J. Meas. Sci. Instrum., 2025, 16(1): 11-25.
DOI: 10.62756/jmsi.1674-8042.2025002
http://jmsi.publish.founderss.cn/
jmsi@nuc.edu.cn

EACNet: Ensemble adversarial co-training neural network for handling missing modalities in MRI images for brain tumor segmentation

RAMADHAN Amran Juma[1], CHEN Jing[1,2*], PENG Junlan[1]

1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China;
2. Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computing Intelligence, Jiangnan University, Wuxi 214122, China

***Corresponding author:** CHEN Jing (chenjing@jiangnan.edu.cn)

Received: September 26, 2024 **Revised:** December 19, 2024 **Accepted:** January 8, 2025

Abstract: Brain tumor segmentation is critical in clinical diagnosis and treatment planning. Existing methods for brain tumor segmentation with missing modalities often struggle when dealing with multiple missing modalities, a common scenario in real-world clinical settings. These methods primarily focus on handling a single missing modality at a time, making them insufficiently robust for the additional complexity encountered with incomplete data containing various missing modality combinations. Additionally, most existing methods rely on single models, which may limit their performance and increase the risk of overfitting the training data. This work proposes a novel method called the ensemble adversarial co-training neural network (EACNet) for accurate brain tumor segmentation from multi-modal magnetic resonance imaging (MRI) scans with multiple missing modalities. The proposed method consists of three key modules: the ensemble of pre-trained models, which captures diverse feature representations from the MRI data by employing an ensemble of pre-trained models; adversarial learning, which leverages a competitive training approach involving two models; a generator model, which creates realistic missing data, while sub-networks acting as discriminators learn to distinguish real data from the generated "fake" data. Co-training framework utilizes the information extracted by the multimodal path (trained on complete scans) to guide the learning process in the path handling missing modalities. The model potentially compensates for missing information through co-training interactions by exploiting the relationships between available modalities and the tumor segmentation task. EACNet was evaluated on the BraTS2018 and BraTS2020 challenge datasets and achieved state-of-the-art and competitive performance respectively. Notably, the segmentation results for the whole tumor (WT) dice similarity coefficient (DSC) reached 89.27%, surpassing the performance of existing methods. The analysis suggests that the ensemble approach offers potential benefits, and the adversarial co-training contributes to the increased robustness and accuracy of EACNet for brain tumor segmentation of MRI scans with missing modalities. The experimental results show that EACNet has promising results for the task of brain tumor segmentation of MRI scans with missing modalities and is a better candidate for real-world clinical applications.

Key words: deep learning; magnetic resonance imaging (MRI); medical image analysis; semantic segmentation; segmentation accuracy; image synthesis

0 Introduction

Brain tumors pose a significant and ongoing public health challenge worldwide. These complex and life-threatening conditions affect millions of people annually, cutting across all ages, races, ethnicities, and genders[1]. In the United States alone, brain tumors were estimated to affect over 25 400 individuals in 2023, with nearly 18 800 succumbing to the disease[2]. Magnetic resonance imaging (MRI) is a versatile imaging tool with broad applications in medical diagnostics, including ophthalmology. Its ability to produce detailed, non-invasive images makes it indispensable for detecting and monitoring various conditions throughout the body. In oncology, MRI plays a crucial role in identifying tumors and tracking the effectiveness of cancer treatments[2]. Computer-aided diagnosis (CAD) has emerged as a valuable tool in ophthalmology, leveraging machine learning and deep learning techniques to assist clinicians in disease diagnosis. To streamline this process, the development of efficient deep learning models is essential. In practice, MRI scans often encounter the challenge of missing modalities, which can hinder the accurate segmentation and localization of tumor regions[3].

Addressing the issue of missing modalities is crucial for improving the diagnostic capabilities of MRI in clinical settings. By recovering missing modalities, we can enhance the quality and reliability of tumor segmentation, ultimately leading to more accurate and timely diagnoses. This further highlights the critical need for advancements in this area. Early detection and accurate localization of brain tumors are critical for timely intervention and personalized treatment planning, leading to improved patient outcomes and survival rates[3]. MRI plays a vital role in brain tumor diagnosis, providing detailed anatomical information of the brain[4]. Multi-modal MRI scans are used to integrate complementary features from different sequences, which is particularly useful when some data (modalities) are missing. However, acquiring a complete set of MRI modalities T1, contrast-enhanced T1-weighted (T1ce), T2 and fluid attenuated invension recovery (FLAIR) can be time-consuming due to lengthy scan protocols or impractical due to patient limitations (e.g., claustrophobia) or cost constraints[5].

In recent years, deep learning has demonstrated remarkable success in various computer vision tasks, including image recognition, object detection, and segmentation. Leveraging the power of deep neural networks, it has shown promising results in brain tumour segmentation, surpassing traditional methods in accuracy and robustness[6]. Various MRI modalities offer complementary information. For instance, T1 images excel at revealing anatomical structures, while T1ce, T2 images and FLAIR sequences are better suited for visualizing tumors due to their superior contrast between tumors and surrounding tissues[7]. The absence of any modality can hinder the ability to definitively identify the presence or absence of a tumor.

Existing works[7-10] that addressed missing modalities demonstrated promising results in handling missing modalities for brain tumor segmentation. However, most of these existing works[11-14] primarily focus on scenarios with a single missing modality. Real-world clinical settings encounter cases with multiple missing modalities, potentially impacting the effectiveness of models.

The proposed ensemble adversarial co-training neural network (EACNet) framework addresses the challenge of missing MRI modalities in brain tumor segmentation. By employing an ensemble of pre-trained models, co-training, and adversarial learning, the method demonstrates robust performance and high accuracy. As illustrated in Fig.1, the framework utilizes four modalities of MRI scans to handle missing data. The ensemble approach leverages the strengths of multiple pre-trained models, enhancing the models' robustness and generalization ability. Co-training allows the model to learn from different subsets of available modalities, while adversarial learning ensures that the model can effectively handle variations in missing data patterns. Rigorous evaluations on the BraTS challenge datasets validate the effectiveness of the EACNet framework, demonstrating its capability to accurately segment brain tumors even in the presence of missing modalities.

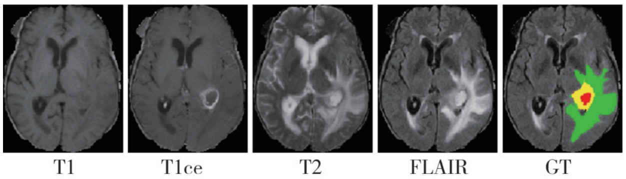


Fig. 1 Four modalities of MRI images and ground truth label

1 Related works

The works conducted by Shen et al.[13] developed a brain tumor segmentation algorithm resilient to missing modalities, employing a channel-independent encoding path and a feature-fusion decoding path. Their approach incorporates self-supervised training via channel dropout and introduces a novel domain adaptation technique on feature maps to compensate for absent channels. Findings reveal varying segmentation quality based on the missing modality. Meanwhile, Feng et al. [14] proposed a deep learning model for brain tumor segmentation using a modified 3D U-Net architecture with adjustments in training and testing strategies, network structures, and model parameters. Instead of selecting a single optimal model, they utilized an ensemble of multiple models trained with varying hyper-parameters to mitigate random errors and enhance performance.

Initial findings indicate the efficacy of this approach, leading to the 9th place ranking in the fiercely competitive 2018 Multimodal Brain Tumor Segmentation BraTS Challenge. However, their method is computationally expensive. Furthermore, the method proposed by Liu et al. [15] introduced a mixture of experts and a semantic-guided network for brain tumor segmentation with missing modalities. Their model employs a transformer-based encoder with modality-specific feature learning through mixture-of-experts blocks, aided by learnable modality embedding. Additionally, a semantic-guided decoder enhances attention to tumor regions. Extensive experiments on the BraTS2018 dataset demonstrate superior segmentation performance, achieving top or near-top results in most cases with lower computational costs. The model accurately segments brain tumor sub-regions, indicating a promising potential for clinical applications.

