Computerized Medical Imaging and Graphics 104 (2023) 102167

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

A literature survey of MR-based brain tumor segmentation with missing modalities

Tongxue Zhou [a], Su Ruan [b], Haigen Hu [c,d,*]

[a] School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China
[b] Université de Rouen Normandie, LITIS - QuantIF, Rouen 76183, France
[c] College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China
[d] Key Laboratory of Visual Media Intelligent Processing Technology of Zhejiang Province, Hangzhou 310023, China

ARTICLE INFO

Keywords:
Brain tumor segmentation
Missing modalities
Multi-modalities
Magnetic Resonance Imaging
Deep learning

ABSTRACT

Multimodal MR brain tumor segmentation is one of the hottest issues in the community of medical image processing. However, acquiring the complete set of MR modalities is not always possible in clinical practice, due to the acquisition protocols, image corruption, scanner availability, scanning cost or allergies to certain contrast materials. The missing information can cause some restraints to brain tumor diagnosis, monitoring, treatment planning and prognosis. Thus, it is highly desirable to develop brain tumor segmentation methods to address the missing modalities problem. Based on the recent advancements, in this review, we provide a detailed analysis of the missing modality issue in MR-based brain tumor segmentation. First, we briefly introduce the biomedical background concerning brain tumor, MR imaging techniques, and the current challenges in brain tumor segmentation. Then, we provide a taxonomy of the state-of-the-art methods with five categories, namely, image synthesis-based method, latent feature space-based model, multi-source correlation-based method, knowledge distillation-based method, and domain adaptation-based method. In addition, the principles, architectures, benefits and limitations are elaborated in each method. Following that, the corresponding datasets and widely used evaluation metrics are described. Finally, we analyze the current challenges and provide a prospect for future development trends. This review aims to provide readers with a thorough knowledge of the recent contributions in the field of brain tumor segmentation with missing modalities and suggest potential future directions.

1. Introduction

A brain tumor is a growing abnormal cell in the brain or central spine canal, which is one of the most life-threatening cancer. In the United States, the National Brain Tumor Society estimates that 13,000 patients die from brain tumor every year, and 29,000 patients undergo primary brain tumors (Singh et al., 2012; Saouli et al., 2018). The median overall survival time remains only 15–17 months after diagnosis and standard treatment (Goel et al., 2021; Farmanfarma et al., 2019; Rouse et al., 2015). In addition, brain tumors can also lead to a significant financial impact with a mean cost of $2788 ± 3719 for the treatments, which depends on the various imaging techniques (Das et al., 2022). Therefore, the accurate diagnosis of brain tumors plays an important role in planning surgeries and treatments, which may potentially prolong a patient's survival time.

Magnetic resonance imaging (MRI) is a widely used medical imaging technique in radiology. It can form pictures of the anatomy and the physiological processes of the body. MRI can offer a delicate aspect of the brain, vascular anatomy and spinal cord. Besides, it can visualize the brain structure in all planes (axial, sagittal and coronal). MRI is considered as one of the most promising and fastest ways to diagnose brain tumors, because of its soft-tissue contrast and extensive availability. A variety of MRI modalities can be obtained by changing the time of excitation during image acquisition (Rao and Karunakara, 2021). In a clinical scenario, the widely used MR modalities are Fluid-attenuated inversion recovery (FLAIR), T2-weighted images (T2), T1-weighted images (T1), and T1-weighted images with contrast enhancement (T1c), These different modalities can provide complementary information to help distinguish brain tumors from normal brain tissues (Agravat and Raval, 2021; Fang and Wang, 2022). However, we often confront the absence of certain modalities due to the different acquisition protocols, image corruption, scanner availability, scanning cost, or allergies to certain contrast materials. This poses challenges to both physicians and automated diagnosis systems since complementary information provided by the missing modalities is missing.

* Corresponding author.
E-mail address: hghu@zjut.edu.cn (H. Hu).

https://doi.org/10.1016/j.compmedimag.2022.102167
Received 25 July 2022; Received in revised form 1 November 2022; Accepted 22 December 2022
Available online 27 December 2022

Following the breakthroughs in brain tumor segmentation in the past decades, many noteworthy surveys have been published discussing the progressions in brain tumor segmentation (Das et al., 2022; Wadhwa et al., 2019; Agravat and Raval, 2021). Nevertheless, there is still a lack of dedicated surveys focusing on the missing modalities issue in MR-based brain tumor segmentation. In recent years, a lot of research works and various technologies have been presented to solve the missing modality issue in MR-based brain tumor segmentation. In this paper, we provide a technical review for MR-based brain tumor segmentation with missing modalities. Particularly, we define the problem as follows: full MR modalities are available during training, while one or more modalities can be absent during tests. We aim to cover the most influential works in this field published until 1st March 2022. The objective of this review is to help the readers to know the recent contributions and explore the future research directions. Our contributions can be divided three-fold:

(1) We provide a technical overview of the recent methodologies and categorize these publications into five groups: image synthesis-based method, latent feature space-based model, multi-source correlation-based method, knowledge distillation-based method, and domain adaptation-based method. We give an in-depth analysis of these methods across the principles, architecture, benefits and limitations.

(2) We introduce the widely used brain tumor segmentation datasets and the corresponding evaluation metrics. Furthermore, we provide concise tabular descriptions of the reviewed methods and compare their performance in the case of missing modalities.

(3) We discuss the current challenges and provide significant information to the readers about the future development directions in the field of MR brain tumor segmentation with missing modalities.

The remainder of this paper is structured as follows. In Section 2, we first provide a biomedical background regarding the brain tumor, MR imaging techniques, and the current challenges in brain tumor segmentation. Then, in Section 3, we offer an in-depth analysis of the recent state-of-the-art methods across principles, architecture, benefits and limitations. Following that, in Section 4, we introduce the prominent brain tumor segmentation datasets, along with a description of the widely used evaluation metrics in brain tumor segmentation. In addition, we provide concise tabular descriptions of the reviewed methods and compare their performance in the case that one or more modalities are absent. Finally, we discuss current major challenges and future development trends in Section 5, and give a conclusion of this paper in Section 6.

2. Biomedical background review

2.1. Brain tumor

Brain tumors are cancerous or non-cancerous growth of abnormal cells in the brain Wadhwa et al. (2019), Bahadure et al. (2017). Brain tumors are one of the most aggressive cancers, which seriously threaten the life and health of patients, and account for substantial morbidity and mortality. World Health Organization (WHO) assigned grades for brain tumors, ranging from Grade I (least malignant) to Grade IV (most malignant). In general, grade I and grade II are benign brain tumors (low-grade); grade III and grade IV are malignant brain tumors (high-grade) (Tiwari et al., 2020). Benign tumors cover 29.7% while malignant tumors take up about 70.3% (Miller et al., 2021). Considering malignant tumors, glioblastoma tumors are the most common ones (48.6%), which advance rapidly and can lead to death quickly, while other classes combined cover 51.4% of all the cases. Fig. 1 presents a detailed distribution of brain tumors. Brain tumors can also be classified as primary tumors and secondary tumors (brain metastasis tumors). The former one starts within the brain, such as meninges, brain cells, nerve cells, and glands. The latter begins as cancer elsewhere and spreads to the brain. For example, lung cancer, skin cancer and kidney cancer can metastasize to the brain. Most brain tumors are secondary forms

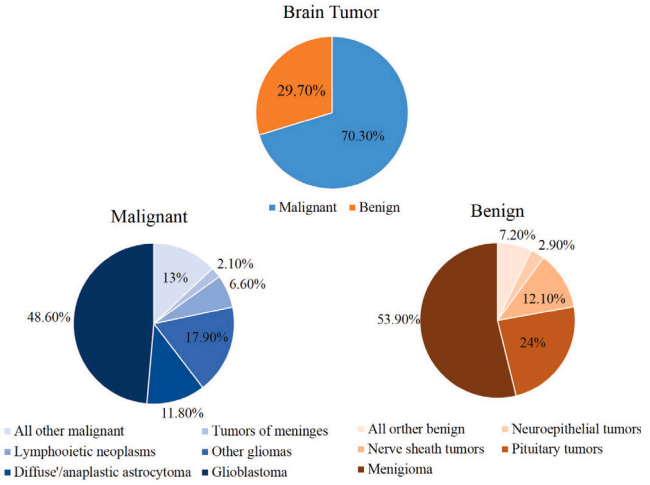

Fig. 1. Distribution of brain and other central nervous system (CNS) tumors by behavior and major histology type (Miller et al., 2021).

of tumors. It is estimated that 20%–40% of all patients diagnosed with a primary tumor will develop a secondary tumor in the brain. 10%–20% of patients have only one secondary brain tumor, but the vast majority of patients (80+%) typically have more than one secondary tumor. Brain metastases are considered malignant forms of cancer.[1] The 5-year survival rate indicates what percent of people live at least 5 years after the tumor is found. It is reported that the 5-year survival rate of brain tumors is only about 36% (Das et al., 2022; Rouse et al., 2015). Therefore, early diagnosis of brain tumors is a critical issue in real clinical practice.

2.2. Magnetic resonance imaging

Magnetic Resonance Imaging (MRI) is a commonly used medical imaging technique in radiology for medical diagnosis and treatment monitoring. The MRI scanner is essentially a giant magnet, and the strength of the magnet is measured by Tesla (T). Most MRI scanners used in hospitals and medical research clinics are 1.5 or 3 T. A 3 T MRI scanner is around 50,000 times stronger than the earth's magnetic field which is around 0.00006 T. To obtain an MR image, a patient is placed inside a large magnet. The human body contains a lot of water and hydrocarbons, which contain many hydrogen atoms. The hydrogen nuclei (protons) act as tiny magnets. Fig. 2 illustrates the process of obtaining signals from protons in the body. When outside of an external magnetic field, protons will point randomly in any direction, thus yielding a null magnetization moment. When it is placed in a strong magnetic field, B0, protons exhibit a weak tendency to align with the direction of B0. When the Radio Frequency (RF) pulse is applied, B1, perpendicular to B0, the protons become deflected by 90°. When the RF signal stops, the defected protons return to equilibrium such that it is parallel again to B0. The recovery process is called the relaxation process, which emits stored energy and can be received by a coil as a signal. Finally, an MR image is obtained by analyzing this signal and the time duration of the return.

