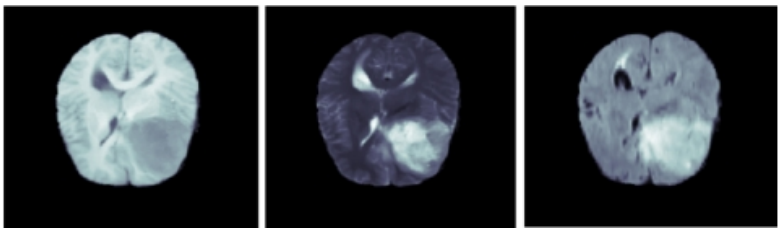

Figure 1. Comparison of different MRI sequences. Left to right, T1, T2, and FLAIR

Each MRI modality reveals different characteristics of the soft tissue of the human brain [7, 8].

2.2. Image quality issues and preprocessing methods

The quality of an MRI image is the premise of correct diagnosis. Usually, the reduction of image quality is due to artifacts and noise. Noise is usually caused by wrong system input, which is observed as fuzzy areas, random changes and artifacts. At the same time, it will also blur the anatomic boundaries. Therefore, in the process of denoising, it is necessary to reduce the intensity of noise and ensure that the edges do not show artifacts [9]. Also, motion is a challenge in MRI, as patients' head motion may lead to artifacts in k-space. Thus, deep learning is proposed to improve non pathological differences, which improves the image quality through image registration, resampling and data enhancement [3]. However, this can only repair the quality problem of the existing image, but not restore the pathological information in the completely missing sequence.

2.3. Evolution and evaluation of brain tumor segmentation methods

Traditional brain-tumor segmentation usually includes several steps. The first step is image preprocessing. After that, feature extraction and classification are performed. At last, the segmentation result is produced. This kind of method depends a lot on handcrafted features. It is also sensitive to noise, intensity inhomogeneity, and tumor boundaries that are complex. In contrast, deep-learning-based approaches are different. They can automatically learn hierarchical features from the input images, and then they generate segmentation labels [1]. U-shaped networks (U-Net) use an encoder–decoder architecture. In this architecture, skip connections are included, so the network can achieve precise localization. Three-dimensional convolutional neural networks (3D CNNs) make use of spatial continuity across adjacent slices. Cascaded networks and multiscale networks can help to identify different tumor regions in a progressive way. Transformer-based models can also provide help. They assist in selecting salient features and in modeling long-range dependencies.

For evaluation, the BraTS dataset is commonly used by many researchers. In this dataset, the whole tumor, the tumor core, and the enhancing tumor are regarded as the primary segmentation targets. When the performance needs to be assessed, several metrics are typically used. These metrics include the Dice similarity coefficient (Dice), the 95th-percentile Hausdorff distance

DOI: 10.54254/2755-2721/2026.AST37396

(HD95), sensitivity, and specificity. Besides these metrics, some other factors are also taken into consideration. They are the number of model parameters, the computational cost, and the inference time [2].

2.4. Performance and limitations of segmentation models under complete-modality settings

In the task of brain tumor segmentation, the right multimodal MRI can clearly distinguish the cancerous regions from the healthy tissues. That is to say, it is able to help separate the cancerous areas from the healthy areas in a clear way. 3D U-Net, attention network, multi-scale model can make full use of complementary information and obtain better segmentation results. However, due to the influence of many factors in practical applications, if the relevant sequence is missing, the input information and feature distribution will change, and the model may not recognize the tumor subregion closely related to the mode, which will affect image segmentation and imaging. At the same time, 3D networks also have unstable problems such as a large amount of calculation and weak interpretability, so it's necessary to develop a more stable multimodal segmentation algorithm [6].

3. Impact of missing modalities on brain tumor segmentation and deep learning-based mitigation strategies

3.1. Causes of missing modalities and impact on segmentation

Clinically, prolonged scan times, patient motion, variations in imaging protocols across hospitals, and allergies or contraindications to contrast agents may lead to incomplete acquisition of one or more MRI sequences. Missing modality involves a single sequence or multiple sequences. Due to the difference of MRI sequences provide complementary information, the absence of T1-weighted imaging may impair anatomical localization, whereas the absence of T2-weighted or FLAIR imaging may compromise the assessment of edema and whole-tumor extent. Accordingly, the loss of any modality can substantially affect image segmentation and lesion localization. Segmentation performance should therefore be compared under different missing-modality conditions using the Dice and HD95 for the whole tumor and tumor core, thereby clarifying the relationships among the type and number of missing modalities and the resulting performance degradation.

