2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

Deformation Tumor Synthesis with Modal-Data Adaptive Supervision for 3D Brain Tumor Segmentation

Xiaofeng Peng[1], Feng Yang[1,2](✉)

[1] School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi, 530004, China
[2] Guangxi Key Laboratory of Multimedia Communications Network Technology, Guangxi University, Nanning, Guangxi, 530004, China
`yf@gxu.edu.cn`

Abstract. The accurate segmentation of brain tumor is important not only for treatment planning, but also for follow-up evaluations. However, the inadequacy of annotated medical images poses challenges in training the brain tumor segmentation models. This paper addresses this issue by presenting a new method called Deformation Tumor Synthesis with Model-Data Adaptive Supervision (DSMA). DSMA consists of data synthesis and weight allocation. The Deformation Tumor Synthesis (DTSS) strategy combines the morphological features of real tumors and adopts a unique iteration synthesis and fusion mechanism to generate diverse derived synthetic data customized for each set of real data. The Model-Data Adaptive Supervision (MAS) strategy dynamically filters and allocates the loss weights of synthetic data based on the real-time performance of the segmentation model to ensure the positive effects of adding synthetic data. The experimental results on the publicly available MRI brain imaging datasets BraTS2019 and BraTS2020 indicate that the proposed method achieves high-quality data synthesis and effectively improves the performance of the segmentation model.

Keywords: brain tumor segmentation, Deformation Tumor Synthesis, iteration synthesis, Model-Data Adaptive Supervision.

1 Introduction

Using artificial intelligence methods for high-quality segmentation of medical images to delineate organs and lesion areas can greatly improve doctors' work efficiency and diagnostic reliability [1-4].

2D medical image segmentation is efficient and easy to implement, but it has limitations in accurately capturing spatial information and handling complex anatomical structures. The significant advantage of 3D medical image segmentation lies in its ability to capture spatial information, making it possible to extract spatial relationships and contextual information [5-6]. However, to train the AI segmentation model to achieve the desired effect, a large amount of labeled data is needed. Compared to the field of natural images, labeling medical images not only relies on extensive medical expertise but also involves privacy protection and high data costs [7-8]. At this point, data synthesis can alleviate the problem of scarce labeled data.

Methods based on GAN have also been proposed for synthesizing medical images [9-13], capturing features that represent high-level semantic information of images through generator networks and discriminator networks. Diffusion models [14] have recently performed well in generating high-quality images, but their performance is highly dependent on sampling frequency, and the time cost of generating images is significantly higher than that of other generative models. Synthetic images, as new samples, broaden the scale, shape, and positional information, enhancing the feature distribution covered by the original dataset and providing more training material for segmentation networks [15].

Due to the diversity of tumor types, conventional methods struggle to limit the information of the affected areas, making it difficult to obtain synthetic images that meet the requirements for model training. Generating synthetic tumors [16-21] to expand data samples represents a synthetic strategy. The main challenge of synthetic tumors lies in integrating comprehensive information to narrow the gap with real data, which limits their application in model training. To obtain high-quality synthetic data for segmentation model training, this paper proposes a Deformation Tumor Synthesis with Modal-Data Adaptive Supervision (DSMA) method. Inspired by the tumor generation strategy [16], we randomly micro-modify the tumors in real data through a unique iterative synthesis and fusion mechanism, obtaining rich and realistic derived synthetic data based on real data to provide abundant training samples for the segmentation model. Then, by designing a mechanism for dynamically allocating loss weights, the model can assess the quality of synthetic data based on real-time performance, allowing it to better filter and utilize synthetic data, ensuring that the inclusion of synthetic data positively impacts model performance. The contributions of this paper are as follows:

-We designed a Deformation Tumor Synthesis (DTSS) strategy. First, we constructed a deformation tumor entity by referencing the morphological characteristics of real tumors, and then obtained the deformation tumor through an iterative synthesis mechanism, which was then fused with real data to obtain diversified derived synthetic data. This synthesis strategy does not require annotation costs, and the randomness of iterative synthesis expands the characteristics of the derived synthetic data tumors. At the same time, since the synthetic data is based on real data, the generated data will be more realistic.

-We designed a Model-Data Adaptive Supervision (MAS) strategy for dynamically screening and allocating loss weights for synthetic data. The quality of synthetic data is judged adaptively by the model, and the advantage of this strategy is that it can prevent the model from learning towards overfitting, and it can selectively extract features from low-quality samples. The model can extract more features from derived synthetic data, thus achieving more effective model training.

-The experimental results show that the method in this paper achieves high-quality data synthesis, and combined with the model data adaptive supervision strategy, effectively improves the performance of the segmentation model.

2 Related work

2.1 Medical image segmentation

Çiçek et al. [22] expanded the U-Net [23] architecture by Ronneberger et al., replacing all 2D operations with 3D operations while implementing dynamic elastic deformation to effectively enhance data during the training process. Xing et al. proposed a novel nested modality-aware transformer called NestedFormer [24] to explicitly explore the intra-model and inter-modal relationships in brain tumor segmentation from multi-modal MRI. Lin et al. introduced a new FCN architecture called RefineU-Net [25] to improve the performance of U-Net in medical image segmentation. Mecheter et al. proposed a 3D CNN-based network [26] that includes channel squeeze and excitation (cSE) and spatial squeeze and excitation (sSE) modules, which are used to dynamically adjust and emphasize important features during the feature extraction process. Bruzadin et al. introduced a new deep learning method called "Learning Label Diffusion" (LLD) [27] for semi-automatic segmentation of lungs in CT images, particularly targeting lung infections from COVID-19. Medical image data synthesis.