MRI has a wide range of applications in medical diagnoses, such as in neuroimaging and Nichols et al. (2017), cardiovascular (Petersen et al., 2017) and systemic muscle diseases (Schmidt et al., 2007). MRI is among the most preferred imaging technique for diagnosing brain tumors due to the following reasons. First, it can provide more precise and detailed images of soft tissues than Computed Tomography

[1] https://www.brainlab.org/

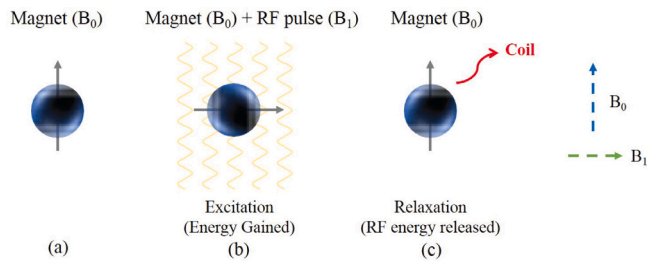

Fig. 2. Conceptual diagram of obtaining signals from protons in the body.

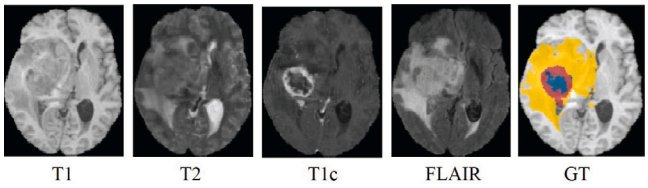

Fig. 3. Four commonly used MRI modalities with the corresponding sub-tumor regions from the BraTS 2018 dataset. denotes edema, denotes enhancing tumor, and denotes necrotic and non-enhancing tumor.

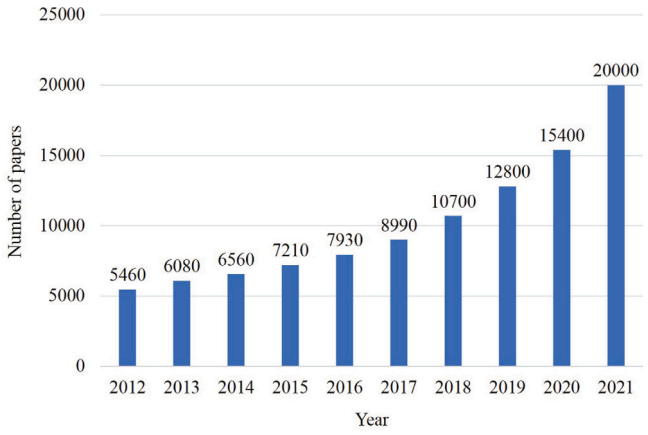

Fig. 4. The growth of research papers with the keyword 'brain tumor segmentation' in the last decade, surveyed by Google Scholar.

(CT) (Das et al., 2022). Second, MRI is a non-invasive imaging technique, since the RF pulses used in MRI do not cause ionization and have no harmful effects of ionizing radiation. Third, MRI, as a multimodal imaging technique, can express various contrasts between different tissues according to the difference between transverse relaxation time and longitudinal relaxation time. In clinical practice, the most commonly used MR modalities are Fluid-attenuated inversion recovery (FLAIR), T2-weighted images (T2), T1-weighted images (T1), and T1-weighted images with contrast enhancement (T1c) (Menze et al., 2014; Zhang et al., 2021a). Different modalities carry specific image information, and these multi-modalities can provide complementary information to distinguish brain tumors from normal brain tissues. Fig. 3 presents these four MR modalities with the corresponding three sub-tumor regions: edema, enhancing tumor, necrotic and non-enhancing tumor. For example, FLAIR and T2 modalities can highlight the brain edema regions (yellow region). T1 and T1c modalities can give a better visualization of tumor core regions (red and blue regions). Thus, the joint analysis of the multimodal MR images can aid doctors to judge the pathological and histomorphological changes to optimize the segmentation results. Additionally, it is helpful to extract brain tumors and customize the following treatment planning (Zhang et al., 2021b; Zhou et al., 2022). Finally, MRI is a multi-planar imaging technique, which can produce axial, sagittal and coronal images, while it is impossible with radiography and CT.

2.3. Brain tumor segmentation

MR-based brain tumor segmentation is an essential basic step in clinical practice, which has many applications in neurology such as quantitative analysis, operational planning, and functional imaging (Ben Rabeh et al., 2017; Işın et al., 2016). Fig. 4 presents the growth of research papers with the keyword 'brain tumor segmentation' in the last decade, surveyed by Google Scholar. We can observe that the number of papers is increasing by the year, and doubling from 2018 to 2021, implying that brain tumor segmentation is attracting more and more attention in recent years.

MRI can give a better description of the structures of the brain. However, MR brain tumor segmentation is always a tough task in the real clinical scenario due to the following facts. First, considering the properties of brain tumors: the brain anatomy structure and brain tumor characteristics vary from patient to patient. Besides, brain tumors are very heterogeneous, the intensity value of brain tumors may overlap with the intensity value of healthy brain tissue (Zhou et al., 2019b; Ghaffari et al., 2019). Second, the limitations of MRI pose challenges for brain tumor segmentation: (1) The low contrast from MRI results in obscure tumor boundaries. (2) MR images contain various artifacts, such as irregularities, pixel variations, and inhomogeneity due to the different scanners (Cong et al., 2016). (3) It is challenging to design an effective approach to fuse different MR modalities to maximum utilize the complementary information from multi-source (Zhou et al., 2020). (4) The lack of one or more MR modalities can heavily decrease the segmentation accuracy, compared with the complete modalities (Wadhwa et al., 2019; Sharma and Shukla, 2021). In this paper, we focus on the missing modalities issue. Third, a very important technical limitation for deep learning-based brain tumor segmentation is the scarcity of available datasets. Since deep learning network requires extensive training data, it makes the dataset acquisition more demanding. And different hospitals have different scanners and image protocols, which complicates standardization and data quality. Besides, the tumor annotation is manually delineated by experts, which leads to some intra- and inter-rater variability. Therefore, there is no certain truth but only a gold standard (Renard et al., 2020). Lastly, the brain tumor segmentation process usually consumes significant time and computing resources. Therefore, speeding up the segmentation process is crucial to obtain accurate and efficient segmentation results. For example, Xiong et al. (2021) proposed an FPGA-based brain tumor segmentation accelerator to increase the segmentation speed and reduce power consumption.

Deep learning models can bring certain advantages for brain tumor segmentation in real clinical scenarios. First, it can automatically learn complex characteristics from images which are difficult to specify manually. Second, it can accelerate the process of brain tumor segmentation, and guide clinical decision-making to alleviate the burden on doctors. However, using deep learning models still requires clinical validation due to the black box nature of deep learning. Since once the network has been trained, it behaves like a black box, providing predictions which are not directly interpretable. This issue makes the model unpredictable, intractable for model verification, and ultimately untrustworthy (Chen et al., 2020). Thus, interpretability could be a significant research direction to understand the features learned by deep learning models.

3. State-of-the-art segmentation methods

In recent years, a large number of research works have been proposed to address MR-based brain tumor segmentation with missing modalities. We divide them into five categories, namely, image

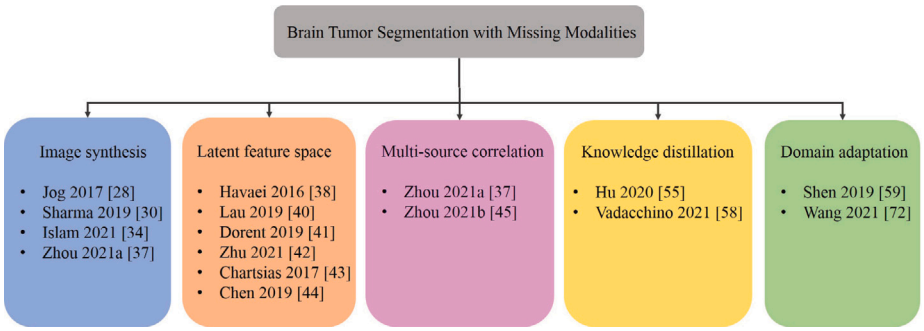

Fig. 5. The classification tree of state-of-the-art methods in the field of MR brain tumor segmentation with missing modalities.

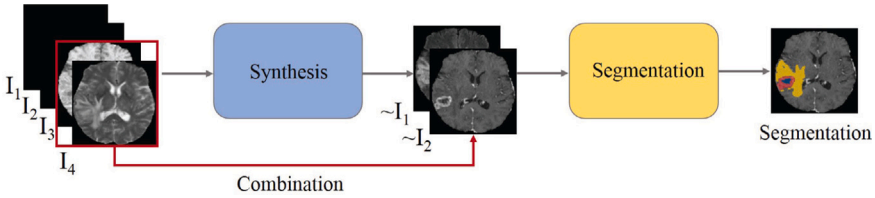

Fig. 6. Overview of the image synthesis-based method.

synthesis-based method, latent feature space-based model, multi-source correlation-based method, knowledge distillation-based method, and domain adaptation-based method. The classification tree of these methods is depicted in Fig. 5. In the following, we will elaborate on these methods regarding the principles, architectures, benefits and limitations.

3.1. Image synthesis-based method

An intuitive solution to address the problem of missing modalities is to adopt image synthesis techniques to compensate for the missing data. Many works have been proposed by first synthesizing the missing modalities, and then segmenting brain tumors using existing and synthesized modalities. The overview of this strategy is depicted in Fig. 6. First, a specific image synthesis method is used to generate the missing modalities. Then, the synthesized and existing images are combined to perform brain tumor segmentation.

Jog et al. (2017) proposed a supervised random forest-based (Breiman, 1996) image synthesis approach called REPLICA (Regression Ensembles with Patch Learning for Image Contrast Agreement). REPLICA first extracts multi-scale features and then performs a multi-resolution analysis via random forests. Finally, the learned nonlinear mapping is applied to test images to synthesize the missing modality. Experimental results indicate that REPLICA can synthesize both T2-weighted images of the full head and FLAIR images, and can be beneficial as a pre-processing step for subsequent image segmentation.

Sharma and Hamarneh (2019) proposed a Multi-Modal Generative Adversarial Network (MM-GAN) for the task of synthesizing missing MR modalities. The generator is based on U-Net (Ronneberger et al., 2015), and the discriminator is adopted from PatchGAN architecture (Isola et al., 2017). The network is a multi-input and multi-output architecture, generalizing to any combination of available and missing modalities. In addition, the model is trained via curriculum learning (Bengio et al., 2009), where easier cases are shown at the beginning, and followed by increasingly difficult cases. Experimental results demonstrate the applicability of the proposed method in simultaneously synthesizing all missing modalities in any possible missing modalities situations on two brain MRI datasets.

Islam et al. (2021) proposed to first synthesize the missing MR modalities using a fully convolutional network (FCN) (Long et al., 2015) based generator, and a modified discriminator via a conditional generative adversarial network (cGAN) (Mirza and Osindero, 2014). Finally, the network achieves brain tumor segmentation by combining the available and the synthesized MR modalities. Experimental results validate the effectiveness of the proposed method in both image synthesis and brain tumor segmentation.