Similarly, the work conducted by Jeong et al. [16], employed a training technique called adversarial learning to create features for missing information based on complete data. They combines these features using an "attention" mechanism to create a single, unified representation. Additionally, Li et al.[17] introduced a local-global modelling module to improve intramodal feature representation in the encoder. To address irregular tumor shapes, a deformation-adaptive perceptual multimodal representation learning module is developed, learning deformation information from incomplete image sets for accurate tumor localization. Furthermore, a reconstruction-driven key information mining module aids in extracting essential tumor discriminate features. During inference, this module is omitted to reduce computational load. Experimental results on two benchmark datasets demonstrate superior performance compared to existing methods for brain tumor segmentation with missing modalities. Yuan. et al. [18] proposed a unified generative adversarial network (GAN) featuring pairwise modality shared feature disentanglement. Their approach incorporates a multi-pooling feature fusion module to merge features from all modalities, complemented by a distance loss and a margin loss to ensure feature symmetry regularization. Demonstrating superior performance, their model surpasses existing state-of-the-art methods for missing modality completion, particularly enhancing generation quality and aiding brain tumor segmentation, especially in regions requiring detailed information from multiple modalities for accurate diagnosis.

Besides the success shown by these methods, many of the methods, such as ensemble learning and transformer-based models, are computationally intensive. This can limit their practical deployment in real-time clinical settings where speed and efficiency are critical. Furthermore, as medical imaging technology advances, new modalities and imaging techniques emerge. Adapting existing models to incorporation and effectively utilizing these new modalities while maintaining performance and accuracy is an ongoing challenge. Handling missing modalities requires robust generalization across different datasets and clinical scenarios. Variability in data quality, imaging protocols, and tumor characteristics can affect the model's performance, and reliability is still a challenge for these methods.

2 Methods

This section highlights the structure of the proposed framework for brain tumor segmentation, the network aims to handle missing modalities in MRI scans which normally contain some missing data. The general framework utilizes adversarial learning and a co-trained approach to handle missing modalities. Fig. 2 illustrates the framework of the proposed method, detailing the components integrated within the EACNet architecture.

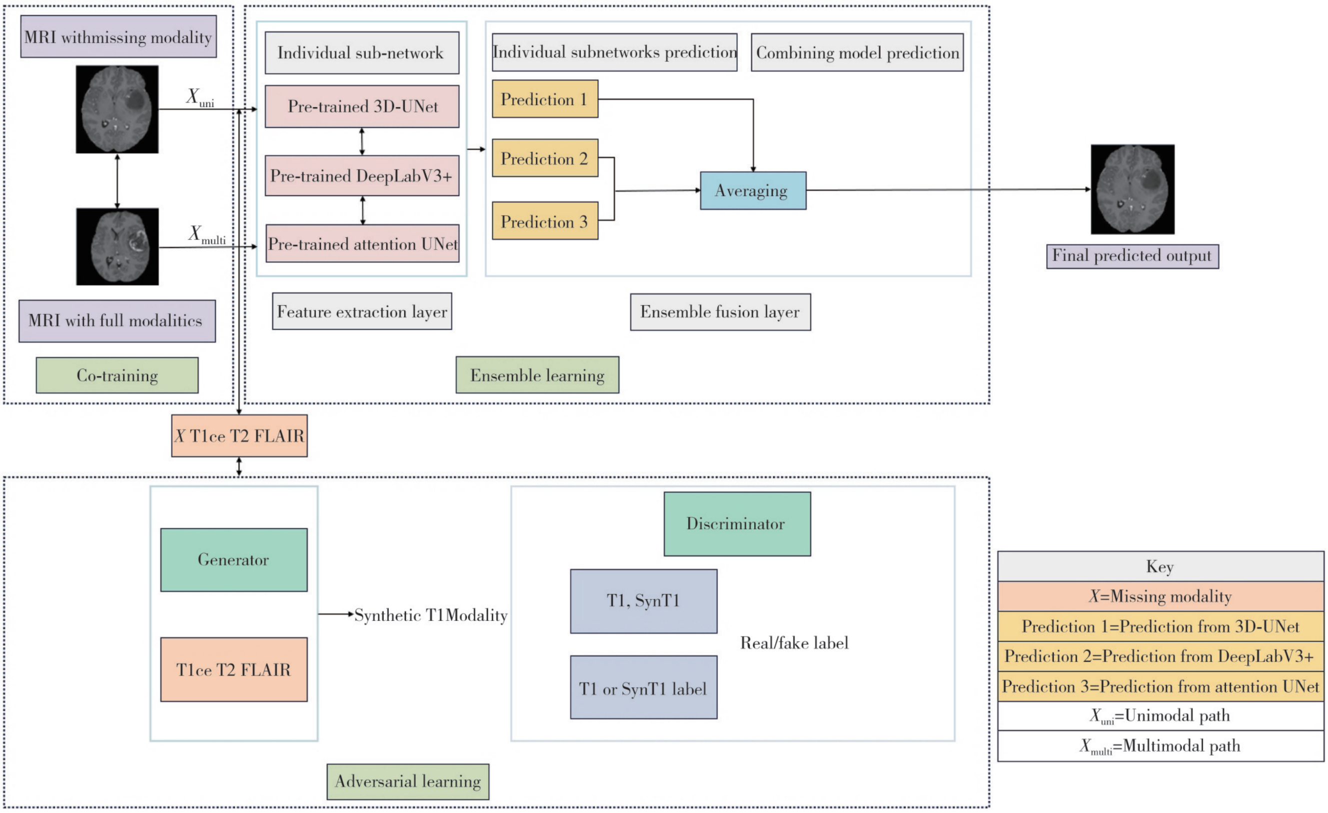


Fig. 2 Overall architectural framework of ensemble adversarial co-training network (EACNet) for brain tumor segmentation

Fig.2 provides a clear overview of the core modules and their interactions. Each component is designed to enhance the network's ability to handle missing MRI modalities for accurate brain tumor segmentation.

2.1 Overview

EACNet's architecture is designed to tackle the complex problem of brain tumor segmentation in MRI scans, specifically addressing the frequent issue of missing modalities. The approach is divided into three interlinked modules, each contributing to the overall effectiveness of the segmentation process.

1) Co-training module

This module is crucial for handling incomplete MRI data. It employs two separate processing paths: one for complete scans, and another for incomplete scans. Each path utilizes pre-trained deep learning models, which have been trained on diverse MRI datasets.

The complete scan path provides a baseline for generating high-quality soft segmentation maps, indicating the probability of each pixel belonging to specific tissue classes, such as tumor or healthy tissue. For incomplete scans, the model generates initial soft segmentation maps based on available modalities. This dual-path strategy allows the model to maintain segmentation performance even when certain modalities are missing, thereby accommodating real-world variability in MRI data.

2) Ensemble learning module

This module enhances the segmentation accuracy by integrating predictions from various pre-trained models, each with distinct architectures and training conditions. By combining outputs from these diverse models, EACNet captures a broader spectrum of features and patterns present in the MRI data.

The ensemble approach not only mitigates the risk of overfitting to any single model but also increases the robustness of the segmentation results. This is particularly beneficial in medical imaging, where variations in data can significantly impact the reliability of predictions.

3) Adversarial learning module

The adversarial learning component introduces a discriminator that evaluates the entropy maps generated from soft segmentation maps. The entropy maps provide insight into the confidence of the model regarding pixel classifications.

In this setup, the incomplete scan path functions as a generator, aiming to produce entropy maps that closely resemble those from the complete scan path. The discriminator's role is to differentiate between the two sets of entropy maps, guiding the generator to improve its outputs. This adversarial training fosters a stronger learning process, allowing the incomplete scan path to refine its predictions and better emulate the performance of the complete scan path.

The training strategy of EACNet employed a multi-stage approach, facilitating seamless integration of the three modules as follows.

1) The co-training module operates first, where pre-trained models analyze both complete and incomplete scans to produce initial soft segmentation maps. This foundational step sets the stage for further refinement.

2) The ensemble learning module combines predictions from the previous stage to create a more accurate refined segmentation map. This collective intelligence enhances the model's understanding of the data.