3.2. Deep learning-based strategies for missing modalities

3.2.1. Modality synthesis and generative compensation

Before the segmentation is performed, the missing modalities can be synthesized. They are synthesized from the MRI sequences that are still available. These techniques have evolved over time. In the early stage, they were based on image-to-image mapping and generative adversarial networks (GANs). Later, they developed into unified synthesis frameworks, hypernetworks, and diffusion models. In this way, different combinations of input and output modalities can be supported. This strategy gives an intuitive form of compensation. However, it may also bring some problems. For example, some lesions may be omitted, or some spurious structures may appear. Therefore, when evaluation is carried out, two kinds of metrics should be considered. One kind is the synthesis-quality metrics, such as peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). The other kind is the metrics for the downstream segmentation task, including the Dice and the HD95. This means that the evaluation should not only look at the quality

DOI: 10.54254/2755-2721/2026.AST37396

of the synthesized image itself. It should also examine how the synthesis result affects the later segmentation performance [8].

3.2.2. Shared representation learning and knowledge transfer

Shared latent-space learning captures information common across MRI modalities, while feature disentanglement further separates shared representations from modality-specific features. Knowledge distillation uses a teacher network trained with complete modalities to guide a student network operating on incomplete modalities. Methods enhanced by Hölder divergence and mutual information can constrain discrepancies between predictions and strengthen knowledge transfer. However, they may still depend on complete training data and remain susceptible to teacher bias and increased training complexity [10].

3.2.3. Modality-adaptive fusion and global context modeling

The Modality-Masked Fusion Transformer (M²FTrans) makes use of fusion tokens. It also uses masked self-attention. In this way, it can disregard the inputs that are missing. Pre–Post Interaction Learning (PPIL) is another method. It strengthens the associations between different modalities. This is done through interactions between the representations before fusion and the representations after fusion. M²SegMamba combines Mamba with masked autoencoding. It is designed to model long-range dependencies. This modeling is carried out under conditions where some modalities are missing and only a few samples are available. These methods have an advantage. They can directly fuse the modalities that are available. They do not need to explicitly synthesize the modalities that are missing. However, their compensation mechanisms are not so intuitive. In addition, their three-dimensional implementations are still computationally expensive [6].

3.3. Current limitations and future directions

Nowadays, in order to deal with the lack of MRI modality, existing research mostly adopts the strategy of manually deleting modality, but it's difficult to cover the problem of low-quality clinical data. Shared representation and knowledge transfer methods do not need to generate images, but usually rely on complete modal data or teacher models. At the same time, the model generated based on deep learning may introduce false structures, making it difficult to ensure accuracy, and it is also easily affected by hospital environmental equipment. In addition, the cost of 3D network and diffusion models is high, and most studies lack external experience and uncertainty analysis, which is also difficult to apply directly. Therefore, in the future, we should build a real missing data set, develop a unified model that can adapt to any combination of inputs, and estimate and control errors with data consistency, anatomical constraints and uncertainty. We should also combine segmentation with tumor grading, treatment planning and prognosis prediction, and turn to clinical effectiveness evaluation [11, 12].

4. Conclusion

This review investigates multimodal MRI-based brain-tumor segmentation under missing-modality conditions. It describes the complementary roles of T1-weighted, T2-weighted, and FLAIR sequences and compares generative compensation, knowledge transfer, and modality-adaptive fusion strategies. Complete-modality inputs provide an important foundation for stable segmentation: the absence of T1-weighted imaging may impair anatomical localization, whereas the

DOI: 10.54254/2755-2721/2026.AST37396

absence of T2-weighted or FLAIR imaging may compromise the delineation of edema and the whole-tumor region. Existing deep-learning methods can partially compensate for missing information; however, generative models may produce hallucinated structures or omit lesions, knowledge-distillation methods generally rely on complete-modality teacher networks, and Transformer- and Mamba-based models remain computationally demanding and insufficiently interpretable. Consequently, no single deep-learning approach has yet fully resolved the problem of missing modalities.