However, the main drawback of this strategy is that it is computationally cumbersome, especially in the case that many modalities (e.g., three modalities) are absent. To overcome this problem, Zhou et al. (2021a) proposed to first generate an average image of the missing modalities, considering two facts: (1) The averaging operation is simple to realize. And one decoder is sufficient to cope with any number of missing modalities; (2) Since the features extracted from different MR modalities are correlated for the same patient. Thus, the averaging operation can extract the overall feature information from multi-modalities, providing discriminative feature information for the subsequent segmentation network. Then, the generated modality is combined with the existing ones to perform the brain tumor segmentation. Experimental results verify the advantage of the proposed method.

3.2. Latent feature space-based method

Many researchers attempted to retrieve the missing information by exploiting the multimodal latent feature space. The overview of the method is depicted in Fig. 7. First, individual encoders are applied to extract the modality-specific feature representations. Subsequently, these feature representations are used to learn a shared latent feature representation. Finally, a decoder is used to achieve brain tumor segmentation.

The early segmentation network designed for missing modalities is from HeMIS (Havaei et al., 2016). The network architecture is based on a convolutional neural network (CNN) (LeCun et al., 1998), which consists of three main parts: back-end layer, abstract layer and front-end layer. First, each modality is processed by independent convolutional layers in the back-end layer. Then, in the abstract layer, the obtained feature maps from all available modalities are merged by computing the mean and the variance. Finally, the mean and variance feature maps are concatenated and fed into the front-end layer to obtain network output. Experiments evaluated on two challenge datasets demonstrate the promising segmentation results of the proposed method.

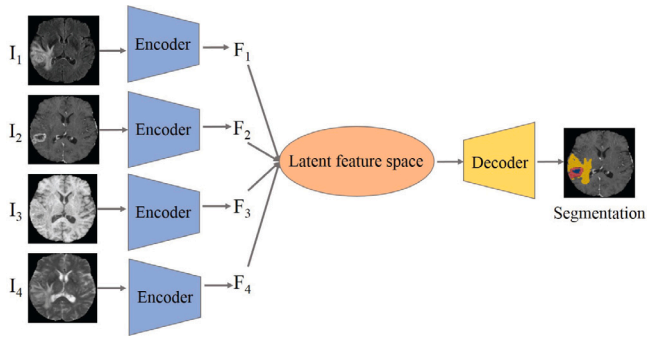

Fig. 7. Overview of the latent feature space-based method. F_1, F_2, F_3, and F_4 denote the latent feature representations learned from the input modalities.

Similar to this work, Lau et al. (2019) proposed a unified representation network (URN) to map the multi-modalities into a unified representation for segmentation. The baseline is U-Net, consisting of modality-specific encoders, a fusion module and a segmentation decoder. The fusion module combines the outputs of all individual encoders into a unified representation by only calculating the mean. Furthermore, the authors propose to synthesize multi-modal images by designing additional decoders for each modality. Experimental results demonstrate the robustness of the proposed method on missing modalities.

Instead of using arithmetic operations to fuse feature maps, Dorent et al. (2019) proposed a network named U-HVED, which applies multimodal variational auto-encoders to cope with the absence of modalities. Encoders are first used to extract the context information. At each level, the means and the variances of the modalities are combined via the product of Gaussian. Besides, skip connections are applied to capture more rich features. Then, decoders are used to decode the multi-scale latent representation to obtain the final segmentation map as well as generate the missing modalities. Experimental results validate that the proposed model can outperform the current state-of-the-art method HeMIS for dealing with missing modalities on the BraTS 2018 dataset.

Inspired by the aforementioned work (Dorent et al., 2019), Zhu et al. (2021) proposed to extend the multimodal variational auto-encoders to residual form, while this work can only tackle missing one modality situation.

Another solution to explore the latent feature space is to minimize the L1 or L2 distance of features from different modalities to achieve the shared latent feature representation. Chartsias et al. (2017) proposed a multi-input multi-output fully convolutional neural network model for MRI synthesis. The code is publicly available.[2] It is composed of three major parts: encoder, latent representation fusion and decoder. These modules are trained end-to-end by using a cost function that encourages representations to be modality-invariant, whilst the individual reconstruction error is kept low. Experimental results demonstrate the benefits of learning a shared latent feature space, resulting in a statistically significant improvement over the current state-of-the-art method REPLICA.

Since different MRI modalities have different intensity distributions, using arithmetic operations, such as mean and variance or simply encouraging the features from different modalities to be close under L1 or L2 distance, could not guarantee the network can learn a shared latent representation. Recently, Chen et al. (2019) introduced feature disentanglement to address the missing modality problem. The network first disentangles multimodal features into modality-specific appearance code and modality-invariant content code. The modality-specific appearance code is special for each modality. The modality-invariant content code shares multimodal information for segmentation. Then, the content code of each modality is fused into a shared feature representation via a learning-based fusion strategy. To enhance its modality-invariance, the shared feature representation is required to reconstruct any modality given the corresponding appearance code, even with missing modalities. Experimental results demonstrate that the proposed method can achieve competitive performance to the state-of-the-art methods for both full modality and various missing modalities situations.

[2] https://github.com/agis85/multimodal_brain_synthesis

3.3. Multi-source correlation-based method

Since the same tumor region can be presented in different MR modalities, Zhou et al. (2021b) found that there was a strong correlation in the intensity distribution between each pair of MR modalities in the tumor regions (Lapuyade-Lahorgue et al., 2017). To this end, they assume that a strong correlation also exists in the latent feature representation between different MR modalities. The overview of the method is shown in Fig. 8.

The proposed network first takes the four modalities as inputs in each encoder. The independent encoders can not only learn modality-specific feature representation but also can avoid false adaptation between modalities. To take into account the strong correlation between multi-modalities, a correlation model is developed to discover the correlation between modalities. Then, the correlation representations across modalities are fused via an attention mechanism (Hu et al., 2018; Roy et al., 2018) to emphasize the most discriminative representation for segmentation. Finally, the fused latent representation is decoded to form the final segmentation result. At the same time, the four reconstruction decoders are applied to provide more supervision to enhance the segmentation. Experimental results demonstrate the effectiveness of the proposed method on two brain tumor segmentation datasets.

However, in the above-mentioned method, the authors only replaced missing modalities with the existing ones, which leads to unsatisfactory results when more modalities are missing. Also, it did not take into account if the estimated correlated feature representation and the original feature representation are similar. Thus, to solve these issues and further enhance the segmentation performance, Zhou et al. (2021a) proposed an improved version, the overview of the method is depicted in Fig. 9. It first generates an average of the missing modalities via an encoder–decoder-based generator, forming new complete modalities. Then, a correlation model is presented to explore the multi-source correlation on the new complete set of modalities. In addition, the missing feature representations can be retrieved. Finally, a decoder is applied to obtain the final brain tumor segmentation. Experimental results validate the effectiveness of the proposed method compared with the previous and other state-of-the-art methods.

3.4. Knowledge distillation-based method

Knowledge distillation (Hinton et al., 2015) is a way of transferring knowledge from a cumbersome model to a compact model, aiming to improve the performance of compact networks. It has received increasing attention from the research community in recent years. And it has a wide range of applications, such as image classification (Ba and Caruana, 2014; Urban et al., 2016), object detection (Li et al., 2017) and semantic segmentation (Xie et al., 2018; Liu et al., 2019). The overview of the method is depicted in Fig. 10. First. training the Teacher model till full convergence. And the loss function is a segmentation-based loss function. Then, training the Student model in coordination with the Teacher model. It is noted that the forward propagation is achieved on both Teacher and Student models and back-propagation is only done on the Student network. For the Student model, there are two loss functions: student segmentation loss function and distillation loss function.

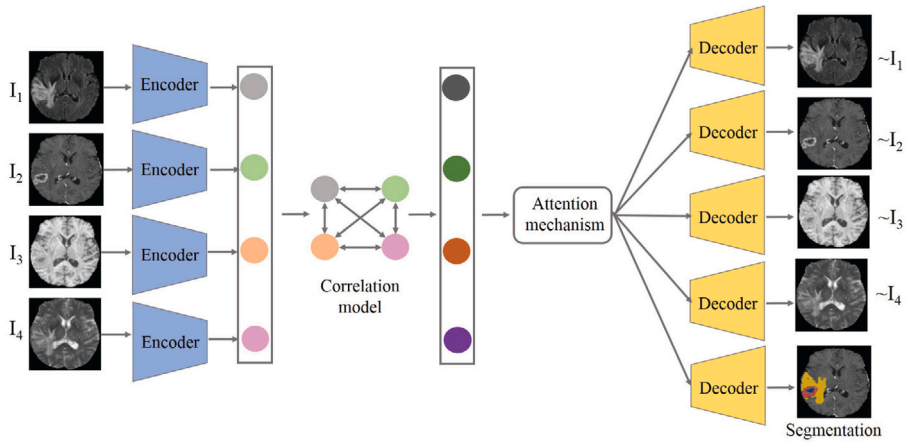

Fig. 8. Overview of the multi-source correlation-based method. The colorful circles denote the latent feature representations learned from the input modalities.

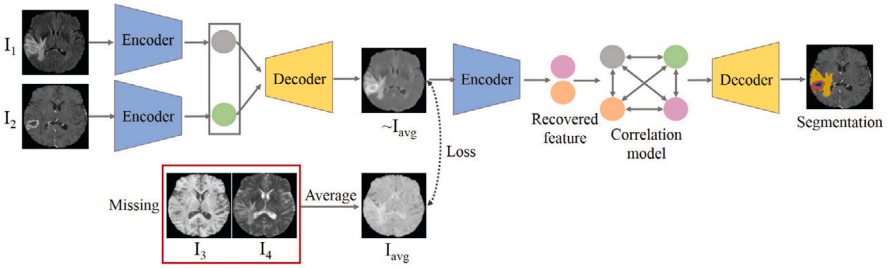

Fig. 9. Overview of the image generation and multi-source correlation-based method. The colorful circles denote the latent feature representations learned from the input modalities.

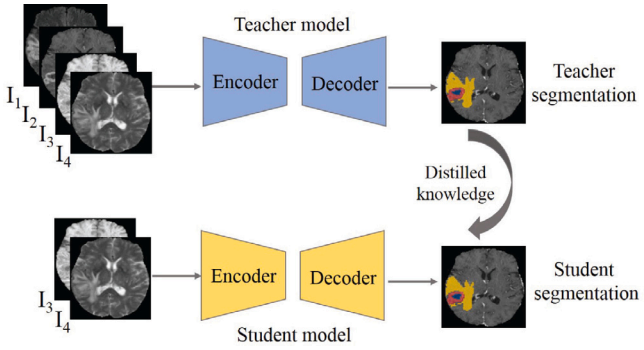

Fig. 10. Overview of the knowledge distillation-based method.