3) Finally, adversarial learning is used to improve the performance of the incomplete scan path. Through iterative training, the model learns to produce outputs that mimic those of the complete scan path, thus enhancing segmentation accuracy.

Overall, EACNet's robust design allows it to achieve reliable brain tumor segmentation even in the presence of missing MRI modalities, marking a significant contribution to medical imaging and deep learning.

2.2 Co-training approach

This section describes the co-training approach within EACNet for brain tumor segmentation with missing modalities in MRI scans. Co-training is a technique that is used when data have multiple perspectives or features. Each perspective is trained separately, and together they offer a more complete understanding. In the same way, combining different brain MRI scans gives a full picture of the brain's structure and function. The co-training approach in this context involves training two separate models.

Inspired by the work of Wang et al.[8], we denote X_{uni} as an incomplete model input, representing a subset of MRI modalities available for a given sample; X_{multi} as full model input, containing all MRI modalities for a sample; and X_{g} as ground truth labels, indicating the correct segmentation for each voxel in the image.

In terms of co-training with incomplete modalities, we trained two separate models: full modality model (FMM) trained on X_{multi} and X_{g}, incomplete modality model (IMM) trained on X_{uni} and X_{g}, and this was done concurrently (co-training).

Via concurrent training, we enable the model IMM to learn complementary information from the full and incomplete modalities. By leveraging the strengths of both models through concurrent training process, segmentation performance is improved even when the model faces missing data in the MRI scans.

Both FMM (multimodal path) and IMM (unimodal path) are passed to the common backbone architecture of pre-trained models of U-Net, 3D-UNet, and Deeplabv3+ to extract generic features from the MRI data, which helps to ensure consistency and reduce the computational overhead.

As for the framework of co-training block, it takes MRI data as input, which are either complete or incomplete. Complete MRI data (X_{multi}) include all available modalities for a given MRI scan, typically T1, T2, T1ce, and FLAIR. Incomplete MRI data (X_{uni}) contain only a subset of the available modalities, with at least one modality missing.

In the co-training block consisting of FMM and IMM, FMM processes the complete MRI data (X_{multi}) by extracting features from all modalities and learning a comprehensive representation of the brain tumor, and IMM processes the incomplete MRI data (X_{uni}) for the purpose of handling missing modalities and learning from the available information.

For FMM and IMM, each model has modality-specific convolutional layers, processing incomplete MRI scans with incomplete modality and MRI scans with complete modality effectively. These layers were designed to capture the unique characteristics of each modality while handling missing information. Dilated convolutions were applied to capture long-range dependencies in incomplete MRIs.

IMM and FMM were designed to incorporate modality-specific layers which are crucial for capturing the unique characteristics of different modalities and handling missing data effectively. Also, modality-specific layers were designed to focus on the available modalities and reduce the impact of missing data in MRI. A dilation rate was employed to determine the spacing between the filter elements.

In this setup, a dilation rate of 3 was used to expand the receptive field, a 3×3 filter to reduce computational complexity, 32 filters per convolutional layer to balance model capacity, and ReLU activation to introduce non-linearity.

The output from the co-training block is segmentation predictions (soft segmentation maps) denoted as Y_{multi} and Y_{uni} resulting from FMM and IMM respectively, and they are expressed as

$$Y_{\text{multi}} = O_{\text{FMM}}(X_{\text{multi}}), \tag{1}$$

$$Y_{\text{uni}} = O_{\text{IMM}}(X_{\text{uni}}). \tag{2}$$

2.3 Adversarial learning

The EACNet framework shown in Fig. 2, designed for brain tumor segmentation with multimodal missing modalities in MRI scans, is further augmented by incorporating an adversarial training component.

Adversarial training hinges on competitively training two models, one model is the generator G, which creates realistic representations of missing modalities based on the available data in each sub-network's training scenario; the other model is the discriminator, which acts as the sub-networks themselves in this case, and aims to distinguish between genuine missing modalities and the generator's synthetic versions.

In our setup, the inputs of adversarial learning are segmentation maps of IMM and FMM (Y_{uni} and Y_{multi}). The aim is to allow generator G to learn and create realistic synthetic representations of the missing modality that can fool the corresponding sub-network acting as the discriminator. Each sub-network incorporates a generator-discriminator pair for adversarial training. The generator creates synthetic data (synthetic T1 images for a network handling missing T1 data) to deceive the discriminator. The discriminator, in turn, strives to distinguish real missing modality data from the synthetic data. This iterative process helps sub-networks better handle unseen missing modality patterns during testing, leading to a more robust EACNet framework.

2.3.1 Network structure of discriminator

Input: Ground-truth segmentation mask (real) and predicted segmentation mask generated by the FMM or IMM (fake).

Architecture: The discriminator is based on convolutional layers, and a PatchGAN allows focusing on local patches of the image rather than the entire image.

Convolutional layers: Multiple layers of convolutions are used to extract features from the input soft segmentation mask, and these layers are followed by batch normalization and Leaky ReLU activation.

Patch-based output: The discriminator produces a probability map (patch-level) where each value corresponds to whether a particular patch of the input is real or fake. Patch-based output is a probability score for each patch (or the whole segmentation mask) where the output closer to 1 means real and 0 means fake.

Loss functions: Discriminator loss is evaluted by binary cross-entropy (BCE) loss for classifying real vs.

fake segmentations. The discriminator's role is to provide feedback by correctly distinguishing between real and generated augmentations, thus improving the generator over time.

2.3.2 Adversarial training loop

In the adversarial training component of the co-training framework, the combined predictions from the FMM and IMM (Y_{multi} and Y_{uni}) were used as input to a discriminator network. The discriminator's task is to distinguish between the segmentation predictions generated by FMM and IMM. This forces the models to learn complementary information and improve their ability to generate accurate and diverse predictions.

The generator and discriminator networks were trained iteratively. The generator aims to produce more realistic synthetic T1-weighted images to deceive the discriminator, while the discriminator aims to accurately distinguish between real and synthetic T1-weighted images. We employed concepts from GANs and deep learning, specifically from works conducted by Chen et al.[19], Sahoo et al.[20], and Kalejah et al.[21].

For generator operation, generator network G takes input modalities from IMM and FMM (T2, T1ce, FLAIR images) and generates synthetic images as

$$Synthetic\ I_{\text{T1}} = G(I_{\text{T2}}, I_{\text{T1ce}}, I_{\text{FLAIR}}), \tag{3}$$

where I_{T1} stands for the synthetic image, I_{T2} stands for T2-weighted image, I_{T1ce} stands for the T1ce-weighted image, and I_{FLAIR} stands for the FLAIR-weighted image.

In terms of discriminator operation, the discriminator network D receives both real T1-weighted images and synthetic T1 images, and then predicts whether the input is real or fake. The discriminator's output can be represented as a probability distribution over real or fake classes, namely

$$p_D = D(Real\ I_{\text{T1}} / Synthetic\ I_{T1}). \tag{4}$$

The discriminator's output is interpreted as the probability that the input is a real T1-weighted image. During training, D learns to maximize this probability for real images and minimize it for synthetic ones. The adversarial training process optimizes both the generator and discriminator networks iteratively. The generator aims to produce synthetic T1 images that are indistinguishable from real ones, while the discriminator aims to correctly classify real and synthetic images. The generator's objective function G is formulated as

$$G = -\log(D(synthetic\ I_{\text{T1}})). \tag{5}$$

The generator aims to minimize this objective, i.e., it tries to generate synthetic T1 images that have high probability of being classified as real by the discriminator. Similarly, the discriminator's objective function is formulated as a binary cross-entropy loss:

$$L_{\text{Cross-entropy}} = -\log(Real\ I_{\text{T1}}) - \log(1 - D(synthetic\ I_{\text{T1}})). \tag{6}$$

The discriminator aims to minimize this objective, i. e., it tries to correctly classify real and synthetic images. During training process, the parameters of generator and discriminator networks are updated using gradient descent. The generator updates its parameters to minimize G, while the discriminator updates its parameters to minimize $L_{\text{Cross-entropy}}$.