2.2 Medical image data synthesis

Hong et al. expanded on StyleGAN2, called 3D-StyleGAN [28], which inherits the controllability and interpretability of the style vector research from the original StyleGAN2, including projection and reconstruction of unseen real images in the latent space and style mixing. Sun et al. proposed a hierarchical GAN model [29] that generates low-resolution versions of images and randomly selected high-resolution image sub-volumes while employing an encoder with a similar hierarchical structure to extract features from the images. Zia et al. introduced VANT-GAN [12], which combines the concept of residual GAN and utilizes the capabilities of cycle-consistent GAN to generate normal images related to anomalous input images. Bargshady et al. applied CycleGAN technology to achieve the mutual conversion generation of COVID-19 X-ray images and normal images [13]. Shin et al. proposed a framework utilizing Conditional Generative Adversarial Networks (Conditional GAN) and additional conditions, introducing a network for training composite input condition images based on edge filtering to synthesize polyp images [30]. Bai et al. proposed a new method called BSGAN-ADD [31], which enhances the extraction of more representative high-dimensional brain features by generating augmented two-dimensional brain slice images, thereby improving diagnostic performance for early detection of Alzheimer's disease (AD). Lyu et al. used pseudo-labels to guide the synthesis of COVID-19 images, integrating synthetic images into model training by combining single-stage synthetic auxiliary cross pseudo-supervision (SACPS) and multi-stage synthetic auxiliary self-training (SAST) methods [32]. Wang et al. introduced a novel Poisson mixture data augmentation (PBDA) algorithm [33] for generating synthetic diabetic retinopathy images to assist in segmentation tasks. Horvath et al. designed a dual-channel generator trained through cycle consistency to generate tumor image labels using real anatomical data [34].

3 Method

DSMA uses deformation tumor fusion to expand data, combining Model-Data Adaptive Supervision to filter data, and applies it to 3D brain tumor segmentation. In terms of data expansion, deformation tumor monomers are created based on the basic morphology of the tumor, and multiple iterations of synthesis are performed by combining the lesion locations and texture information from real data to obtain deformation tumors. Finally, these are fused with real data to make minor modifications to the tumors, resulting in derived synthetic data. In terms of data filtering, the model adaptively supervises the participation level of synthetic data in real-time based on its performance at different stages on real data during the training process, allowing the model to autonomously and dynamically utilize the information from synthetic data and ensuring that the inclusion of synthetic data has a positive impact on the model's training.

DSMA includes two parts: Deformation Tumor Synthesis (DTSS) and Model-Data Adaptive Supervision (MAS). The model training is divided into two stages. In the first stage, only real data is used to train the segmentation model to achieve a certain level of performance. In the second stage, the DTSS strategy is introduced, generating corresponding derived synthetic data while training with real data, and both are included in the training. At the same time, the MAS strategy is introduced, where the model dynamically selects and allocates the loss weights of synthetic data based on real-time segmentation performance. The training framework of the experiment is shown in Fig. 1.

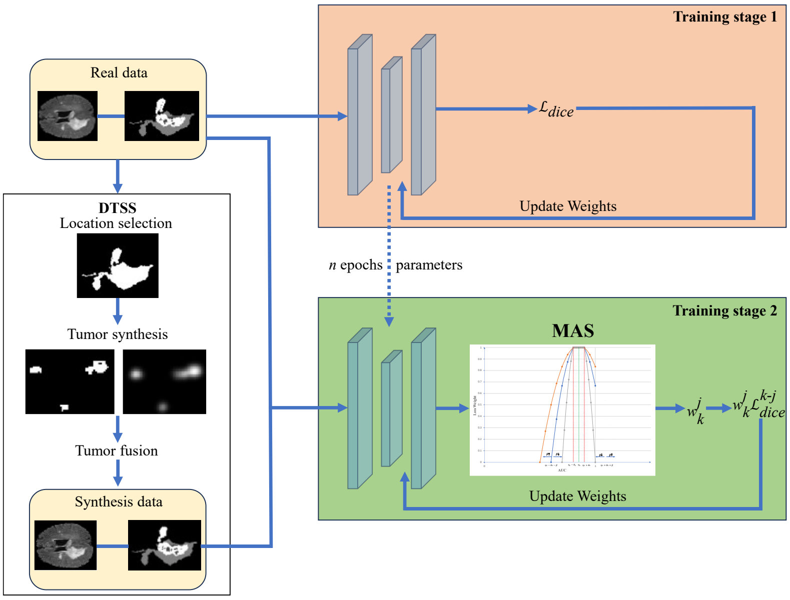
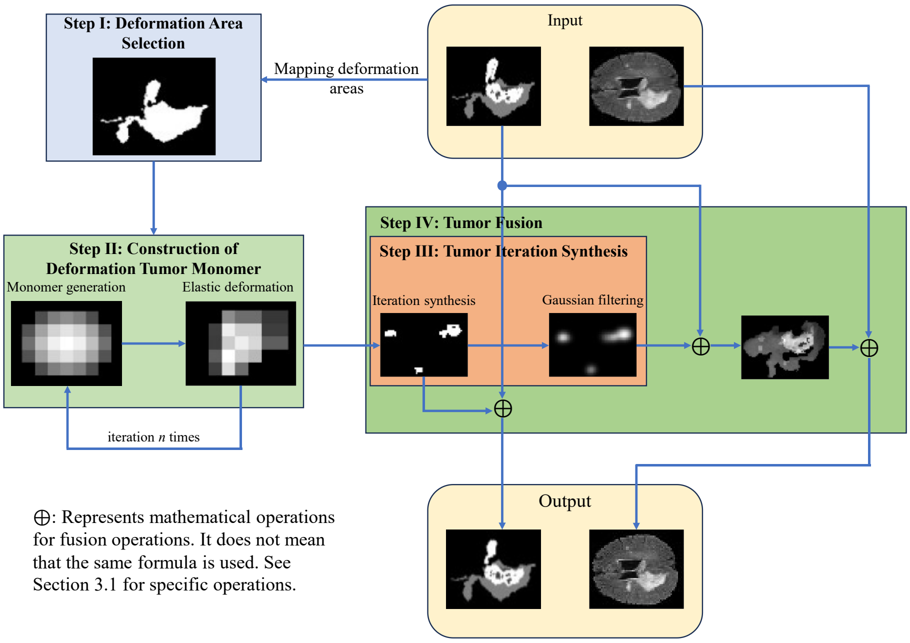