Hu et al. (2020b) proposed to use generalized knowledge distillation (Lopez-Paz et al., 2015; Vapnik et al., 2015) to transfer knowledge from a trained multi-modal network (teacher) to a mono-modal one (student). The proposed network is based on two same encoder–decoder networks, one for student and one for teacher. While the teacher learns from multiple input modalities whereas the student from only one. In addition, to constrain the latent representation between student and teacher to be similar, three different loss functions are used in student: the knowledge distillation loss, the Kullback–Leibler (KL) divergence-based loss, and the reference segmentation loss. Experimental results show that the proposed student network can effectively learn from the teacher and can outperform the baseline mono-modal network.

Vadacchino et al. (2021) introduced a Hierarchical Adversarial Distillation Network (HAD-Net) to overcome the absence of modalities in brain tumor segmentation. The code is publicly available.[3] The proposed network consists of three main components: the teacher network, the student network, and the hierarchical discriminator (HD). The teacher network takes all available MRI modalities as input, while the student network only uses the pre-contrast MRI modalities (exclude T1c) to train the model. The HD component is proposed to bridge the domain gap between the student and the teacher by mapping their segmentation as well as their multi-scale feature maps to a common space. However, similar to the work (Shen and Gao, 2019), this method specifically focuses on the missing T1c modality. Experimental results show the significant performance and improvements of the proposed method in segmenting enhancing tumors without T1ce over other methods.

3.5. Domain adaptation-based method

Domain adaptation solves a learning problem in a target domain by utilizing the training data in a different but related source domain. In recent years, the domain adaptation method has attracted much attention in many fields, such as point clouds (Achituve et al., 2021; Wu et al., 2019), object detection (Hsu et al., 2020; Inoue et al., 2018; Li et al., 2019a), image classification (Ren et al., 2018; Peng et al., 2019; Rietzler et al., 2019) and semantic segmentation (Li et al., 2019b; Vu et al., 2019; Choi et al., 2019; Liu et al., 2021). The overview of the method is depicted in Fig. 11. First, a model is trained on source domain data (full modalities). Meanwhile, another model is trained on target

[3] https://github.com/SaverioVad/HAD_Net

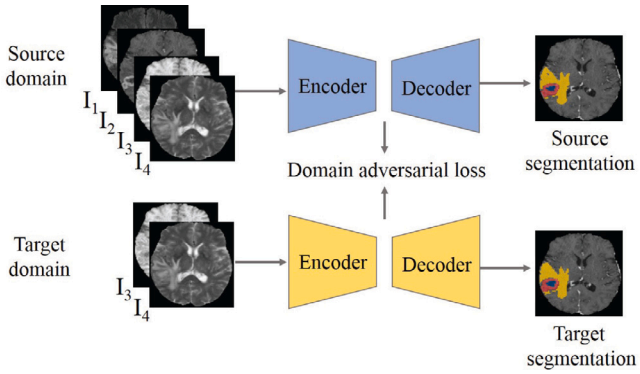

Fig. 11. Overview of the domain adaptation-based method.

domain data (missing modality). Subsequently, the domain adversarial loss function is adopted to minimize the gap between the two domains to achieve the segmentation.

Shen and Gao (2019) proposed a domain adaptation approach to ensure the model generates similar features under missing modalities as well as in full modality situations. First, it randomly removes one modality during training to enhance the robustness of the absence of one modality. Then, the network is trained with one missing modalities, at the same time, the adversarial loss is applied to adapt feature maps from the target domain (missing modalities) to the one from the source domain (full modalities). Experimental results show the effectiveness of the proposed method, and also validate that the quality of the segmentation depends on which modality is missing. Furthermore, the proposed method is capable to interpret the segmentation results by visualizing the contribution of each modality.

Wang et al. (2021b) proposed an Adversarial Co-training Network (ACN) to align domain and feature distributions between full modality and missing modalities. Specifically, the proposed network consists of a multimodal path to obtain rich modality information, and a unimodal path to generate modality-specific feature representations. Then, a co-training approach is introduced to establish a coupled learning process between them. It includes three modules: an entropy adversarial learning module (EnA) to match the distributions between two paths; a knowledge adversarial learning module (KnA) to encourage features distribution alignment in latent space; a modality-mutual information knowledge transfer module (MMI) to retain high mutual information between two paths via variational information maximization. Experiments demonstrate that the proposed method outperforms all existing methods on the multimodal BraTS 2018 dataset across all the missing situations.

3.6. Comparative analysis

In this section, we will present a comparative analysis of the above-mentioned methods regarding the principle, architecture, benefits and limitations.

Utilizing the image synthesis technique to compensate for the missing modalities is an intuitive solution, which can recover the missing data information. However, first, additional image synthesis will require more computational time and network complexity. Second, Generative Adversarial Network (GAN) (Goodfellow et al., 2014; Nie et al., 2018; Yu et al., 2019) is usually the most preferred approach for image synthesis. While training GAN will confront with some common issues, such as mode collapse, non-convergence, instability, and high sensibility to hyper-parameters. Thus, the unstable training will result in unsatisfactory synthesis results, which will finally affect the segmentation results. Third, combining the synthesis path and the segmentation path in the same architecture is also challenging (Xia et al., 2021).

Most of the works focused on latent feature space analysis, exploiting a shared feature representation between multi-modalities to retrieve the missing information. This method can be implemented easily via computing the arithmetic operations, such as mean and variance, or simply encouraging the features from different modalities to be close under the L1 or L2 distance. However, the limitation is that it cannot recover the missing image information. Besides, inadequate available modalities will lead to poor feature recovery, resulting in unsatisfactory segmentation results.

The third solution is to discover the multi-source correlation between modalities. On the one hand, it takes advantage of the strong correlation between modalities to compensate for the missing modalities. On the other hand, it can guide the segmentation network to learn the correlated feature representation to improve the segmentation performance. However, when most of the modalities are absent, this method cannot guarantee the restoration of the missing information due to insufficient feature information. In addition, the lost image information usually cannot be recovered without incorporating the image synthesis path.

Another solution is inspired by knowledge distillation, which can transfer knowledge from a trained multi-modal network (teacher) to a mono-modal one (student). However, there are two hindrances to this strategy (Gou et al., 2021): (1) to extract rich knowledge from the teacher. It is essential to understand the influence of each type of knowledge and to know how different kinds of knowledge help each other in a complementary manner. For example, the knowledge from different hidden layers may have different influences on the training of the student model. Thus, it is challenging to extract rich knowledge in a unified and complementary framework. (2) to transfer the knowledge from the teacher to guide the training of the student. In general, the dedicated loss functions or modules are desired to constrain the feature information between student and teacher to be similar.

The last solution is to adopt the domain adaptation method to generate similar features between missing modalities and full modalities situations. The main issue in this method is to reduce the difference between the distributions of source and target domain data since it can directly affect the following segmentation results. Moreover, the missing image information cannot be retrieved without joining an additional image synthesis sub-network.

In conclusion, we provide a brief discussion about each type of method by analyzing its benefits and limitations. The selection of the method will depend on the practical research needs. In general, combining diverse strategies can benefit from the advantages of different methods. The image synthesis-based method can be combined with other four methods (latent feature space-based model, multi-source correlation-based method, knowledge distillation-based method, and domain adaptation-based method). In this way, the network cannot only recover the missing modalities but also achieve brain tumor segmentation. For example, Zhou et al. (2021a) proposed to combine image synthesis and multi-source correlation to address brain tumor segmentation. Nevertheless, the undesirable effects should also be considered. For example, merging the synthesis path can help to recover the missing image formation, but the unstable training and computational cost should be aware. More details about the model design will be mentioned in Section 5.

In the aforementioned introduction, we focus on the models with normal contrast MR images for brain tumor segmentation. While for the low-contrast MR images used in real clinical scenarios, several works have been proposed to address the brain tumor segmentation problem in low-contrast images. For example, Zhou et al. (2019a) proposed a novel high-resolution dense encoder–decoder network for low-contrast medical image segmentation, which is a general plug-in module for encoder–decoder networks to improve performance on low-contrast image segmentation tasks. Another solution is to integrate contour information in the network to focus on the boundary of tumor region, to overcome the limitation of low image contrast. For example, Ilhan et al.

(2022) proposed a nonparametric tumor localization and enhancement method to detect the tumor regions. Hamad et al. (2018) proposed to use of Fuzzy C-Means to detect the tumor regions in MR images.

Additionally, some commercial products for brain tumor segmentation have also been developed and applied in clinical scenarios. For example, 3D Slicer (Kikinis and Pieper, 2011) provides a set of interactive tools for fast and reproducible brain tumor segmentation from MR images. GIST (Cates et al., 2004) is an interactive, GPU-based level-set segmentation tool for 3D medical images. BraTumIA (Brain Tumor Image Analysis)[4] is a software dedicated to multimodal image analysis of brain tumor studies. It can achieve segmentation of healthy and tumor tissues using multimodal MR images (T1, T1-contrast, T2-contrast, and FLAIR), including Gray Matter (GM), White Matter (WM), Cerebrospinal Fluid (CSF), necrotic core, edema, non-enhancing tumor and enhancing tumor. BraTS Toolkit (Kofler et al., 2020) can translate state-of-the-art computational methods into the clinical routine and scientific practice.

4. Performance evaluation

In this section, we will introduce the performance evaluation in the task of brain tumor segmentation with missing modalities. We will describe the prominent brain tumor segmentation datasets, and the widely used evaluation metrics, and also some comparative analysis of the experimental results will be given.

4.1. Datasets

We provide a summary of the commonly used dataset in the field of brain tumor segmentation with missing modalities, which is described in Table 1.

4.1.1. BraTS

The most commonly used public dataset is the BraTS[5] (Brain Tumor Segmentation) challenge dataset (Menze et al., 2014). It focuses on the evaluation of state-of-the-art methods for brain tumor segmentation from multimodal MRI. It utilizes multi-institutional pre-operative MRI scans for the segmentation of gliomas, which are the most frequent primary brain tumors. All the provided images have been pre-processed including co-registered to the same anatomical template, interpolated to the same resolution (1 mm^3) and skull-stripped. The dataset contains four MRI sequences: T1-weighted (T1), post-contrast T1-weighted (T1c), T2-weighted (T2), and Fluid Attenuated Inversion Recovery (FLAIR). The Annotations consist of three sub-tumor regions: enhancing tumor (ET), peritumoral edema (ED), and necrotic and non-enhancing tumor core (NCR/NET). These annotations have been segmented manually by one to four raters, following the same annotation protocol, and the annotations have been approved by experienced neuro-radiologists (Bakas et al., 2017a, 2018, 2017b; Baid et al., 2021).