2.4 Ensemble learning

As for ensemble approach within the EACNet framework shown in Fig.2, EACNet utilizes an ensemble architecture consisting of multiple sub-networks. Each sub-network within the EACNet architecture incorporates a pre-trained model specifically selected to address unique missing modality challenges.

Selection criteria for sub-networks involves three criteria as follows:

1) Capturing local and contextual information. It is crucial to capture spatial dependencies and local features. 3D U-Net excels at this due to its U-shaped architecture.

2) Multi-scale contextual awareness. DeepLabV3+ with atrous convolutions is well-suited for understanding broader context in scans, especially important with missing information.

3) Focusing on critical regions. Attention U-Net excels at directing focus to these crucial regions using attention mechanisms. The final prediction for brain tumor presence is derived by combining the outputs from all sub-networks. Each sub-network independently processes the MRI scans and generates individual predictions, focusing on learning complementary features from available modalities. This approach aims to extract diverse and informative features even with incomplete data.

Afterwards, the predictions from each sub-network are aggregated using an averaging approach. Let p_i represent the probability of brain tumor presence predicted by the ith sub-network within the ensemble, and N is the total number of sub-networks within the EACNet ensemble. p_{ensemble} presenting ensemble prediction for brain tumor presence is calculated as the average probability across all sub-networks.

In the co-training framework of EACNet, the pre-trained models serve as the shared backbone for both the FMM and the IMM. This means that the same set of pre-trained models are used to extract features from both complete and incomplete MRI data.

Feature extraction process is as follows: The ensemble receives the complete and incomplete MRI data as input. Each pre-trained model within the ensemble processes the input data independently. Each model extractes its own set of features using its internal architecture (e.g., convolutional layers, pooling layers, attention mechanisms). We set model parameters in each model as described in subsection 3.2 including pooling, maxpooling, convolutions, down-sampling and up-sampling layers in U-Net.

Then, ensemble fusion is performed. The individual model predictions of U-Net, 3D U-Net, and DeepLabV+ in the ensemble block are denoted as p_1, p_2 and p_2, respectively.

For the unimodal path, there is

$$p_{\mathrm{uni}}^{n\mathrm{th}} = X_{\mathrm{uni}}(m_{\mathrm{indi}}). \tag{7}$$

For multimodal path, there is

$$p_{\mathrm{multi}}^{n\mathrm{th}} = X_{\mathrm{multi}}(m_{\mathrm{indi}}), \tag{8}$$

where $P^{n\mathrm{th}}$ is the prediction corresponding to each model in the ensemble block. The individual network (pre-trained model) predictions are fused via averaging techniques to obtain the final result as

$$p_{\mathrm{final}} = \frac{1}{N} p_i. \tag{9}$$

3 Experiment and results

The BraTS2018 and BraTS2020 training datasets[21,22] consist of 285 multi-contrast MRI scans with four modalities: native (T1), post-contrast (T1ce), T2-weighted (T2), and T2 FLAIR volumes. Each of these modalities captures different properties of brain tumor sub-regions which are the GD-enhancing tumor (ET), the peritumoral oedema (ED), and the necrotic and non-enhancing tumor core (NCR/NET).

These sub-regions are combined into three nested sub-regions, WT, tumor core (TC), and ET. All the volumes have been co-registered to the same anatomical template and interpolated to the same resolution.

Experiments were implemented in Pytorch and performed on a computer with memory 12 GB of GPU of NVIDIA GTX 1080 and a processor of Intel i7 with 2.5 GHz clock speed 12th generation family. The training process consisted of 300 epochs in each run with an input patch size of 160×192×128 and a batch size of 1. The Adam optimizer was applied, and the initial learning rate was set as 0.000 1 and progressively decreased according to a poly policy as

$$\eta_t = \eta_0 \left(1 - \frac{N_{\mathrm{epoch}}}{N_{\mathrm{epoch,\,max}}}\right)^{0.9}. \tag{10}$$

We randomly split the dataset into training and validation sets at a ratio of 2 : 1. The detection performance was evaluated on each nested sub-region of brain tumors using "The Dice similarity coefficient (DSC)" and "Hausdorff distance (HD95)". A higher DSC and a lower HD95 indicate a better prediction performance.

3.1 Results

This section presents experimental results of EACNet in segmenting brain tumors on MRI scans with missing modalities. The results are expressed in terms of DSC scores and HD95, key metrics for segmentation accuracy, reported across various regions and modalities.

3.1.1 Quantitative results on BraTS2018

Experimental results on BraTS2018 dataset show that the EACNet achieves a competitive performance on DSC in ET, WT, and TC regions. For ET, this region represents the most critical part of the tumor, often exhibiting aggressive growth. High and consistent DSC across modalities for EACNet indicate its ability to accurately segment this vital area, even when data are missing. WT encompasses the entire tumor volume. The strong performance of 83.67% in this region suggests that EACNet effectively captures the overall tumor size and shape, which is crucial for treatment planning and monitoring. For TC, this region denotes the central, most malignant part of the tumor. High DSC in this region showcases EACNet's ability to precisely delineate the core, which is essential for guiding treatment decisions.

Table 1 shows the results of the proposed method in DSC on ET, WT, and TC regions of MRI scans based on each modality, in the table missing modalities are denoted by × and modalities present by √. EACNet demonstrates superior performance in all three MRI regions (ET, TC, and WT) for brain tumor detection with complete MRI data. This highlights the effectiveness of the ensemble approach and the potential benefits of adversarial co-training.

Table 1 Results of EACNet on BraTS2018 in DSC

Modality				DSC/%			
FLAIR	T1	T1ce	T2	ET	TC	WT	Average
×	×	×	√	45.67	69.95	69.95	68.45
×	×	√	×	80.91	85.51	86.34	83.67
×	√	×	×	44.47	83.36	84.56	68.89
√	×	×	×	46.39	65.34	89.48	68.89

Table 2 shows the results of the proposed method in HD95 on ET, WT, and TC regions of MRI scans based on each modality, in the table missing modalities are denoted by ×, and modalities present by √.

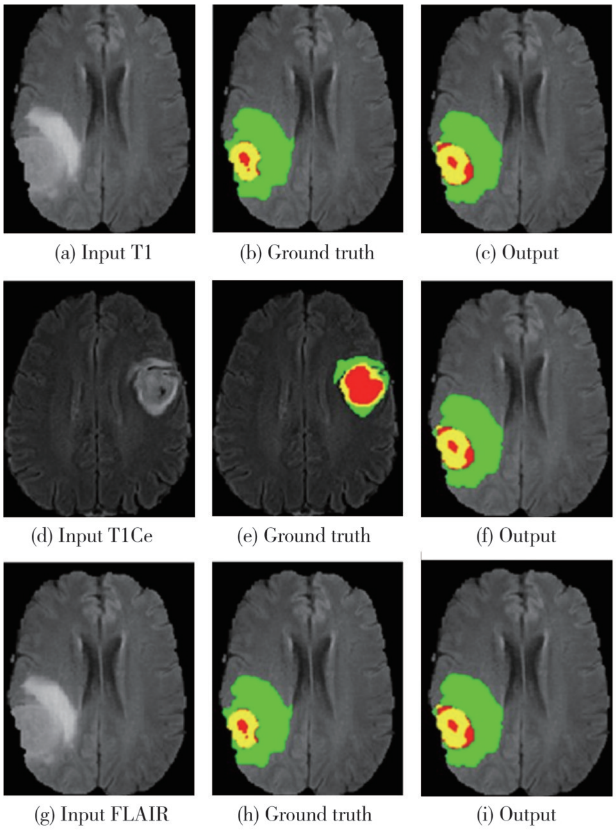
Table 2 Results of EACNet on BraTS2018 in HD95

Modality				HD95/mm			
FLAIR	T1	T1ce	T2	ET	TC	WT	Average
×	×	×	√	7.21	5.62	6.98	6.45
×	×	√	×	8.43	6.78	7.62	6.67
×	√	×	×	5.62	5.94	6.76	6.89
√	×	×	×	5.95	7.75	7.75	6.86

EACNet also demonstrates superior performance in all three MRI regions (ET, TC, and WT) for brain tumor detection with complete MRI data. This also highlights the effectiveness of the ensemble approach and the potential benefits of adversarial learning.