Fig. 1. DSMA training framework. The entire training process is divided into two stages.

3.1 Deformation Tumor Synthesis

The Deformation Tumor Synthesis mainly includes four steps: Deformation Area Selection, Construction of Deformation Tumor Monomer, Tumor Iteration Synthesis, and Tumor Fusion. The DTSS process is shown in Fig. 2.



⊕: Represents mathematical operations for fusion operations. It does not mean that the same formula is used. See Section 3.1 for specific operations.

Fig. 2. The process of Deformation Tumor Synthesis.

Deformation Area Selection. The deformation area is a blank space used for the iterative generation and collision of tumor monomers, mapping to the lesion location in the original data. We extract the lesion location information from the real data to construct the deformation area, which is determined by the following formula:

$$d(x, y, z) = l(x, y, z) \times m(x, y, z), \tag{1}$$

where $l(x, y, z)$ represents the annotated data and $m(x, y, z)$ is a threshold mask whose threshold is determined by the lesion area of the annotated data.

Construction of Deformation Tumor Monomer. Tumor cells are mostly close to a spherical shape, and then they continuously grow, gather, and spread from the center. Based on this principle, we choose an ellipsoid as the basic shape for constructing the deformation tumor monomer. The axis lengths of the ellipsoid in the x, y, and z directions are randomly selected from the given range (r_{min}, r_{max}) of the size r of the deformation tumor monomer. Subsequently, a certain degree of elastic deformation is applied to each generated ellipsoid to enhance the diversity of tumor textures, making them closer to the characteristics of real tumors [23,35], ultimately resulting in the deformation tumor monomer.

Tumor Iterative Synthesis. We extract texture information from the lesion locations of real data and fill the tumor monomers at the pixel level. Then, we iteratively generate tumor monomers in the deformation area. When the real tumor collides with the deformation tumor monomer in position, the texture of the conflict area is randomly retained, overlapped, or reduced, thereby affecting the morphology and texture of the tumor. Retention means no changes are made. Overlapping means adding at the pixel level. Reduction means subtracting or deleting at the pixel level. Finally, an iterative tumor prototype is generated, denoted as $t(x, y, z)$. Subsequently, a Gaussian filter with a standard deviation of $\sigma(g(x, y, z; \sigma))$ is applied to blur the iterative tumor prototype to generate the deformation tumor $T(x, y, z)$, where σ represents the standard deviation. The deformation tumor is given by the following formula:

$$T(x, y, z) = t(x, y, z) \otimes g(x, y, z; \sigma). \tag{2}$$

Tumor Fusion. In order to obtain derived synthetic data, it is necessary to fuse the deformation tumor into the real data to modify the tumors. For the labels, the deformation tumor are directly superimposed onto the labels to obtain a new label G_L, which is fused by the following formula:

$$G_L = L + G_t, \tag{3}$$

where L represents the label data, G_t represents the deformation tumor, and G_L represents the synthesis label. For images, we first extract the texture i of the real image and the lesion area l in the label, then add Gaussian noise to G_t to obtain $G_t^{'}$, and obtain the fuzzy texture $l^{'}$ of the lesion area through the following formula:

$$l^{'} = (i - G_t^{'}) \odot l, \tag{4}$$

where $\odot$ represents point-wise multiplication. Then $l^{'}$ is fused to the lesion area of the real image through the following formula:

$$i^{'} = i \odot (1 - l) + l^{'}. \tag{5}$$

Finally, replace the lesion part in the real data with the fused lesion data $i^{'}$, thereby obtaining the derived synthetic image data. The actual synthesis effect comparison example is shown in Fig. 3.

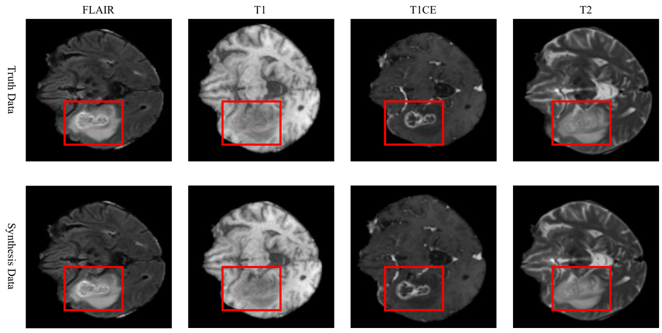


Fig. 3. Example of comparing synthesized data with real data.