4.1.2. ISLES 2015

Ischemic Stroke Lesion Segmentation (ISLES)[6] challenge dataset (Maier et al., 2017) consists of two sub-challenges: Sub-Acute Stroke Lesion Segmentation (SISS) and Stroke Perfusion Estimation (SPES). All MRI sequences are skull-stripped, re-sampled to an isotropic spacing of 1 mm^3 (SISS), 2 mm^3 (SPES) and co-registered to the FLAIR (SISS) and T1w contrast (SPES) sequences, respectively. The SISS sub-dataset contains FLAIR, T2w TSE, T1w TFE/TSE, and DWI sequences, and the SPES sub-dataset contains T1c, T2, DWI, CBF, CBV, TTP, and Tmax sequences.

4.1.3. TCGA

The Cancer Genome Atlas (TCGA)[7] dataset (Tomczak et al., 2015; Bakas et al., 2017a) is a multi-institutional comprehensive collection of various molecularly characterized tumor types. It contains both MRI and genomic data for all pre-operative multimodal MRI, including glioblastoma (TCGA-GBM, n = 262) and low-grade-glioma (TCGA-LGG, n = 199) collections via radiological assessment. The glioma sub-region labels were produced by an automated state-of-the-art method and manually revised by an expert board-certified neuroradiologist. A large amount of radiomic features have been extracted according to the manually-revised labels.

4.2. Evaluation metrics

In this section, we will introduce the commonly used standard evaluation metrics for brain tumor segmentation, including the following metrics. Before the introduction, some notations are given:

- TP (True Positive): The prediction is a brain tumor, and the prediction is correct.
- TN (True Negative): The prediction is not a brain tumor, and the prediction is correct.
- FP (False Positive): The prediction is a brain tumor, but the prediction is not correct.
- FN (False Negative): The prediction is not a brain tumor, but the prediction is not correct.

(1) Dice Similarity Coefficient (DSC): It is used to calculate the overlap rate of prediction result and the ground truth. The better predict result will have a larger value. The definition is described in Eq. (1).

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{1}$$

(2) Jaccard Similarity Coefficient: It is used to measure the intersection over the union of two different sets. The better predict result will have a larger value. The definition is described in Eq. (2).

$$Jaccard = \frac{TP}{TP + FP + FN} \tag{2}$$

(3) Sensitivity (recall): It is used to measure the percentage of correct positive values divided by the number of all relevant samples that should have been identified as positive. The better predict result will have a larger value. The definition is described in Eq. (3).

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

(4) Specificity: It is used to measure the percentage of correct negative values divided by the number of all relevant samples that should have been identified as negative. The better predict result will have a larger value. The definition is described in Eq. (4).

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

(5) Precision: It is used to measure the percentage of correct positive results divided by the number of positive results. The better predict result will have a larger value. The definition is described in Eq. (5).

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

(6) Hausdorff Distance (HD): It is computed between the boundaries of the prediction result and the ground truth, it is an indicator of the largest segmentation error. The better predict result will have a smaller value. The definition is described in Eq. (6).

$$HD = \max\{\max_{s\in S}\min_{r\in R} d(s,r), \max_{r\in R}\min_{s\in S} d(r,s)\} \tag{6}$$

where S and R are the two sets of the surface points of the prediction and the real annotation, and d is the Euclidean distance.

[4] https://www.nitrc.org/projects/bratumia
[5] http://braintumorsegmentation.org/
[6] https://www.isles-challenge.org/
[7] https://www.genome.gov/Funded-Programs-Projects/Cancer-Genome-Atlas

Table 1
The summary of the commonly used dataset about training, validation and test images, modalities, image size and time point, # denotes the number of images.

Datasets	# Training	# Validation	# Test	Modality	Image size	Time point
BraTS 2012	35	N/A	15	T1, T1c, T2, FLAIR	$160\times216\times176$ $176\times176\times216$	Pre-operative
BraTS 2013	35	N/A	25	T1, T1c, T2, FLAIR	$160\times216\times176$ $176\times176\times216$	Pre-operative
BraTS 2014	200	N/A	38	T1, T1c, T2, FLAIR	$160\times216\times176$ $176\times176\times216$	Longitudinal
BraTS 2015	200	N/A	53	T1, T1c, T2, FLAIR	$240\times240\times155$	Longitudinal
BraTS 2016	200	N/A	191	T1, T1c, T2, FLAIR	$240\times240\times155$	Longitudinal
BraTS 2017	285	46	146	T1, T1c, T2, FLAIR	$240\times240\times155$	Pre-operative
BraTS 2018	285	66	191	T1, T1c, T2, FLAIR	$240\times240\times155$	Pre-operative
BraTS 2019	335	125	166	T1, T1c, T2, FLAIR	$240\times240\times155$	Pre-operative
BraTS 2020	369	125	166	T1, T1c, T2, FLAIR	$240\times240\times155$	Pre-operative
BraTS 2021	1251	219	570	T1, T1c, T2, FLAIR	$240\times240\times155$	Pre-operative
ISLES 2015	SISS: 28	N/A	36	FLAIR, T2w TSE, T1w TFE/TSE, DWI	$230\times230\times154$	N/A
	SPES: 30	N/A	20	T1c, T2, DWI, CBF, CBV, TTP, Tmax	N/A	N/A
TCGA	262	N/A	N/A	T1, T1c, T2, FLAIR	N/A	Pre-operative
	199	N/A	N/A	T1, T1c, T2, FLAIR	N/A	Pre-operative

Table 2
Summary of brain tumor segmentation methods with missing modalities reviewed in this literature.

Reference	Dataset	Data dimension	Network architecture
Image synthesis-based method			
Jog et al. (2017)	MPRAGE, SPGR	3D	Random forest
Sharma and Hamarneh (2019)	ISLES 2015, BraTS 2018	2D	U-Net, GAN
Islam et al. (2021)	TCGA GBM, BraTS 2017	2D	GAN, FCN
Latent feature space-based method			
Havaei et al. (2016)	MSGC, RRMS, BraTS 2013, BraTS 2015	2D	CNN
Lau et al. (2019)	BraTS 2018	2D	U-Net
Dorent et al. (2019)	BraTS 2018	3D	U-Net, MVAE
Zhu et al. (2021)	BraTS 2018	2D	U-Net, MVAE
Chartsias et al. (2017)	ISLES 2015, BraTS 2015	2D	FCN
Chen et al. (2019)	BraTS 2015	3D	U-Net
Multi-source correlation-based method			
Zhou et al. (2021b)	BraTS 2018, BraTS 2019	3D	U-Net
Zhou et al. (2021a)	BraTS 2018	3D	U-Net
Knowledge distillation-based method			
Hu et al. (2020b)	BraTS 2018	3D	U-Net
Vadacchino et al. (2021)	BraTS 2019	3D	U-Net
Domain adaptation-based method			
Shen and Gao (2019)	BraTS 2017	2D	U-Net, GAN
Wang et al. (2021b)	BraTS 2018	3D	U-Net, GAN

4.3. Comparative analysis

First, we summarize the methods mentioned in Section 3 across datasets, data dimensions and network architecture, shown in Table 2. Then, we selected some state-of-the-art brain tumor segmentation methods, and present a quantitative analysis of these methods. To offer a fair comparison, we compare the methods which are trained on the same dataset with the same experimental settings.

From Table 2, we can observe that most methods utilize the BraTS challenge dataset since it provides a large amount of available multi-modal data. From the view of data dimension, 2D or 3D are both possible choices depending on the model complexity. Besides, U-Net is the most used network architecture, because its contracting path can capture contextual information, the symmetric expanding path can enable precise localization, and the skip connections allow for faster convergence during training (Szegedy et al., 2017).

Table 3 presents a quantitative comparison of two methods trained on the BraTS 2018 dataset. These two methods focus on brain tumor segmentation in the case of missing only one modality. The DSC is denoted for three tumor sub-regions including whole tumor, tumor core and enhancing tumor. The work from Zhu et al. (2021) adopts the latent feature space-based method, while the work from Shen and Gao (2019) used the domain adaptation-based method. It can be seen that the method from Zhu et al. (2021) obtains better DSC scores in most cases than the one from Shen and Gao (2019). We explain that the method from Zhu et al. (2021) utilizes the residual MVAE to embed more information into the latent space and decoder the features with a residual form to stabilize the features, which can reduce the difficulty of learning and achieve stable performance in the case of missing modality.

Table 4 provides the comparison in the mono-modal brain tumor segmentation based on the BraTS 2018 dataset. The DSC is denoted for three tumor sub-regions, including whole tumor, tumor core and enhancing tumor. ACN (Wang et al., 2021b) is based on domain adaptation, while KD-Net (Hu et al., 2020b) is based on knowledge distillation. We can observe the recent method ACN (Wang et al., 2021b) achieves better results than KD-Net (Hu et al., 2020b) across all the cases. The reason for that is, in the KD-Net (Hu et al., 2020b), the teacher model may contain some irrelevant information to the mono-modal model, and it can weaken the feature learning ability of the network. On the contrary, the method ACN (Wang et al., 2021b) establishes a coupled learning process from both full modality and missing modality to reduce the domain gap as well as recover the most relevant representations.

Table 5 summarizes the recent brain tumor segmentation methods trained on the BraTS 2018 dataset in the case of missing any number of modalities. The DSC is denoted for three tumor sub-regions, including whole tumor, tumor core and enhancing tumor. We can observe the domain adaptation-based method ACN (Wang et al., 2021b) outperforms all the other compared methods, especially the early work

Table 3
Comparison results in terms of DSC in the case of missing only one modality on the BraTS 2018 dataset, • denotes the present modality and ∘ denotes the missing one. Notice that Shen and Gao (2019) belongs to domain adaptation-based method, Zhu et al. (2021) belongs to the latent feature space-based method.

Modality				Whole tumor		Tumor core		Enhancing tumor	
FLAIR	T1	T1c	T2	Shen and Gao (2019)	Zhu et al. (2021)	Shen and Gao (2019)	Zhu et al. (2021)	Shen and Gao (2019)	Zhu et al. (2021)
•	•	•	∘	89.3	87.4	77.5	77.8	64.3	79.1
•	•	∘	•	87.9	87.6	57.0	65.4	48.4	44.8
•	∘	•	•	89.0	87.9	77.8	78.2	64.2	78.5
∘	•	•	•	61.6	81.8	68.0	77.3	55.2	78.7
•	•	•	•	89.4	87.9	79.0	78.8	65.3	79.1

Table 4
Comparison results in terms of DSC for the mono-modal segmentation on the BraTS 2018 dataset, • denotes the present modality and ∘ denotes the missing one. Notice that KD-Net (Hu et al., 2020b) belongs to knowledge distillation-based method, ACN (Wang et al., 2021b) belongs to the domain adaptation-based method.