3.1.2 Qualitative results on BraTS2018 dataset

Fig.3 presents segmentation results from the EACNet model with ground truth data across different MRI modalities.



(a) Input T1 (b) Ground truth (c) Output

(d) Input T1Ce (e) Ground truth (f) Output

(g) Input FLAIR (h) Ground truth (i) Output

Fig. 3 Segmentation results of EACNet

Fig. 3 includes three columns: (a) input modality, (b) ground truth annotated by experts, and (c) segmented output from EACNet.

Input modality represents the raw MRI images presented for segmentation. This column furthermore displays the original MRI images that serve as input to the segmentation process, providing critical contextual information necessary for accurate tumor delineation. For ground truth, this column of figures present expert-annotated segmentations that serve as the benchmark for evaluating accuracy. These annotations reflect precise and clinically verified delineations of tumor boundaries, which are crucial for performance assessment. For EACNet segmented output, this column presents the predicted tumor regions generated by the EACNet model. These outputs represent the model's capacity to segment and identify tumor boundaries with precision, closely mirroring the expert ground truth segmentations. This visual comparison highlights EACNet's effectiveness in accurately delineating tumor regions, showcasing a strong correspondence between EACNet's outputs and the expert-annotated ground truths across multiple MRI modalities, and demonstrating strong alignment with the ground truth across various MRI inputs. It underscores EACNet's capability to generalize well even under diverse and complex input conditions. The comprehensive comparison offers compelling evidence of EACNet's performance, showcasing its consistency, adaptability, and precision in handling multimodal input data while maintaining high segmentation accuracy.

3.1.3 Quantitative results on BraTS2020 dataset

EACNet was further rigorously evaluated on more complex BraTS2020 dataset, which includes a validation set comprising 50 additional training samples and 59 additional validation samples. The incorporation of this expanded dataset led to a notable enhancement in segmentation performance.

Experimental results show that EACNet achieves the highest DSC for the ET sub-region, demonstrating superior delineation of this critical tumor component. Competitive DSC scores are also attained for the WT and TC sub-regions.

Furthermore, EACNet exhibits the best HD95 across all tumor types, indicating improved localization accuracy. These results underscore the effectiveness of EACNet in handling the intricacies of brain tumor segmentation and its potential for clinical translation.

Table 3 below shows that the inclusion of different modalities significantly influences the segmentation performance. The combination of FLAIR, T1ce, and T2 modalities (row 1) yields the best overall DSC scores, indicating the importance of these modalities for accurate brain tumor segmentation. The T1ce modality appears to be crucial for achieving high DSC scores, especially for the ET sub-region. Removing T1ce (rows 2 and 3) leads to a substantial drop in ET DSC, highlighting its significance in tumor delineation. While removing individual modalities (T1 or T2) degrades performance compared to the full modality setup, the EACNet model still demonstrates reasonable results, suggesting its robustness to missing data scenarios.

Table 3 Results for EACNet on BraTS2020 in DSC

Modality				DSC/%			
FLAIR	T1	T1ce	T2	ET	TC	WT	Average
×	×	×	√	88.21	69.95	79.27	78.45
×	×	√	×	70.91	75.51	76.34	73.67
×	√	×	×	71.45	67.36	74.56	78.89
√	×	×	×	77.51	75.34	79.48	78.89

EACNet exhibits a relatively balanced performance across different tumor sub-regions (ET, TC, WT), indicating its ability to segment various tumor components accurately. EACNet achieves promising results on the BraTS2020 dataset, particularly when all modalities are available. The model demonstrates a certain level of robustness to missing modalities, although performance is generally better with more complete data. These findings highlight the potential of EACNet as a valuable tool for clinical applications in brain tumor segmentation.

Table 4 presents the model's segmentation results in the HD95 metric for different MRI modalities used in tumor segmentation. The modalities assessed include FLAIR, T1, T1ce, and T2. The table highlights the segmentation performance for different tumor regions, ET, TC, and WT, along with an overall average HD95 across these regions. A checkmark (√) indicates the modality used for a specific test configuration, while a cross (×) denotes its exclusion. The values show the accuracy of tumor boundary predictions, where lower HD95 values indicate better segmentation performance.

Table 4 Results of EACNet on BraTS2020 in HD95

Modality				HD95/mm			
FLAIR	T1	T1ce	T2	ET	TC	WT	Average
×	×	×	√	5.21	6.92	6.98	6.65
×	×	√	×	7.34	7.18	7.62	6.67
×	√	×	×	6.62	7.94	8.76	7.89
√	×	×	×	6.45	6.95	7.98	6.86

Furthermore, the table compares the effect of different modality combinations on the segmentation results, helping evaluate the impact of each modality on the overall accuracy.

3.2 Ablation study

3.2.1 Segmentation results of several models

This section investigates the effectiveness of each proposed module within the EACNet framework for brain tumor segmentation when dealing with a single available modality (T1ce).

The primary goal of this study is to assess the individual contributions of different components within the EACNet framework to the overall segmentation performance. The first row establishes a baseline performance using only the L_{con} loss function, without incorporating any additional modules such as adversarial learning (AM), co-training (CM), or ensemble learning (EM). A co-training baseline network is established that solely utilizes the consistency loss (L_{con}) function. This baseline serves as a reference point for evaluating the impact of the proposed modules. Each proposed module was progressively added to the baseline network to analyze their individual and combined contributions.

Table 5 presents the results of the study. The results show that the performance of the unimodal path, responsible for handling the missing modality (all modalities except T1ce), improves steadily with the addition of each module (AM, CM, and EM). This observation suggests a mutually beneficial relationship between the co-training approach and the proposed modules. Each module contributes to an overall improvement in the network's ability to handle the missing modality.

Table 5 Segmentation results in DSC from an ablation study

Model	Modality	DSC/%			
		ET	TC	WT	Average
L_{Con}+AM	T1ce	76.56	85.71	79.23	80.50
L_{Con}+CM		78.23	82.90	78.98	80.04
L_{Con}+EM		78.34	84.45	79.34	81.71
L_{Con}+CM+EM		77.45	85.96	79.56	81.04
L_{Con}+CM+EM+AM		79.79	85.98	91.97	81.78

When an AM is added to the model, a consistent improvement in all evaluation metrics (ET, TC, WT, and Average) is exhibited, demonstrating its effectiveness in enhancing segmentation performance.

Furthermore, when a CM is added, the segmentation accuracy is further increased, particularly for the ET and WT metrics. This suggests that the co-training strategy, which leverages information from both complete and incomplete data, contributes significantly to the model's accuracy.

The addition of the EM further improves the segmentation accuracy, especially for the ET and TC metrics. This indicates that combining multiple models can enhance the model's ability to capture diverse features and improve segmentation accuracy.

The final row shows the performance of the full EACNet model, incorporating all three modules. This configuration achieves the best overall performance across all evaluation metrics, highlighting the synergistic effect of combining the different components. The ablation study results demonstrate the effectiveness of each module within the EACNet framework.

The adversarial, co-training, and ensemble learning modules all contribute to improved segmentation performance. The combination of these modules in the full EACNet model yields the best overall results, emphasizing the importance of a comprehensive approach to address the challenges of brain tumor segmentation.

3.2.2 Evaluation of unimodal vs. multimodal path contributions in EACNet segmentation

To assess the impact of both unimodal and multimodal paths on the segmentation performance of EACNet, we extended our ablation study by conducting experiments designed to isolate and evaluate the contributions of each approach. This comprehensive evaluation sought to determine how the fusion of multiple MRI modalities influences segmentation accuracy, compared to processing single modalities independently.

1) Unimodal path analysis

In this study, EACNet was firstly trained using each MRI modality separately—T1, T1ce, T2, and FLAIR. By isolating each modality, we aimed to understand how EACNet effectively segments brain tumors when it relies solely on the features extracted from a single input source. This process offered valuable insight into the unique contribution and discriminative power of each modality. For instance, certain modalities, like FLAIR, are known to better highlight edema regions, while T1ce images provide more defined tumor core boundaries. By quantifying the segmentation performance for each unimodal input, we were able to observe the strengths and weaknesses of using each modality in isolation.