3.2 Model-Data Adaptive Supervision

We enrich tumor features by constructing derived synthetic data through random iterations of tumor synthesis, which gives the model a chance to learn more features with limited data, thus improving segmentation performance. However, the differences between the derived synthetic data and the real data are not explored from a specialized perspective, so it is difficult to ensure that the quality of the synthetic data is all optimal. For example, the real data may be overly modified so that it seriously deviates from the actual situation of the brain, or the real data may be too slightly modified so that it differs too little from the original data. Therefore, in order to maximize the positive impact of derived synthetic data on model training, we propose a Model-Data Adaptive Supervision (MAS) strategy that allows the model to autonomously choose the data to learn based on its own performance. The method dynamically adjusts the participation level of each synthetic data in different training stages, which is reflected in the adjustment of weights in the loss function. The weight w_k^j assigned to the j-th synthetic data in the k-th epoch of training is determined by the following formula:

$$w_k^j = \begin{cases} 1, \text{ where } L_{\mathrm{k}}^j \in \left[\eta_k - \theta_k, \eta_k + \theta_k\right], \\ \left(\eta_k - \theta_k\right) > 0, \left(\eta_k + \theta_k\right) < 1. \\ -\frac{1}{\beta(2\theta_k + \beta)}\left[L_k^j - (\eta_k - \theta_k - \beta)\right]\left[L_k^j - (\eta_k + \theta_k + \beta)\right], \\ \text{where } L_k^j \in \left(max\left(0, \eta_k - \theta_k - \beta\right), \eta_k - \theta_k\right), \\ \text{or } L_k^j \in \left(\eta_k + \theta_k, \ min\left(\eta_k + \theta_k + \beta, 1\right)\right). \\ 0, \text{ where } L_k^j \le (\eta_k - \theta_k - \beta) \\ \text{or } L_k^j \ge (\eta_k + \theta_k + \beta). \end{cases}, \tag{6}$$

where L_k^j represents the segmentation accuracy of the model on the j-th synthetic data after k epochs of training. η_k represents the average segmentation accuracy of the model on real data during the k-th epoch of training, which is determined by the following formula:

$$\eta_k = \frac{\sum_{j=1}^{n} L_j^{'}}{n}, \tag{7}$$

where n represents the scale of the real data involved in training, and $L_j^{'}$ represents the segmentation accuracy of the model on the j-th real data after the k-th epoch of training.

θ_k represents the expansion width in the k-th epoch of training, used to broaden the model's acceptance range for data quality. It is determined by the following formula:

$$\theta_k = \frac{L_{kmax} - L_{kmin}}{2} + c, \tag{8}$$

where L_{kmax} and L_{kmin} represent the maximum and minimum segmentation accuracy of the model on the real data participating in the k-th epoch of training, respectively. The parameter c is an adjustable hyperparameter that can be further adjusted based on actual conditions to refine the expansion width.

The weight function we designed is a piecewise function that is centrally symmetric, composed of constant functions and quadratic functions. The constant function is used to determine high-quality data that is suitable for the current model training iteration. η_k determines the midpoint of the constant function, and θ_k expands the model's adaptability range. Synthetic data corresponding to the segmentation accuracy interval within $[\eta_k - \theta_k, \eta_k + \theta_k]$ will fully participate in training, with a loss weight allocation of 1.

Considering that the learning value of low-quality synthetic data is low, adding it to the training may reduce the model's performance, and high-quality data should not tend to be memorized, which may lead the model towards overfitting. Therefore, the model's segmentation accuracy for synthetic data in the k-th training epoch is adaptively weighted using the properties of a concave function when it is in the ranges $(max(0, \eta_k - \theta_k - \beta), \eta_k - \theta_k)$ or $(\eta_k + \theta_k, min(\eta_k + \theta_k + \beta, 1))$. In this weight function, a downward-opening quadratic function is used for adaptation, with its shape determined by η_k, θ_k, and β. The coefficients of the quadratic function are mainly determined by three data points: the adaptation midpoint m, the adaptation boundary point u, and the discard point q. There is only one adaptation midpoint m, and the weight function does not pass through this point, but the corresponding quadratic function does, with its value being η_k, representing the model's average segmentation accuracy for real data in the k-th training round, determined by the following formula:

$$\eta_k = -\frac{b}{2a}, \tag{9}$$

where a and b represent the coefficients of the quadratic term and the linear term of the quadratic function, respectively. The boundary point u corresponds to the point where the loss weight allocation is 1, and there are two such points, representing the boundary points of the high-quality interval identified by the current stage model, reflected in the function as $u_1((\eta_k - \theta_k), 1)$ and $u_2((\eta_k + \theta_k), 1)$. The discarded point q corresponds to the point where the loss weight allocation is 0, and there are also two such points, indicating that the quality of the samples is low and should not be included in training, serving as the boundary for evaluating whether data should be included in training, reflected in the function as $q_1((\eta_k - \theta_k - \beta), 0)$ and $q_2((\eta_k + \theta_k + \beta), 0)$. Whether these two points can be reached on the weight function depends on whether $(\eta_k - \theta_k - \beta)$ is greater than 0 and whether $(\eta_k + \theta_k + \beta)$ is less than 1. β is an adjustable parameter that can adjust the sampling range for including synthetic data in training and the rate at which weights decrease as data quality declines. Through the above three sets of data, the shape and parameters of the quadratic function can be determined. At this point, let this quadratic function be:

$$y = a(x - (\eta_k - \theta_k - \beta))(x - (\eta_k + \theta_k + \beta)). \tag{10}$$