Modality				Whole tumor		Tumor core		Enhancing tumor	
FLAIR	T1	T1c	T2	KD-Net (Hu et al., 2020b)	ACN (Wang et al., 2021b)	KD-Net (Hu et al., 2020b)	ACN (Wang et al., 2021b)	KD-Net (Hu et al., 2020b)	ACN (Wang et al., 2021b)
∘	∘	∘	•	82.32	85.55	66.01	67.94	39.04	42.98
∘	∘	•	∘	76.79	80.52	81.89	84.18	75.32	78.07
∘	•	∘	∘	77.28	79.34	70.02	71.18	39.87	41.52
•	∘	∘	∘	85.14	87.30	65.97	67.72	40.99	42.77

Table 5
Comparison results in terms of DSC in the case of missing any number of modalities on the BraTS 2018 dataset, • denotes the present modality and ∘ denotes the missing one. Notice that U-HeMIS (Havaei et al., 2016), URN (Lau et al., 2019) and U-HVED (Dorent et al., 2019) belong to the latent feature space-based method, Zhou et al. (2021a) belongs to the multi-source correlation-based method, and ACN (Wang et al., 2021b) belongs to the domain adaptation-based method.

Modality				Whole tumor					Tumor core					Enhancing tumor				
FLAIR	T1	T1c	T2	U-HeMIS (Havaei et al., 2016)	URN (Lau et al., 2019)	U-HVED (Dorent et al., 2019)	(Zhou et al., 2021a)	ACN (Wang et al., 2021b)	U-HeMIS (Havaei et al., 2016)	URN (Lau et al., 2019)	U-HVED (Dorent et al., 2019)	(Zhou et al., 2021a)	ACN (Wang et al., 2021b)	U-HeMIS (Havaei et al., 2016)	URN (Lau et al., 2019)	U-HVED (Dorent et al., 2019)	(Zhou et al., 2021a)	ACN (Wang et al., 2021b)
∘	∘	∘	•	79.2	77.5	80.9	80.4	85.5	50.0	43.6	54.1	59.5	67.9	23.3	20.3	30.8	35.2	42.9
∘	∘	•	∘	58.5	62.2	62.4	72.0	80.5	58.5	58.5	66.7	83.1	84.1	60.8	55.8	65.5	75.0	78.0
∘	•	∘	∘	54.3	50.4	52.4	74.9	79.3	37.9	34.2	37.2	55.5	71.1	12.4	19.1	13.7	53.2	41.5
•	∘	∘	∘	79.9	84.8	82.1	85.9	87.3	49.8	50.4	50.4	64.6	67.7	24.9	23.6	24.8	39.5	42.7
∘	∘	•	•	81.0	80.3	82.7	81.7	86.4	69.1	68.9	73.7	84.8	84.4	68.6	67.6	70.2	75.5	75.6
∘	•	•	∘	63.8	69.8	66.8	76.2	80.0	64.0	65.9	69.7	84.2	84.5	65.3	66.5	67.0	75.8	75.2
•	•	∘	∘	83.9	85.5	84.3	86.5	87.4	56.7	52.6	55.3	67.0	71.3	29.0	25.3	24.2	43.2	43.7
∘	•	∘	•	80.8	80.8	82.2	83.4	85.5	53.4	48.6	57.2	62.6	73.2	28.3	25.2	30.7	38.0	47.3
•	∘	∘	•	86.0	86.3	87.5	86.5	87.7	58.7	50.7	59.7	66.6	71.6	28.0	25.2	34.6	45.5	45.9
•	∘	•	∘	83.3	85.8	85.8	86.2	88.2	67.6	72.5	72.9	85.0	83.3	68.0	70.4	70.3	77.1	77.4
•	•	•	∘	85.1	85.6	86.2	86.6	88.9	70.7	72.0	74.2	85.6	84.2	69.9	71.0	71.1	77.2	76.1
•	•	∘	•	87.0	86.1	88.0	87.6	88.3	61.0	52.5	61.5	68.0	67.8	33.4	25.8	34.1	45.6	42.0
•	∘	•	•	87.0	86.5	88.6	86.4	88.3	72.2	72.2	75.6	86.0	82.8	69.7	68.5	71.1	76.2	75.9
∘	•	•	•	82.1	81.1	83.3	82.9	86.9	70.7	69.5	75.3	85.2	84.6	69.7	68.5	71.1	76.2	76.1
•	•	•	•	87.6	86.3	88.8	86.6	89.2	73.4	71.8	76.4	85.8	85.1	70.8	69.9	71.7	76.9	77.0
Average				78.6	79.3	80.1	82.9	85.9	59.7	58.9	64.0	74.9	77.6	48.1	46.9	50.0	59.1	61.2

U-HeMIS (Havaei et al., 2016), it achieves improvements by 9.3% on whole tumor, 30% on tumor core, and 27.2% on enhancing tumor, respectively, in terms of average DSC score. The superior performance of ACN is due to the following factors, (1) the better network architecture: ACN adopts a coupled learning process to enhance the learning ability of both multimodal and uni-modal training. (2) the more efficient objective functions: three major components are presented to bridge the domain gap and encourage the alignment of latent representations between multimodal and uni-modal paths, and also to recover the most relevant features for incomplete modalities.

Table 6 summarizes the performance of image (FLAIR) synthesis between REPLICA and Chartsias et al. (2017) on ISLES dataset in the case of missing modalities. REPLICA (Jog et al., 2017) is a supervised random forest-based image synthesis method. The method from Chartsias et al. (2017) is a multi-input multi-output based fully convolutional neural network. We can observe that the method from Chartsias et al. (2017) outperforms REPLICA in 6 of the 7 experimental setups, with statistically significant improvements in 5 cases. We explain that the method from Chartsias et al. (2017) can learn a shared latent feature space to achieve image synthesis without relying on the input modalities.

Table 6
Synthesis of FLAIR image in terms of MSE in the case of missing modalities on the ISLES dataset between Chartsias et al. (2017) and REPLICA. • denotes the present modality and ∘ denotes the missing one. Notice that Chartsias et al. (2017) belongs to the latent feature space-based method, Jog et al. (2017) belongs to the image synthesis-based method.

Modality			MSE (FLAIR)	
T1	T2	DWI	REPLICA	Chartsias et al. (2017)
•	∘	∘	0.301(0.11)	**0.249(0.09)**
∘	•	∘	0.374(0.16)	**0.321(0.12)**
∘	∘	•	0.278(0.09)	0.285(0.13)
∘	•	•	0.235(0.08)	0.214(0.09)
•	∘	•	0.225(0.08)	**0.198(0.02)**
•	•	∘	0.271(0.12)	**0.214(0.08)**
•	•	•	0.210(0.08)	**0.171(0.06)**
Average			0.271	**0.236**

5. Future directions

In recent years, a great deal of work has been proposed for brain tumor segmentation with missing modalities. Given this rapid process, one may wonder if these methods can be directly applied to real-world applications to reduce the workload of clinicians. The current literature

suggests that there is still a long way to go (Chen et al., 2020). In the following, we will analyze the current major challenges and discuss the future research directions concerning data diversity, data annotation, uncertainty quantification, model complexity and model interpretation.

5.1. Data diversity

Data diversity refers to the continuum of various types of elements in the data. The integration of data plays a significant role in increasing the precious, efficiency and reliability of deep models (Mehmood et al., 2020; Nan et al., 2022). In the field of brain tumor segmentation with missing modalities, most works apply the BraTS dataset to implement the experimental analysis. Although this dataset provides a large number of image data for training, validation and testing, the diversity of the data is not enough. For example, adding the genomic information from the clinical scenarios will aid in discovering the tumor patterns and characteristics (Tomczak et al., 2015; Bakas et al., 2017a). Therefore, a variety of public multi-modal datasets are desired to enrich the data diversity to enhance the performance of the deep models.

5.2. Data annotation

Over the past years, deep learning-based methods have made striking progress in various applications, and the capability heavily relies on a large amount of training data, especially training annotation. However, collecting such a huge dataset of annotated data in medical image segmentation is a tough task and performing the annotation is a time-consuming, labor-intensive, and expensive process (Wang et al., 2021a). Several approaches can be utilized to address this problem: data augmentation, transfer learning, and weakly and semi-supervised learning (Chen et al., 2020). Data augmentation aims to increase the size and the variety of training images by applying a set of affine transformations, such as flip, rotate, scaling, and color augmenting (Hesamian et al., 2019). Transfer learning could be another solution to tackle the limited annotation via reusing a pre-trained model. Weakly and semi-supervised learning methods aim at improving learning accuracy by making use of both labeled and unlabeled or weakly labeled data (Sun et al., 2019; Chen et al., 2021).

5.3. Uncertainty quantification

Uncertainty quantification plays an essential role in decision-making, which is employed in many applications in science and engineering (Abdar et al., 2021). For brain tumor segmentation, first, image acquisition methods can bring noise; Second, different experts may disagree on the tumor annotation in some ambiguous cases. Thus, only predicting a single segmentation is not adequate. It is necessary to obtain both the prediction of tumor location and the confidence of the prediction to assist doctors in making the final decision. Uncertainty quantification is a potential solution which can provide some insight into the confidence of segmentation, and can highlight areas of likely segmentation error to the users (Czolbe et al., 2021; Jungo et al., 2020).

5.4. Model complexity

Model complexity is a fundamental problem in deep learning, which is key to precisely understanding the capability and limitation of the model. It involves model framework, model size, optimization process and data complexity (Hu et al., 2021, 2020a; Nakkiran et al., 2021). Specifically, the factor of the model framework includes the model type (e.g., feed-forward neural network, convolutional neural network), activation function (e.g., Sigmoid (Cybenko, 1989), ReLU (Nair and Hinton, 2010)), and others. The model size influences model complexity mainly through the number of parameters, the number of filters, and the filter size. The optimization process affects model complexity via the objective functions, the selection of learning algorithms, and the setting of hyperparameters. The training data also affect model complexity by data dimension, data distribution (Mohri et al., 2018) and information volume. Particularly, in clinical scenarios, an efficient model is often required to assist doctors in making an accurate and effective decision in the process of tumor diagnosis and surgical treatment.