2) Multimodal path analysis

Next, the segmentation performance was examined when all modalities were fused and processed simultaneously by EACNet. This approach leverages complementary information present across different MRI sequences. For example, integrating data from T1, T1ce, T2, and FLAIR allows the model to gain a more comprehensive understanding of both the tumor and surrounding tissues, leading to enhanced segmentation accuracy and robustness. This multimodal approach aligns with clinical practices where combining various imaging modalities often yields better diagnostic accuracy than any single modality alone.

3) Results comparison and key findings

The results of the study demonstrated that the multimodal path substantially outperformed the unimodal paths across key segmentation metrics, such as DSC, precision, and recall. This improvement can be attributed to the ability of the multimodal approach to capture richer and more diverse features, thereby enhancing the model's ability to generalize and identify complex tumor structures. Conversely, while some unimodal paths (notably FLAIR and T1ce) offered relatively strong segmentation results, they were consistently outperformed by their multimodal counterpart due to the lack of cross-modal feature fusion.

Fig. 4 below presents an ablation study that evaluates the segmentation performance using different input modalities and approaches of the unimodal and multimodal paths, including the full EACNet model, for brain tumor segmentation in MRI images. The results are displayed in multiple columns, with each row corresponding to different input MRI sequences. The detailed description is as follows.

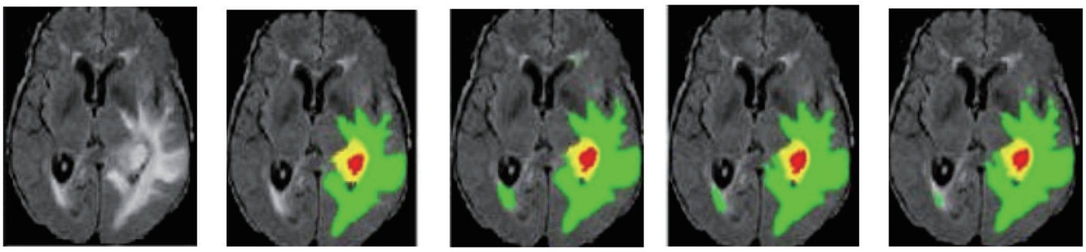
1) Input MRI modalities (column 1). Images (a), (f), and (k) show the input MRI scans in T1, T2, and FLAIR modalities, respectively. These inputs represent the baseline data used for segmentation by different models.

2) Ground truth (column 2). Images in the subfigures labelled (b), (g), and (l) illustrate the ground truth segmentation labels, highlighting different tumor regions such as the necrotic core, and enhancing tumor tissue with distinct colours for accurate visual comparison. Columns 3－5 visualize segmentation results for multimodal and unimodal paths.

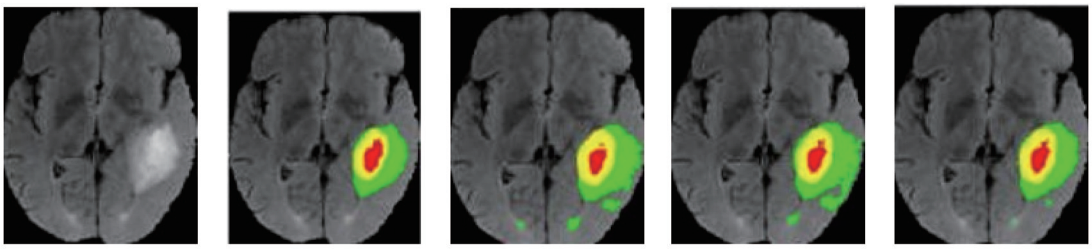
3) Columns with subfigures labelled (c), (d), (h), (i), (m), and (n) depict the segmentation results from unimodal and multimodal paths. These paths use different combinations of MRI input data and provide insights into how each approach affects the segmentation accuracy.

4) Full EACNet. Images in subfigures (e), (j), and (o) present the segmentation output of the full EACNet model. This configuration leverages all input paths and modalities, resulting in more comprehensive and precise tumor segmentation.

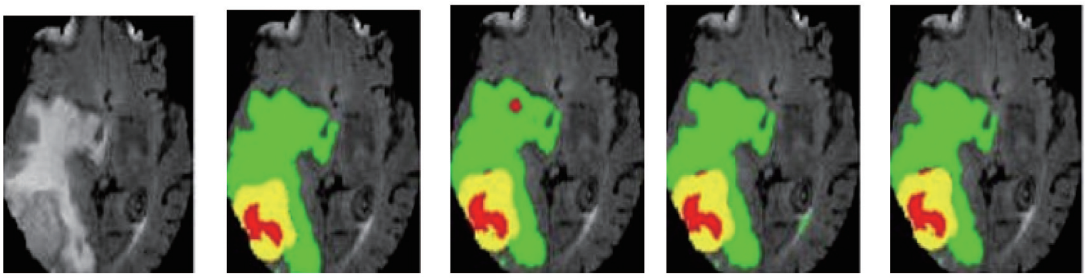
5) Key observations. The ablation study clearly demonstrates the enhanced segmentation accuracy of the full EACNet model compared to unimodal or partial combinations of modalities. In each instance, the full EACNet model (shown in images (e), (j), and (o)) offers the most accurate and complete delineation of tumor regions, reflecting the advantage of combining multiple input paths and modalities in a full EACNet model in achieving superior segmentation performance for brain tumors in multimodal MRI data.



(a) Input T1 (b) Ground truth (c) Multimodal (d) Unimodal (e) Full EACNet



(f) Input T2 (g) Ground truth (h) Unimodal path (i) Multimodal path (j) Full EACNet



(k) Input FLAIR (l) Ground truth (m) Unimodal path (n) Multimodal (o) Full EACNet

Fig. 4 Segmented results of an ablation study

3.3 Comparison with other recent methods

This section presents a comparative analysis of the proposed EACNet method with other recent methods and approaches for brain tumor segmentation with missing modalities. The primary goal of this comparison is to highlight the unique advantages and contributions of EACNet in handling multiple missing modalities and improving the overall accuracy of brain tumor segmentation.

Comparison results demonstrate that EACNet outperforms other methods in handling multiple missing modalities, achieving DSC of 89.27% for WT segmentation on the BraTS2018 dataset. This significant improvement can be attributed to the ensemble approach and adversarial co-training, which enhance the model's robustness and accuracy in handling incomplete data as shown in Tables 6—8.

Tables 6—8 show the segmentation results of various tumor regions including ET, TC, and WT, demonstrating the superior performance of EACNet on the BraTS2018 dataset. Compared to three recent methods for brain tumor segmentation, M3AE[22], SMU-Net[23], and ACN[24], EACNet achieves the highest DSC score, specifically for the WT segmentation task. This quantitative advantage highlights the effectiveness of EACNet's ensemble adversarial co-training strategy in handling missing modalities within MRI scans.

Table 6 DSC scores of various missing modalities of MRI scans in ET region

Modality				DSC/%			
FLAIR	T1	T1ce	T2	EACNet	M3AE	SMU-Net	ACN
×	×	×	√	76.5	75.3	70.6	67.8
×	×	√	×	82.1	83.4	80.4	79.8
×	√	×	×	62.5	70.8	61.8	60.6
√	×	×	×	79.9	78.6	86.7	70.8
√	×	√	√	71.82	70.9	66.9	66.7
×	√	√	×	85.43	74.2	80.4	78.9
×	√	×	√	74.2	72.2	60.4	63.9
√	√	√	×	73.98	73.8	67.2	66.1
√	×	√	×	48.90	84.7	73.3	92.6
√	√	×	√	75.52	75.7	68.2	67.2
√	×	√	√	86.26	84.8	75.5	68.3
×	√	√	√	75.75	73.8	67.2	66.1

Furthermore, superior performance in other segmentation metrics (e.g., TC and ET DSC scores) suggests that EACNet offers a robust and generalizable approach for brain tumor segmentation in challenging clinical scenarios.

Table 8 presents the segmentation results of various

MRI scan modality combinations (FLAIR, T1, T1ce, and T2) for the WT region, and compares the performance of four segmentation models: EACNet, M3AE[22], SMU-Net[23], and ACN[24].