By substituting any adaptive boundary point u into the following equation, the undetermined coefficients of the quadratic function can be determined:

$$a((\eta_k - \theta_k) - (\eta_k - \theta_k - \beta))((\eta_k - \theta_k) - (\eta_k + \theta_k + \beta)) = 1. \tag{11}$$

This yields a:

$$a = -\frac{1}{\beta(2\theta_k + \beta)}. \tag{12}$$

Finally, we can obtain the weight function of the quadratic function model in the segmentation accuracy interval $(max(0,\ \eta_k - \theta_k - \beta),\ \eta_k - \theta_k)$ and $(\eta_k + \theta_k,\ min(\eta_k + \theta_k + \beta,\ 1)$:

$$w_k^j = -\frac{1}{\beta(2\theta_k + \beta)}\left[L_k^j - (\eta_k - \theta_k - \beta)\right]\left[L_k^j - (\eta_k + \theta_k + \beta)\right]. \tag{13}$$

In the low accuracy range, as the data quality decreases, the rate of decline in the assigned weight values also becomes faster, until the discard point q is reached and data is discarded. The advantage of this approach is that it can fully utilize data of varying quality. A similar effect is observed in the high accuracy range, which can to some extent prevent the model from excessively memorizing data and thus shifting towards overfitting.

When the segmentation accuracy of the model in the k-th epoch of training is lower than $(\eta_k - \theta_k - \beta)$ or higher than $min((\eta_k + \theta_k + \beta),\ 1)$, the synthetic data does not participate in training, and the loss weight allocation is 0. The weight function model is shown in Fig. 4.

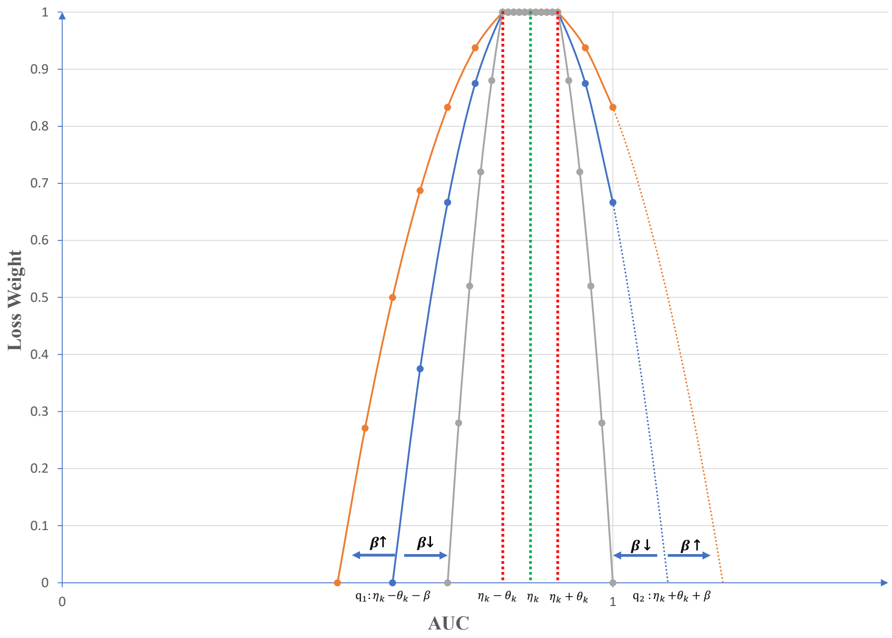


Fig. 4. Weighting function model diagram

3.3 Loss function

We use the Dice loss as the loss function. Let the model's prediction be represented by P and the true label by G. The Dice loss is defined as:

$$\mathcal{L}_{dice} = 1 - \frac{2\sum_v P_v G_v}{\sum_v (P_v + G_v)}, \tag{14}$$

where P_v and G_v represent the predicted result and the true label value of the v-th pixel point, respectively. According to the method proposed in Section 3.2, dynamic loss weights are assigned to each synthetic data for each training epoch. The loss function $\mathcal{L}_{dice}^{k\text{-}j}$ of the j-th synthetic data in the k-th training epoch is defined as:

$$\mathcal{L}_{dice}^{k\text{-}i} = w_k^j \mathcal{L}_{dice}. \tag{15}$$

4 Experiments

4.1 Datasets and evaluation metrics

To verify the effectiveness of this method, this study evaluated the performance of DSMA on two publicly available 3D multimodal datasets, BraTS2019 and BraTS2020 [36,37]. We randomly divided each dataset into two parts: 80% for training and 20% for testing. DSMA was applied to the training part of the dataset, and the comparative methods were also applied to the training part of the dataset. To obtain labels for the comparative methods, an additional teacher model identical to the training model was constructed during actual training. The teacher model updates the network parameters through exponential moving average (EMA), and synthetic data is processed through the teacher model to obtain pseudo-labels. Each synthetic data generated by the comparative methods was subjected to different random noise to obtain four noisy datasets, simulating the quantity of multimodal data. The experiments used the Dice similarity coefficient (DSC) as the evaluation metric.