5.5. Model interpretation

Deep learning-based networks have exhibited remarkable performance in a variety of tasks, particularly in computer vision and Nature language processing (Voulodimos et al., 2018; Young et al., 2018). However, due to their black-box nature, it is often difficult to understand the prediction results of deep models. There has been an increasing interest in finding ways to interpret the features discovered by neural networks (Yang et al., 2022; Li et al., 2021; Zeiler and Fergus, 2014; Grün et al., 2016; Singh et al., 2017; Zhang and Zhu, 2018). In clinical practice, for example, in the task of multimodal brain tumor segmentation, model interpretation plays an essential role in guiding the researchers and also the clinical doctors to explain or reveal the ways that deep models make decisions, such as the combination of features from multi-modalities, or the importance of each modality in the training process (Shen and Gao, 2019). Besides, Model interpretation can be incorporated with the clinical factors (patient history and gene expression data) to further boost the accuracy of deep models.

6. Conclusion

The paper presents a comprehensive literature survey for brain tumor segmentation in the case of missing modalities. The main goal is to help the readers to know the latest contributions and explore future research directions. To achieve it, we first introduce the biomedical background review about a brain tumor, the MRI imaging technique, and current challenges in brain tumor segmentation. Then, we provide a taxonomy of the recent state-of-the-art methods with five categories: image synthesis-based method, latent feature space-based model, multi-source correlation-based method, knowledge distillation-based method, and domain adaptation-based method. We offer an in-depth analysis of the principles, architectures, benefits and limitations of these methods. Subsequently, we describe the frequently used brain tumor datasets and evaluation metrics. Finally, we analyze the existing challenges and suggest prospects.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 6220071321).

References

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U.R., et al., 2021. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Inf. Fusion 76, 243–297.

Achituve, I., Maron, H., Chechik, G., 2021. Self-supervised learning for domain adaptation on point clouds. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 123–133.

Agravat, R.R., Raval, M.S., 2021. A survey and analysis on automated glioma brain tumor segmentation and overall patient survival prediction. Arch. Comput. Methods Eng. 28 (5), 4117–4152.

Ba, J., Caruana, R., 2014. Do deep nets really need to be deep? Adv. Neural Inf. Process. Syst. 27.

Bahadure, N.B., Ray, A.K., Thethi, H.P., 2017. Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int. J. Biomed. Imaging 2017.

Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., et al., 2021. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C., 2017a. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data 4 (1), 1–13.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C., 2017b. Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection. The cancer imaging archive. Nat. Sci. Data 4, 170117.

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., et al., 2018. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint arXiv:1811.02629.

Ben Rabeh, A., Benzarti, F., Amiri, H., 2017. Segmentation of brain MRI using active contour model. Int. J. Imaging Syst. Technol. 27 (1), 3–11.

Bengio, Y., Louradour, J., Collobert, R., Weston, J., 2009. Curriculum learning. In: Proceedings of the 26th Annual International Conference on Machine Learning. pp. 41–48.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.

Cates, J.E., Lefohn, A.E., Whitaker, R.T., 2004. GIST: an interactive, GPU-based level set segmentation tool for 3D medical images. Med. Image Anal. 8 (3), 217–231.

Chartsias, A., Joyce, T., Giuffrida, M.V., Tsaftaris, S.A., 2017. Multimodal MR synthesis via modality-invariant latent representation. IEEE Trans. Med. Imaging 37 (3), 803–814.

Chen, C., Dou, Q., Jin, Y., Chen, H., Qin, J., Heng, P.-A., 2019. Robust multimodal brain tumor segmentation via feature disentanglement and gated fusion. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 447–456.

Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., Rueckert, D., 2020. Deep learning for cardiac image segmentation: a review. Front. Cardiovasc. Med. 25.

Chen, G., Ru, J., Zhou, Y., Rekik, I., Pan, Z., Liu, X., Lin, Y., Lu, B., Shi, J., 2021. MTANS: Multi-scale mean teacher combined adversarial network with shape-aware embedding for semi-supervised brain lesion segmentation. NeuroImage 244, 118568.

Choi, J., Kim, T., Kim, C., 2019. Self-ensembling with gan-based data augmentation for domain adaptation in semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6830–6840.

Cong, W., Song, J., Luan, K., Liang, H., Wang, L., Ma, X., Li, J., 2016. A modified brain MR image segmentation and bias field estimation model based on local and global information. Comput. Math. Methods Med. 2016.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Math. Control Signals Systems 2 (4), 303–314.

Czolbe, S., Arnavaz, K., Krause, O., Feragen, A., 2021. Is segmentation uncertainty useful? In: International Conference on Information Processing in Medical Imaging. Springer, pp. 715–726.

Das, S., Nayak, G., Saba, L., Kalra, M., Suri, J.S., Saxena, S., 2022. An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review. Comput. Biol. Med. 105273.

Dorent, R., Joutard, S., Modat, M., Ourselin, S., Vercauteren, T., 2019. Hetero-modal variational encoder-decoder for joint modality completion and segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 74–82.

Fang, L., Wang, X., 2022. Brain tumor segmentation based on the dual-path network of multi-modal MRI images. Pattern Recognit. 124, 108434.

Farmanfarma, K.K., Mohammadian, M., Shahabinia, Z., Hassanipour, S., Salehiniya, H., 2019. Brain cancer in the world: an epidemiological review. World Cancer Res. J. 6 (5).

Ghaffari, M., Sowmya, A., Oliver, R., 2019. Automated brain tumor segmentation using multimodal brain scans: a survey based on models submitted to the BraTS 2012–2018 challenges. IEEE Rev. Biomed. Eng. 13, 156–168.

Goel, N.J., Bird, C.E., Hicks, W.H., Abdullah, K.G., 2021. Economic implications of the modern treatment paradigm of glioblastoma: an analysis of global cost estimates and their utility for cost assessment. J. Med. Econ. 24 (1), 1018–1024.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Adv. Neural Inf. Process. Syst. 27.

Gou, J., Yu, B., Maybank, S.J., Tao, D., 2021. Knowledge distillation: A survey. Int. J. Comput. Vis. 129 (6), 1789–1819.

Grün, F., Rupprecht, C., Navab, N., Tombari, F., 2016. A taxonomy and library for visualizing learned features in convolutional neural networks. arXiv preprint arXiv:1606.07757.

Hamad, Y.A., Simonov, K., Naeem, M.B., 2018. Brain's tumor edge detection on low contrast medical images. In: 2018 1st Annual International Conference on Information and Sciences. AiCIS, IEEE, pp. 45–50.

Havaei, M., Guizard, N., Chapados, N., Bengio, Y., 2016. Hemis: Hetero-modal image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 469–477.

Hesamian, M.H., Jia, W., He, X., Kennedy, P., 2019. Deep learning techniques for medical image segmentation: achievements and challenges. J. Digit. Imaging 32 (4), 582–596.

Hinton, G.E., Vinyals, O., Dean, J., 2015. Distilling the knowledge in a neural network. Deep learning and representation learning workshop. In: Neural Information Processing Systems.

Hsu, H.-K., Yao, C.-H., Tsai, Y.-H., Hung, W.-C., Tseng, H.-Y., Singh, M., Yang, M.-H., 2020. Progressive domain adaptation for object detection. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 749–757.

Hu, X., Chu, L., Pei, J., Liu, W., Bian, J., 2021. Model complexity of deep learning: A survey. Knowl. Inf. Syst. 63 (10), 2585–2619.

Hu, X., Liu, W., Bian, J., Pei, J., 2020a. Measuring model complexity of neural networks with curve activation functions. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 1521–1531.

Hu, M., Maillard, M., Zhang, Y., Ciceri, T., La Barbera, G., Bloch, I., Gori, P., 2020b. Knowledge distillation from multi-modal to mono-modal segmentation networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 772–781.

Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7132–7141.

Ilhan, A., Sekeroglu, B., Abiyev, R., 2022. Brain tumor segmentation in MRI images using nonparametric localization and enhancement methods with U-net. Int. J. Comput. Assist. Radiol. Surg. 17 (3), 589–600.

Inoue, N., Furuta, R., Yamasaki, T., Aizawa, K., 2018. Cross-domain weakly-supervised object detection through progressive domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5001–5009.

Işın, A., Direkoğlu, C., Şah, M., 2016. Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Comput. Sci. 102, 317–324.

Islam, M., Wijethilake, N., Ren, H., 2021. Glioblastoma multiforme prognosis: Mri missing modality generation, segmentation and radiogenomic survival prediction. Comput. Med. Imaging Graph. 91, 101906.

Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1125–1134.

Jog, A., Carass, A., Roy, S., Pham, D.L., Prince, J.L., 2017. Random forest regression for magnetic resonance image synthesis. Med. Image Anal. 35, 475–488.

Jungo, A., Balsiger, F., Reyes, M., 2020. Analyzing the quality and challenges of uncertainty estimations for brain tumor segmentation. Front. Neurosci. 14, 282.

Kikinis, R., Pieper, S., 2011. 3D slicer as a tool for interactive brain tumor segmentation. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 6982–6984.

Kofler, F., Berger, C., Waldmannstetter, D., Lipkova, J., Ezhov, I., Tetteh, G., Kirschke, J., Zimmer, C., Wiestler, B., Menze, B.H., 2020. BraTS toolkit: translating BraTS brain tumor segmentation algorithms into clinical and scientific practice. Front. Neurosci. 125.

Lapuyade-Lahorgue, J., Xue, J.-H., Ruan, S., 2017. Segmenting multi-source images using hidden Markov fields with copula-based multivariate statistical distributions. IEEE Trans. Image Process. 26 (7), 3187–3195.

Lau, K., Adler, J., Sjölund, J., 2019. A unified representation network for segmentation with missing modalities. arXiv preprint arXiv:1908.06683.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.

Li, Z., Dong, M., Wen, S., Hu, X., Zhou, P., Zeng, Z., 2019a. CLU-CNNs: Object detection for medical images. Neurocomputing 350, 53–59.

Li, Q., Jin, S., Yan, J., 2017. Mimicking very efficient network for object detection. In: Proceedings of the Ieee Conference on Computer Vision and Pattern Recognition. pp. 6356–6364.

Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J., Dou, D., 2021. Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond. arXiv preprint arXiv:2103.10689.

Li, Y., Yuan, L., Vasconcelos, N., 2019b. Bidirectional learning for domain adaptation of semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6936–6945.

Liu, Y., Chen, K., Liu, C., Qin, Z., Luo, Z., Wang, J., 2019. Structured knowledge distillation for semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2604–2613.

Liu, Y., Zhang, W., Wang, J., 2021. Source-free domain adaptation for semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1215–1224.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440.

Lopez-Paz, D., Bottou, L., Schölkopf, B., Vapnik, V., 2015. Unifying distillation and privileged information. arXiv preprint arXiv:1511.03643.