Table 7 DSC Scores of various missing modalities of MRI scans in TC region

Modality				DSC/%			
FLAIR	T1	T1ce	T2	EACNet	M3AE	SMU-Net	ACN
×	×	×	√	76.5	75.3	70.6	67.8
×	×	√	×	82.1	83.4	80.4	79.8
×	√	×	×	62.5	70.8	61.8	60.6
√	×	×	×	79.9	78.6	86.7	70.8
√	×	√	√	71.82	70.9	66.9	66.7
×	√	√	×	85.43	74.2	80.4	78.9
×	√	×	√	74.2	72.2	60.4	63.9
√	√	√	×	73.98	73.8	67.2	66.1
√	×	√	×	48.90	84.7	73.3	92.6
√	√	×	√	75.52	75.7	68.2	67.2
√	×	√	√	86.26	84.8	75.5	68.3
×	√	√	√	75.75	73.8	67.2	66.1

Table 8 DSC scores of various missing modalities of MRI scans in WT region

Modality				DSC/%			
FLAIR	T1	T1ce	T2	EACNet	M3AE	SMU-Net	ACN
×	×	×	√	75.6	74.5	72.7	68.5
×	×	√	×	83.1	82.4	81.4	77.8
×	√	×	×	64.5	72.8	66.8	63.6
√	×	×	×	80.9	78.6	88.8	75.8
√	×	√	√	74.82	76.9	68.9	69.7
×	√	√	×	84.33	83.2	79.4	73.9
×	√	×	√	74.2	72.2	62.4	64.9
√	√	√	×	72.98	70.8	68.2	67.1
√	×	√	×	49.90	82.7	75.3	94.6
√	√	×	√	74.52	76.7	66.2	65.5
√	×	√	√	87.26	85.8	75.5	68.3
×	√	√	√	75.75	75.8	68.2	65.1

1) Impact of missing modalities

The table also shows the segmentation performance when different MRI modalities (FLAIR, T1, T1ce, T2) are either present (√) or absent (×). By observing the scores, it can be seen that the absence of certain modalities affects the segmentation performance across different models. When only T2 is present (other modalities are missing), EACNet achieves a DSC of 75.6%, outperforming all other methods, with M3AE at 74.5% and ACN lagging at 68.5%. This suggests that EACNet is more robust to missing modalities and can extract meaningful features from the remaining T2 modality. But when only T1ce is present, EACNet achieves the highest DSC score of 83.1%, followed closely by M3AE at 82.4%., and SMU-Net and ACN show lower performance, with DSC scores of 81.4% and 77.8%, respectively. When only T1 is present, EACNet underperforms compared to the M3AE method in this case, with a DSC of 64.5% (compared to 72.8% for M3AE). This suggests that EACNet is more reliant on other modalities than on T1 alone, where M3AE seems more capable of leveraging T1 data. When only FLAIR is present, SMU-Net achieves a high DSC score of 88.8%, outperforming EACNet (80.9%) and the other methods, showing that SMU-Net benefits more from FLAIR when other modalities are missing. ACN also performs relatively well here with a DSC of 75.8% but still lags behind SMU-Net.

2) Performance with multiple modalities

When T1ce, T1, and FLAIR are present, EACNet outperforms all other methods with a DSC score of 87.26%, suggesting its strength in multimodal learning. M3AE follows with a DSC of 85.8%, while ACN performs worse with 68.3%. When FLAIR, T1ce, and T2 are present, EACNet continues to dominate with a DSC of 84.33%, slightly outperforming M3AE (83.2%). and ACN and SMU-Net trail behind at 73.9% and 79.4%, respectively. Across the different missing modality scenarios, EACNet consistently achieves high scores, indicating its robustness to incomplete modalities in MRI scans. Its ability to maintain competitive performance despite missing inputs suggests that its ensemble, co-training, and adversarial learning strategies are effectively mitigating the absence of certain modalities. M3AE also shows competitive performance, especially when only one modality is missing, but falls behind EACNet when more than one modality is missing. SMU-Net appears to perform exceptionally well in certain scenarios (such as when only FLAIR is present), but overall, it struggles more than EACNet when multiple modalities are absent.

ACN shows the weakest performance across most scenarios, indicating that it is less capable of handling missing data compared to the other methods. Furthermore, EACNet shows its strength across the board, especially when dealing with multimodal inputs. Its ability to integrate information from different modalities and handle missing data proves effective. M3AE is a close competitor, particularly when fewer modalities are available, but tends to underperform slightly when there is more than one missing modality.

SMU-Net shines in cases where FLAIR is present but struggles to match the performance of EACNet when other combinations of modalities are missing. ACN consistently underperforms compared to the other methods, indicating that it may not be well-suited for handling missing modalities in this context.

EACNet delivers the most robust performance across varying levels of missing modalities, proving its value in dealing with incomplete MRI data. Its integration of ensemble learning, co-training, and adversarial learning helps it achieve competitive segmentation results even in challenging scenarios.

M3AE remains a strong competitor but falls short in certain multimodal configurations. SMU-Net has specific strengths (like FLAIR-based segmentation) but struggles more than EACNet with missing data. ACN appears less effective than the other models in handling missing modalities, consistently yielding lower DSC scores.

Finally, these results show segmentation results of various missing modalities of MRI scans in the WT, TC and ET regions, demonstrating that EACNet (our model) achieves the best results in most cases, consistently outperforming the other methods for case where one or other modality is missing in MRI scans.

Table 9 compares the DSC performance of the proposed EACNet model with three other state-of-the-art methods: M3AE, ACN, and SMU-Net. EACNet achieves a DSC of 89.27%, which is the highest among all the models compared.

Table 9 Comparison of EACNet with other recent methods

Method	DSC/%
M3AE[23]	84.8
ACN[8]	86.34
SMU-NET[24]	85.27
EACNet (Ours)	89.27

This improvement demonstrates the effectiveness of the EACNet in accurately segmenting regions of interest, likely due to its advanced ensemble and co-training mechanisms, which handle multimodal data and missing modalities more robustly. The substantial improvement of EACNet over other methods demonstrates its consistent performance across different scenarios, particularly in complex multimodal datasets with challenges such as missing or noisy modalities.

The EACNet's innovative use of ensemble adversarial co-training enhanced its ability to generalize and achieve higher segmentation accuracy. This design ensures that features are learned collaboratively from multiple modalities, even when some data are missing. EACNet exhibits significant advancements in segmentation accuracy compared to existing state-of-the-art models. The improvement in DSC reflects its capability to handle complex datasets and outperform traditional methods in multimodal medical image segmentation. These results validate the contribution of EACNet to the field and its potential for broader application.

4 Conclusions

In conclusion, our work introduces EACNet, a novel deep learning framework, for brain tumor segmentation from multi-modal MRI scans with missing modalities. EACNet leverages the power of an ensemble of pre-trained models to capture diverse feature representations and incorporates adversarial co-training to enhance robustness against data variations. By effectively combining ensemble learning, adversarial training, and co-training strategies, EACNet demonstrates superior performance compared to state-of-the-art methods on the challenging BraTS2018 and 2020 datasets. This combination is highly effective, achieving state-of-the-art performance on the BraTS2018 challenge dataset and achieving competitive performance on BraTS2020 dataset.

The model excels in handling missing modalities, maintaining competitive accuracy even in the absence of crucial imaging information while EACNet has shown promising results. However, further validation on a wider range of datasets is crucial to assess its generalizability across different clinical settings. Additionally, an ablation study provides deeper insights into the individual contributions of each module within EACNet. Exploring methods to reduce computational cost would also be beneficial for real-world implementation.

Despite these considerations, EACNet's promising results pave the way for further advancements in automated brain tumor segmentation, potentially leading to improved clinical outcomes for brain tumor patients.

Acknowledgement

This research neither utilizes any external funding sources nor is supported by a specific project grant. This research was conducted entirely within the laboratories and facilities of Jiangnan University.

Declaration of conflicting interests

The authors declare that they have no conflicts of interest to disclose.