BraTS2019. This dataset contains MRI data from 259 cases of high-grade glioma (HGG) and 76 cases of low-grade glioma (LGG), with four modalities: FLAIR, T1, T1CE, and T2. The task involves segmenting three regions: whole tumor (WT), enhancing tumor (ET), and tumor core (TC). In this study, only HGG data was used for the experiments.

BraTS2020. This dataset contains MRI data of 369 cases of brain tumors, with 4 modalities: FLAIR, T1, T1CE, and T2. Similarly, the task is to segment the WT, ET, and TC regions.

4.2 Implementation details

The experiment was conducted on PyTorch 2.0, CUDA 12.2, the Monai framework, and the NVIDIA GTX 3090 GPU. To ensure a fair comparison, the experiments were executed on the same codebase, using consistent evaluation methods and training procedures. The initial parameters of the segmentation model remained unchanged throughout the experiment. Training phase 1 used only real data, and after 800 epochs, it entered training phase 2, introducing the DSMA method and evaluating it.

The key parameter settings include using the AdamW [38] optimizer with a weight decay set to 10^{-5}. The learning rate for training phase 1 is set to 10^{-4}, and it is reduced to 10^{-5} in training phase 2. The range for deformation tumor monomer size r is set to [3,5]. Each time a tumor monomer is generated, its size in the x, y, and z directions is randomly selected from r, with the number of iterations set to [3,10], and the number of iterations required to generate deformation tumors is randomly selected within this range. The c in MAS is set to 0.00002, which is used to adjust the expansion width of the weight function. The settings of the key parameters are determined based on experimental testing and experience.

4.3 Result

The experiment validates the DSMA method of this paper on the 3D U-Net [22] and NestedFormer [24] segmentation models, and compares it with GAN synthesis methods: 3D-StyleGAN [28] and HA-GAN [29], as well as diffusion model synthesis methods: Wu et al. [39] and Yu et al. [40]. The experimental results are shown in Table 1 and Table 2. The segmentation results are shown in Fig. 5.

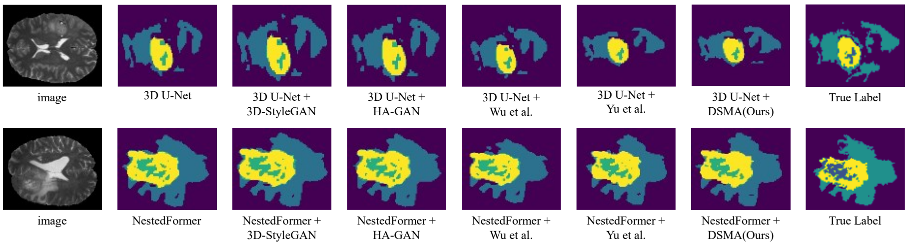


Fig. 5. Visualization of the segmentation results of each method

The results in Table 1 indicate that when using 3D U-Net [22] as the segmentation network, the proposed method DSMA effectively improves the model's segmentation performance on two datasets. In the BraTS2019 and BraTS2020 datasets, the average Dice Similarity Coefficient (DSC) of the proposed method DSMA reached 88.49% and 83.59%, respectively, which is an average increase of 0.59% and 0.96% compared to training with only real data. For the methods that introduced 3D-StyleGAN [28] and HA-GAN [29], the average DSC results did not show significant improvement, mainly due to the limited accuracy of the generative networks and the constraints of the training data, which resulted in insufficient recovery of various details in the anatomical structures of the generated images, affecting the training effectiveness of the model. For the methods introduced by Wu et al. [39] and Yu et al. [40], there was a slight increase in the average DSC results, mainly because the diffusion models have better sampling performance and can generate higher quality synthetic data. However, they are still relatively dependent on the training data, and the pathological features of the synthetic data are still difficult to expand, leading to limited improvement in model performance. Table 2 uses the same dataset as Table 1, but Table 2 employs NestedFormer [24] as the segmentation network, achieving a higher DSC in various segmentation tasks compared to Table 1, mainly due to the more advanced segmentation network architecture having better feature extraction capabilities. In the BraTS2019 and BraTS2020 datasets, the average DSC of the proposed method DSMA reached 91.45% and 87.81%, respectively, which is an improvement of 1.37% and 1.84% compared to training with only real data. In contrast, the proposed method DSMA is more effective in enhancing the performance of the segmentation model.

In summary, the method DSMA in this paper can further enhance the model when its performance reaches or approaches a bottleneck. At the same time, the derived synthetic data is generated one-to-one in combination with real data, allowing it to exhibit textures and features similar to the dataset to a certain extent. By randomly micro-modifying tumors, the derived synthetic data can construct rich tumor features that real data does not possess while maintaining a certain level of data quality. Furthermore, the MAS strategy of this method enables the model to dynamically filter and learn from the derived synthetic data at different stages, fully utilizing its data characteristics and ensuring that the inclusion of derived synthetic data can effectively improve the model's performance.