Maier, O., Menze, B.H., von der Gablentz, J., Häni, L., Heinrich, M.P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., et al., 2017. ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. 35, 250–269.

Mehmood, I., Anwar, S., Dilawar, A., Zulfiqar, I., Abbas, R., 2020. Managing data diversity on the internet of medical things (IoMT). Int. J. Inf. Technol. Comput. Sci.(IJITCS) 12 (6), 49–56.

Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al., 2014. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging 34 (10), 1993–2024.

Miller, K.D., Ostrom, Q.T., Kruchko, C., Patil, N., Tihan, T., Cioffi, G., Fuchs, H.E., Waite, K.A., Jemal, A., Siegel, R.L., et al., 2021. Brain and other central nervous system tumor statistics, 2021. CA: Cancer J. Clin. 71 (5), 381–406.

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Mohri, M., Rostamizadeh, A., Talwalkar, A., 2018. Foundations of Machine Learning. MIT Press.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: Icml.

Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., Sutskever, I., 2021. Deep double descent: Where bigger models and more data hurt. J. Stat. Mech. Theory Exp. 2021 (12), 124003.

Nan, Y., Del Ser, J., Walsh, S., Schönlieb, C., Roberts, M., Selby, I., Howard, K., Owen, J., Neville, J., Guiot, J., et al., 2022. Data harmonisation for information fusion in digital healthcare: A state-of-the-art systematic review, meta-analysis and future research directions. Inf. Fusion.

Nichols, T.E., Das, S., Eickhoff, S.B., Evans, A.C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M.P., Poldrack, R.A., Poline, J.-B., et al., 2017. Best practices in data analysis and sharing in neuroimaging using MRI. Nature Neurosci. 20 (3), 299–303.

Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q., Shen, D., 2018. Medical image synthesis with deep convolutional adversarial networks. IEEE Trans. Biomed. Eng. 65 (12), 2720–2730.

Peng, J., Sun, W., Ma, L., Du, Q., 2019. Discriminative transfer joint matching for domain adaptation in hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 16 (6), 972–976.

Petersen, S.E., Aung, N., Sanghvi, M.M., Zemrak, F., Fung, K., Paiva, J.M., Francis, J.M., Khanji, M.Y., Lukaschuk, E., Lee, A.M., et al., 2017. Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in caucasians from the UK biobank population cohort. J. Cardiovasc. Magn. Resonance 19 (1), 1–19.

Rao, C.S., Karunakara, K., 2021. A comprehensive review on brain tumor segmentation and classification of MRI images. Multimedia Tools Appl. 80 (12), 17611–17643.

Ren, J., Hacihaliloglu, I., Singer, E.A., Foran, D.J., Qi, X., 2018. Adversarial domain adaptation for classification of prostate histopathology whole-slide images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 201–209.

Renard, F., Guedria, S., Palma, N.D., Vuillerme, N., 2020. Variability and reproducibility in deep learning for medical image segmentation. Sci. Rep. 10 (1), 1–16.

Rietzler, A., Stabinger, S., Opitz, P., Engl, S., 2019. Adapt or get left behind: Domain adaptation through bert language model finetuning for aspect-target sentiment classification. arXiv preprint arXiv:1908.11860.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Rouse, C., Gittleman, H., Ostrom, Q.T., Kruchko, C., Barnholtz-Sloan, J.S., 2015. Years of potential life lost for brain and CNS tumors relative to other cancers in adults in the United States, 2010. Neuro-Oncol. 18 (1), 70–77.

Roy, A.G., Navab, N., Wachinger, C., 2018. Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 421–429.

Saouli, R., Akil, M., Kachouri, R., et al., 2018. Fully automatic brain tumor segmentation using end-to-end incremental deep neural networks in MRI images. Comput. Methods Programs Biomed. 166, 39–49.

Schmidt, G.P., Reiser, M.F., Baur-Melnyk, A., 2007. Whole-body imaging of the musculoskeletal system: the value of MR imaging. Skeletal Radiol. 36 (12), 1109–1119.

Sharma, A., Hamarneh, G., 2019. Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. IEEE Trans. Med. Imaging 39 (4), 1170–1183.

Sharma, P., Shukla, A.P., 2021. A review on brain tumor segmentation and classification for MRI images. In: 2021 International Conference on Advance Computing and Innovative Technologies in Engineering. ICACITE, IEEE, pp. 963–967.

Shen, Y., Gao, M., 2019. Brain tumor segmentation on MRI with missing modalities. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 417–428.

Singh, A., Bajpai, S., Karanam, S., Choubey, A., Raviteja, T., 2012. Malignant brain tumor detection. Int. J. Comput. Eng 4 (6), 1002.

Singh, R., Lanchantin, J., Sekhon, A., Qi, Y., 2017. Attend and predict: Understanding gene regulation by selective attention on chromatin. Adv. Neural Inf. Process. Syst. 30.

Sun, Y., Zhou, C., Fu, Y., Xue, X., 2019. Parasitic GAN for semi-supervised brain tumor segmentation. In: 2019 IEEE International Conference on Image Processing. ICIP, IEEE, pp. 1535–1539.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-First AAAI Conference on Artificial Intelligence.

Tiwari, A., Srivastava, S., Pant, M., 2020. Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019. Pattern Recognit. Lett. 131, 244–260.

Tomczak, K., Czerwińska, P., Wiznerowicz, M., 2015. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp. Oncol. 19 (1A), A68.

Urban, G., Geras, K.J., Kahou, S.E., Aslan, O., Wang, S., Caruana, R., Mohamed, A., Philipose, M., Richardson, M., 2016. Do deep convolutional nets really need to be deep and convolutional? arXiv preprint arXiv:1603.05691.

Vadacchino, S., Mehta, R., Sepahvand, N.M., Nichyporuk, B., Clark, J.J., Arbel, T., 2021. Had-net: A hierarchical adversarial knowledge distillation network for improved enhanced tumour segmentation without post-contrast images. In: Medical Imaging with Deep Learning. PMLR, pp. 787–801.

Vapnik, V., Izmailov, R., et al., 2015. Learning using privileged information: similarity control and knowledge transfer. J. Mach. Learn. Res. 16 (1), 2023–2049.

Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep learning for computer vision: A brief review. Comput. Intell. Neurosci. 2018.

Vu, T.-H., Jain, H., Bucher, M., Cord, M., Pérez, P., 2019. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2517–2526.

Wadhwa, A., Bhardwaj, A., Verma, V.S., 2019. A review on brain tumor segmentation of MRI images. Magn. Resonance Imaging 61, 247–259.

Wang, S., Li, C., Wang, R., Liu, Z., Wang, M., Tan, H., Wu, Y., Liu, X., Sun, H., Yang, R., et al., 2021a. Annotation-efficient deep learning for automatic medical image segmentation. Nature Commun. 12 (1), 1–13.

Wang, Y., Zhang, Y., Liu, Y., Lin, Z., Tian, J., Zhong, C., Shi, Z., Fan, J., He, Z., 2021b. ACN: Adversarial co-training network for brain tumor segmentation with missing modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 410–420.

Wu, B., Zhou, X., Zhao, S., Yue, X., Keutzer, K., 2019. Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud. In: 2019 International Conference on Robotics and Automation. ICRA, IEEE, pp. 4376–4382.

Xia, Y., Zhang, L., Ravikumar, N., Attar, R., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F., 2021. Recovering from missing data in population imaging–Cardiac MR image imputation via conditional generative adversarial nets. Med. Image Anal. 67, 101812.

Xie, J., Shuai, B., Hu, J.-F., Lin, J., Zheng, W.-S., 2018. Improving fast segmentation with teacher-student learning. arXiv preprint arXiv:1810.08476.

Xiong, S., Wu, G., Fan, X., Feng, X., Huang, Z., Cao, W., Zhou, X., Ding, S., Yu, J., Wang, L., et al., 2021. MRI-based brain tumor segmentation using FPGA-accelerated neural network. BMC Bioinformatics 22 (1), 1–15.

Yang, G., Ye, Q., Xia, J., 2022. Unbox the black-box for the medical explainable ai via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Inf. Fusion 77, 29–52.

Young, T., Hazarika, D., Poria, S., Cambria, E., 2018. Recent trends in deep learning based natural language processing. Ieee Comput. Intell. Mag. 13 (3), 55–75.

Yu, B., Zhou, L., Wang, L., Shi, Y., Fripp, J., Bourgeat, P., 2019. Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis. IEEE Trans. Med. Imaging 38 (7), 1750–1762.

Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks. In: European Conference on Computer Vision. Springer, pp. 818–833.

Zhang, D., Huang, G., Zhang, Q., Han, J., Han, J., Yu, Y., 2021a. Cross-modality deep feature learning for brain tumor segmentation. Pattern Recognit. 110, 107562.

Zhang, W., Wu, Y., Yang, B., Hu, S., Wu, L., Dhelimd, S., 2021b. Overview of multi-modal brain tumor MR image segmentation. Healthcare 9 (8), 1051.

Zhang, Q.-s., Zhu, S.-C., 2018. Visual interpretability for deep learning: a survey. Front. Inf. Technol. Electron. Eng. 19 (1), 27–39.

Zhou, T., Canu, S., Ruan, S., 2020. Fusion based on attention mechanism and context constraint for multi-modal brain tumor segmentation. Comput. Med. Imaging Graph. 86, 101811.

Zhou, T., Canu, S., Vera, P., Ruan, S., 2021a. Feature-enhanced generation and multi-modality fusion based deep neural network for brain tumor segmentation with missing MR modalities. Neurocomputing 466, 102–112.

Zhou, T., Canu, S., Vera, P., Ruan, S., 2021b. Latent correlation representation learning for brain tumor segmentation with missing MRI modalities. IEEE Trans. Image Process. 30, 4263–4274.

Zhou, S., Nie, D., Adeli, E., Yin, J., Lian, J., Shen, D., 2019a. High-resolution encoder–decoder networks for low-contrast medical image segmentation. IEEE Trans. Image Process. 29, 461–475.

Zhou, T., Ruan, S., Canu, S., 2019b. A review: Deep learning for medical image segmentation using multi-modality fusion. Array 3, 100004.
Zhou, T., Ruan, S., Vera, P., Canu, S., 2022. A tri-attention fusion guided multi-modal segmentation network. Pattern Recognit. 124, 108417.
Zhu, Y., Wang, S., Hu, Y., Ma, X., Qin, Y., Xie, J., 2021. DRM-VAE: A dual residual multi variational auto-encoder for brain tumor segmentation with missing modalities. In: 2021 IEEE 4th International Conference on Electronics and Communication Engineering. ICECE, IEEE, pp. 82–86.