This study is based primarily on a literature review and methodological comparison; it does not include quantitative experiments on a unified dataset or validation under real-world missing-modality scenarios and cross-institutional conditions. Future research should establish datasets containing genuinely incomplete modalities, develop lightweight models capable of handling arbitrary modality combinations, strengthen anatomical constraints and uncertainty estimation, and conduct multicenter clinical validation. Segmentation outputs should also be further integrated into tumor grading, treatment planning, and prognostic prediction.

References

[1] Soomro, T. A., et al. (2023). Image segmentation for MR brain tumor detection using machine learning: A review. *IEEE Reviews in Biomedical Engineering, 16*, 70–90. https://doi.org/10.1109/RBME.2022.3185292

[2] Alsmadi, H. S., Hamdan, H., Mustapha, N., & Manshor, N. (2025). 3D convolutional neural networks for brain tumor analysis in multimodal MRI: A systematic review. *IEEE Access, 13*, 140771–140790. https://doi.org/10.1109/ACCESS.2025.3597130

[3] Spieker, V., et al. (2024, February). Deep learning for retrospective motion correction in MRI: A comprehensive review. *IEEE Transactions on Medical Imaging, 43*(2), 846–859. https://doi.org/10.1109/TMI.2023.3323215

[4] Azad, R., Dehghanmanshadi, M., Khosravi, N., Cohen-Adad, J., & Merhof, D. (2025, April). Addressing missing modality challenges in MRI images: A comprehensive review. *Computational Visual Media, 11*(2), 241–268. https://doi.org/10.26599/CVM.2025.9450399

[5] Yang, H., Sun, J., & Xu, Z. (2023, December). Learning unified hyper-network for multi-modal MR image synthesis and tumor segmentation with missing modalities. *IEEE Transactions on Medical Imaging, 42*(12), 3678–3689. https://doi.org/10.1109/TMI.2023.3301934

[6] Zhang, X., Bahri, A., Desrosiers, C., Liu, H., & Bao, F. (2026, March). M^{2}SegMamba: Mamba-based incomplete multimodal learning for brain tumor segmentation with few samples. *IEEE Journal of Biomedical and Health Informatics, 30*(3), 2455–2465. https://doi.org/10.1109/JBHI.2025.3600652

[7] Menze, B. H., et al. (2015). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging, 34*(10), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

[8] Ashton, E. A., & Kim, S. H. (2004). Evaluation of reproducibility for perfusion assessment of tumors in MRI. *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, 824–827. https://doi.org/10.1109/ISBI.2004.1398665

[9] Mishro, P. K., Agrawal, S., Panda, R., & Abraham, A. (2022). A survey on state-of-the-art denoising techniques for brain magnetic resonance images. *IEEE Reviews in Biomedical Engineering, 15*, 184–199. https://doi.org/10.1109/RBME.2021.3055556

[10] Zhang, Y., Peng, C., Wang, Q., Song, D., Li, K., & Kevin Zhou, S. (2025). Unified multi-modal image synthesis for missing modality imputation. *IEEE Transactions on Medical Imaging, 44*(1), 4–18. https://doi.org/10.1109/TMI.2024.3424785

[11] Cheng, R., Qiu, X., Li, M., Zhang, Y., Yu, F. R., & Li, C. (2026). Robust brain tumor segmentation with incomplete MRI modalities using Hölder divergence and mutual information-enhanced knowledge transfer. *IEEE/CAA Journal of Automatica Sinica, 13*(4), 939–954. https://doi.org/10.1109/JAS.2025.125609

[12] Shi, J., Yu, L., Cheng, Q., Yang, X., Cheng, K.-T., & Yan, Z. (2024). MFTrans: Modality-masked fusion transformer for incomplete multi-modality brain tumor segmentation. *IEEE Journal of Biomedical and Health Informatics, 28*(1), 379–390. https://doi.org/10.1109/JBHI.2023.3326151