Table 1. Comparative experimental results between BraTS2019 and BraTS2020 on 3D U-Net

Dataset	BraTS2019				BraTS2020			
Methods	DSC(%)							
	WT	TC	ET	AVE	WT	TC	ET	AVE
3D U-Net[22]	88.39	90.15	85.14	87.90	88.94	80.75	78.21	82.63
3D-StyleGAN[28]	88.43	90.01	85.15	87.86	88.68	80.36	77.86	82.30
HA-GAN[29]	88.39	90.07	85.16	87.88	88.82	80.43	77.81	82.35
Wu et al.[39]	88.69	90.34	85.34	88.12	89.15	80.94	78.09	82.72
Yu et al.[40]	88.84	90.45	85.42	88.24	89.25	81.12	77.98	82.79
DSMA(Ours)	**89.17**	**90.75**	**85.54**	**88.49**	**89.43**	**82.62**	**78.73**	**83.59**

Table 2. Comparative experimental results between BraTS2019 and BraTS2020 on Nested-Former

Dataset	BraTS2019				BraTS2020			
Methods	DSC(%)							
	WT	TC	ET	AVE	WT	TC	ET	AVE
NestedFormer[24]	91.29	92.62	87.09	90.33	90.90	86.54	80.47	85.97
3D-StyleGAN[28]	91.24	92.53	87.18	90.32	90.72	86.04	81.18	85.98
HA-GAN[29]	91.27	92.48	87.22	90.32	91.26	86.26	81.23	86.25
Wu et al.[39]	91.47	93.07	87.80	90.78	91.79	87.31	82.21	87.10
Yu et al.[40]	91.85	93.33	87.72	90.93	91.90	87.63	82.49	87.34
DSMA(Ours)	**92.21**	**93.68**	**88.46**	**91.45**	**92.06**	**88.34**	**83.03**	**87.81**

4.4 Ablation studies

The method DSMA mainly consists of two parts: Deformation Tumor Synthesis (DTSS) and Model-Data Adaptive Supervision (MAS), which are used for deriving data synthesis to expand data features and for data selection and weight allocation, respectively. DTSS generates derived synthetic data by merging the designed deformation tumor with real data. MAS dynamically selects and allocates the loss weights of the derived synthetic data through a designed weight function. For the evaluation of the DTSS method, no methods were applied, and only simple data augmentation techniques such as translation, random rotation, and center cropping were applied to the real data before adding it to the training, and MAS was removed separately. The results of the ablation experiments obtained are shown in Table 3 and Table 4.

Table 3 and Table 4's ablation experiments indicate that compared to simple data augmentation, incorporating derivative synthetic data generated by DTSS into training can further improve the model's segmentation performance. When combined with the MAS strategy, the model's segmentation performance is further enhanced, which suggests that due to the lack of design and selection of professional metrics, it is difficult to avoid the inclusion of low-quality derivative synthetic data generated by DTSS in training, thereby affecting the model's performance. Through the supervision strategy provided in this paper, the synthetic data is reasonably selected and assigned loss weights, which can further ensure that the inclusion of derivative synthetic data positively contributes to the improvement of model performance.

Table 3. Ablation study of DSMA in the 3D U-Net model on BraTS2020

Model	DTSS	MAS	DSC(%)			
			WT	TC	ET	AVE
3D U-Net[22]			88.94	80.75	78.21	82.63
	√		89.08	81.51	78.43	83.01
	√	√	**89.43**	**82.62**	**78.73**	**83.59**

Table 4. Ablation study of DSMA in the NestedFormer model on BraTS2020

Model	DTSS	MAS	DSC(%)			
			WT	TC	ET	AVE
NestedFormer[24]			90.90	86.54	80.47	85.97
	√		91.76	87.68	82.09	87.18
	√	√	**92.06**	**88.34**	**83.03**	**87.81**

5 Conclusion

This article proposes a method for generating synthetic data of deformed tumors and combines model data adaptive supervision to expand the data and enhance the performance of segmentation models, addressing the issue of scarce medical images. The characteristics of the research method in this article are: constructing deformed tumors through iterative synthesis, obtaining derived synthetic data by merging with real data, and filtering and selectively learning from the derived synthetic data through dynamic allocation of loss weights to ensure that the introduction of the method has a positive impact on the model's segmentation performance. Experimental results show that the method in this article can generate derived synthetic data, providing the possibility for the model to learn potential features that are not present in the original data, thereby effectively improving model performance; adjusting the participation level of derived synthetic data of different qualities in training through the filtering strategy of model data adaptive supervision further enhances the performance of the segmentation model. In future research, we will further explore the application of image synthesis strategies on more organs to generate richer and higher-quality synthetic data to comprehensively assist segmentation tasks.

Acknowledgments. This study was funded by the National Natural Science Foundation of China (NSFC) through Grants 61861004 and 61962004, as well as by support from the Finnish Cultural