References

[1] ILIC I, ILIC M. International patterns and trends in the brain cancer incidence and mortality: an observational study based on the global burden of disease. Heliyon, 2023, 9 (7): e18222.

[2] OSTROM Q T, PATIL N, CIOFFI G, et al. CBTRUS

statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2016-2020. Neuro-Oncology, 2023, 25(12): 90-99.

[3] OSTROM Q T, PRICE M, NEFF C, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2014-2018. Neuro-Oncology, 2021, 23(3): 61-105.

[4] MARTUCCI M, RUSSO R, SCHIMPERNA F, et al. Magnetic resonance imaging of primary adult brain tumors: state of the art and future perspectives. Biomedicines, 2023, 11(2): 3-64.

[5] ENDERS J, ZIMMERMANN E, RIEF M, et al. Reduction of claustrophobia during magnetic resonance imaging: methods and design of the “CLAUSTRO” randomized controlled trial. BMC Medical Imaging, 2011, 11(4): 4-10.

[6] KHANAL A, ESTRADA R. Dynamic deep networks for retinal vessel segmentation. Frontiers in Computer Science, 2020, 27(8): 51-64.

[7] ABDUSALOMOV A B, MUKHIDDINOV M, WHANGBO T K. Brain tumor detection based on deep learning approaches and magnetic resonance imaging. Cancers, 2023, 15(16): 41-72.

[8] WANG Y X, ZHANG Y, LIU Y, et al. ACN: adversarial co-training network for brain tumor segmentation with missing modalities//Medical Image Computing and Computer Assisted Intervention (MICCAI), September 27-October 1, 2021, Strasbourg: France. Cham: Springer International Publishing, 2021: 410-420.

[9] MA B T, WANG Y S, WANG S. Analyzing deep learning based brain tumor segmentation with missing MRI modalities. 2022: 2208.03470. https://arxiv.org/abs/2208.03470v1.

[10] TING H, LIU M. Multimodal transformer of incomplete MRI data for brain tumor segmentation. IEEE Journal of Biomedical and Health Informatics, 2024, 28 (1): 89-99.

[11] ZHOU T X, CANU S, VERA P, et al. Latent correlation representation learning for brain tumor segmentation with missing MRI modalities. IEEE Transactions on Image Processing, 2021, 30(1): 4263-4274.

[12] RAUT P, BALDINI G, SCHÖNECK M, et al. Using a generative adversarial network to generate synthetic MRI images for multi-class automatic segmentation of brain tumors. Frontiers in Radiology, 2024, 3(2): 1336902.

[13] SHEN Y, GAO M C. Brain tumor segmentation on MRI with missing modalities//The 26th Information Processing in Medical Imaging. June 2-7, 2019, Hong Kong, China. Cham: Springer International Publishing, 2019: 417-428.

[14] FENG X, TUSTISON N J, PATEL S H, et al. Brain tumor segmentation using an ensemble of 3D U-nets and overall survival prediction using radiomic features. Frontiers in Computational Neuroscience, 2020, 1(4): 1662.

[15] LIU S Y, WANG H R, LI S M, et al. Mixture-of-experts and semantic-guided network for brain tumor segmentation with missing MRI modalities. Medical & Biological Engineering & Computing, 2024, 62(10): 3179-3191.

[16] JEONG S W, CHO H H, LEE S, et al. Robust multimodal fusion network using adversarial learning for brain tumor grading. Computer Methods and Programs in Biomedicine, 2022, 226: 107165.

[17] LI Z Y, ZHANG Y F, LI H F, et al. Deformation-aware and reconstruction-driven multimodal representation learning for brain tumor segmentation with missing modalities. Biomedical Signal Processing and Control, 2024, 91: 106012.

[18] YUAN Y X, HUANG Y W, ZHOU Y. Rethinking a unified generative adversarial model for MRI modality completion//The 3rd MICCAI Workshop on Deep Generative Models. October 8, 2023, Vancouver, Canada. Cham: Springer Nature Switzerland, 2024: 143-153.

[19] CHEN H, QIN Z G, DING Y, et al. Brain Tumor Segmentation with Generative Adversarial Nets//2019 2nd International Conference on Artificial Intelligence and Big Data, May 25-28, 2019. Chengdu, China. New York: IEEE, 2019: 301-305.

[20] SAHOO S, MISHRA S, PANDA B, et al. An augmented modulated deep learning based intelligent predictive model for brain tumor detection using GAN ensemble. Sensors, 2023, 23(15): 6930.

[21] KALEJAHI B K, MESHGINI S, DANISHVAR S. Segmentation of brain tumor using a 3D generative adversarial network. Diagnostics, 2023, 13(21): 3344.

[22] LIU H, WEI D, LU D H, et al. M3AE: multimodal representation learning for brain tumor segmentation with missing modalities. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(2): 1657-1665.

[23] AZAD R, KHOSRAVI N, MERHOF D. SMU-Net: Style matching U-Net for brain tumor segmentation with missing modalities. International Conference on Medical Imaging with Deep Learning, July 6-8, 2022, Lürich, Switzerland. New York: IEEE, 2022, 1040(37): 1746-8094.

[24] WANG N, CHEN C W, MA L Z, et al. Latent feature regularization based adversarial network for brain tumor anomaly detection//2023 IEEE International Conference on Multimedia and Expo, July 10-14, 2023, Brisbane, Australia. New York: IEEE, 2023: 954-959.

EACNet:用于处理MRI图像中缺失模态的脑肿瘤分割集成对抗协同训练神经网络

RAMADHAN Amran Juma[1], 陈　璟[1,2*], 彭俊斓[1]

1.江南大学 人工智能与计算机学院,江苏 无锡 214122;
2.江南大学 江苏省模式识别与计算智能工程实验室,江苏 无锡 214122

摘　要: 脑肿瘤分割是临床诊断和治疗计划中的关键步骤。现有针对缺失模态脑肿瘤分割的方法在处理真实世界临床环境中常见的多种缺失模态时往往难以应对。这些方法主要侧重于一次处理单个缺失模态,对于包含各种缺失模态组合的缺失数据,其鲁棒性不足以应对额外的复杂性。此外,大多数现有方法依赖于单一模型,这可能会限制其性能并增加过拟合训练数据的风险。这项工作提出了一种名为集成对抗性共训练神经网络(Ensemble adversarial co-training neural network, EACNet)的新方法,用于从具有多种缺失模态的多模态磁共振成像(Magnetic resonance imaging, MRI)扫描中进行准确的脑肿瘤分割。所提出的方法由三个关键模块组成:“预训练模型集成”,该模块通过使用预训练模型的集成来捕获来自MRI数据的多样化特征表示;对抗学习,该模块利用涉及两个模型的竞争性训练方法;生成器模型,其创建逼真的缺失数据,而作为判别器的子网络则学习区分真实数据和生成的“假”数据。共训练框架利用多模态路径(在完整扫描上训练)提取的信息来指导处理缺失模态的路径中的学习过程。通过共训练交互,模型可以利用可用模态和肿瘤分割任务之间的关系来补偿缺失信息。EACNet在BraTS2018和BraTS2020挑战数据集上进行了评估,分别达到了较先进和竞争性的性能。值得注意的是,全肿瘤(Whole tumor, WT)Dice相似系数(Dice similarity coefficient, DSC)的分割结果达到89.27%,超过了现有方法的性能。分析表明,集成方法具有潜在的优势,而对抗性共训练则有助于提高EACNet对具有缺失模态的MRI扫描进行脑肿瘤分割的鲁棒性和准确性。我们的实验结果表明,EACNet在缺失模态MRI扫描脑肿瘤分割任务上取得了令人满意的结果,是现实世界临床应用的较优候选方法。

关键词: 深度学习;磁共振成像;医学图像分析;语义分割;分割精度;图像合成

引用格式: RAMADHAN Amran Juma, CHEN Jing, PENG Junlan. EACNet: Ensemble adversarial co-training neural network for handling missing modalities in MRI images for brain tumor segmentation. Journal of Measurement Science and Instrumentation, 2025, 16(1): 11-25. DOI: 10.62756/jmsi.1674-8042.2025002