Foundation for North Ostrobothnia Regional Fund (Grant 60231712) and the Instrumentarium Foundation (Grant 240016).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Hatamizadeh A, Tang Y, Nath V, et al. Unetr: Transformers for 3d medical image segmentation[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2022: 574-584.1
2. Isensee F, Jaeger P F, Kohl S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature methods, 2021, 18(2): 203-211.1,3,8
3. Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE transactions on medical imaging, 2019, 39(6): 1856-1867
4. Wang R, Lei T, Cui R, et al. Medical image segmentation using deep learning: A survey[J]. IET Image Processing, 2022, 16(5): 1243-1267
5. Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis[J]. Medical image analysis, 2017, 42: 60-88
6. Taha A A, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool[J]. BMC medical imaging, 2015, 15: 1-28
7. Gkoulalas-Divanis A, Loukides G, Sun J. Publishing data from electronic health records while preserving privacy: A survey of algorithms[J]. Journal of biomedical informatics, 2014, 50: 4-19
8. Dong T, Zhao B, Lyu L. Privacy for free: How does dataset condensation help privacy?[C]//International Conference on Machine Learning. PMLR, 2022: 5378-5396
9. Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1125-1134
10. Jin D, Xu Z, Tang Y, et al. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11. Springer International Publishing, 2018: 732-740
11. Wu E, Wu K, Cox D, et al. Conditional infilling GANs for data augmentation in mammogram classification[C]//Image Analysis for Moving Organ, Breast, and Thoracic Images: Third International Workshop, RAMBO 2018, Fourth International Workshop, BIA 2018, and First International Workshop, TIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 3. Springer International Publishing, 2018: 98-106
12. Zia T, Murtaza S, Bashir N, et al. VANT-GAN: adversarial learning for discrepancy-based visual attribution in medical imaging[J]. Pattern Recognition Letters, 2022, 156: 112-118
13. Bargshady G, Zhou X, Barua P D, et al. Application of CycleGAN and transfer learning techniques for automated detection of COVID-19 using X-ray images[J]. Pattern Recognition Letters, 2022, 153: 67-74
14. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851

15. Han C, Rundo L, Araki R, et al. Combining noise-to-image and image-to-image GANs: Brain MR image augmentation for tumor detection[J]. Ieee Access, 2019, 7: 156966-156977
16. Hu Q, Chen Y, Xiao J, et al. Label-free liver tumor segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7422-7432
17. Wyatt J, Leach A, Schmon S M, et al. Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 650-656
18. Jin Q, Cui H, Sun C, et al. Free-form tumor synthesis in computed tomography images via richer generative adversarial network[J]. Knowledge-Based Systems, 2021, 218: 106753
19. Han C, Kitamura Y, Kudo A, et al. Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection[C]//2019 International Conference on 3D Vision (3DV). IEEE, 2019: 729-737
20. Wang Q, Wang Z, Li C. Liver Tumor Synthesis and Segmentation Based on Generative Adversarial Networks[C]//2023 3rd International Conference on Electronic Information Engineering and Computer (EIECT). IEEE, 2023: 404-408
21. Chen Q, Chen X, Song H, et al. Towards generalizable tumor synthesis[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 11147-11158
22. Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. Springer International Publishing, 2016: 424-432
23. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241
24. Xing Z, Yu L, Wan L, et al. NestedFormer: Nested modality-aware transformer for brain tumor segmentation[C]//International Conference on Medical Image Computing and computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2022: 140-150
25. Lin D, Li Y, Nwe T L, et al. RefineU-Net: Improved U-Net with progressive global feedbacks and residual attention guided local refinement for medical image segmentation[J]. Pattern Recognition Letters, 2020, 138: 267-275
26. Mecheter I, Abbod M, Zaidi H, et al. Brain MR images segmentation using 3D CNN with features recalibration mechanism for segmented CT generation[J]. Neurocomputing, 2022, 491: 232-243
27. Bruzadin A, Boaventura M, Colnago M, et al. Learning label diffusion maps for semi-automatic segmentation of lung CT images with COVID-19[J]. Neurocomputing, 2023, 522: 24-38
28. Hong S, Marinescu R, Dalca A V, et al. 3D-StyleGAN: A style-based generative adversarial network for generative modeling of three-dimensional medical images[C]//Deep Generative Models, and Data Augmentation, Labelling, and Imperfections: First Workshop, DGM4MICCAI 2021, and First Workshop, DALI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 1. Springer International Publishing, 2021: 24-34
29. Sun L, Chen J, Xu Y, et al. Hierarchical amortized GAN for 3D high resolution medical image synthesis[J]. IEEE journal of biomedical and health informatics, 2022, 26(8): 3966-3975

30. Shin Y, Qadir H A, Balasingham I. Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance[J]. IEEE Access, 2018, 6: 56007-56017
31. Bai T, Du M, Zhang L, et al. A novel Alzheimer's disease detection approach using GAN-based brain slice image enhancement[J]. Neurocomputing, 2022, 492: 353-369
32. Lyu F, Ye M, Carlsen J F, et al. Pseudo-label guided image synthesis for semi-supervised covid-19 pneumonia infection segmentation[J]. IEEE Transactions on Medical Imaging, 2022, 42(3): 797-809
33. Wang H, Zhou Y, Zhang J, et al. Anomaly segmentation in retinal images with poisson-blending data augmentation[J]. Medical Image Analysis, 2022, 81: 102534
34. Horvath I, Paetzold J, Schoppe O, et al. Metgan: generative tumour inpainting and modality synthesis in light sheet microscopy[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2022: 227-237
35. Ogden R W. Non-linear elastic deformations[M]. Courier Corporation, 1997.4
36. Bakas S, Akbari H, Sotiras A, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features[J]. Scientific data, 2017, 4(1): 1-13
37. Menze B H, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS)[J]. IEEE transactions on medical imaging, 2014, 34(10): 1993-2024
38. Loshchilov I, Hutter F. Decoupled weight decay regularization[J]. arXiv preprint arXiv:1711.05101, 2017
39. Wu Z, Chen X, Xie S, et al. Super-resolution of brain MRI images based on denoising diffusion probabilistic model[J]. Biomedical Signal Processing and Control, 2023, 85: 104901
40. Yu J, Oh H, Lee Y, et al. Denoising diffusion model with adversarial learning for unsupervised anomaly detection on brain MRI images[J]. Pattern Recognition Letters, 2024, 186: 229-235