



## OPEN ACCESS

## EDITED BY

Nebojsa Bacanin,  
Singidunum University, Serbia

## REVIEWED BY

Zhiwei Ji,  
Nanjing Agricultural University, China  
Miodrag Zivkovic,  
Singidunum University, Serbia

## \*CORRESPONDENCE

Bin Li  
✉ libin781029@sina.com  
Qiushi Sun  
✉ shao\_minghao@126.com  
Xianjin Fang  
✉ xjfang@aust.edu.cn

†These authors share first authorship

RECEIVED 24 January 2024

ACCEPTED 18 September 2024

PUBLISHED 02 October 2024

## CITATION

Li B, Sun Q, Fang X, Yang Y and Li X (2024) A novel metastatic tumor segmentation method with a new evaluation metric in clinic study.  
*Front. Med.* 11:1375851.  
doi: 10.3389/fmed.2024.1375851

## COPYRIGHT

© 2024 Li, Sun, Fang, Yang and Li. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# A novel metastatic tumor segmentation method with a new evaluation metric in clinic study

Bin Li<sup>1\*†</sup>, Qiushi Sun<sup>2\*†</sup>, Xianjin Fang<sup>3\*</sup>, Yang Yang<sup>3</sup> and Xiang Li<sup>3,4</sup>

<sup>1</sup>Department of Neurology, The First Hospital of Anhui University of Science and Technology, Huainan, China, <sup>2</sup>Department of Anesthesiology, Fudan University Affiliated Huashan Hospital, Shanghai, China, <sup>3</sup>Department of Anesthesiology, Fudan University Affiliated Huashan Hospital, Huainan, China, <sup>4</sup>School of Safety Science and Engineering, Anhui University of Science and Technology, Huainan, China

**Background:** Brain metastases are the most common brain malignancies. Automatic detection and segmentation of brain metastases provide significant assistance for radiologists in discovering the location of the lesion and making accurate clinical decisions on brain tumor type for precise treatment.

**Objectives:** However, due to the small size of the brain metastases, existing brain metastases segmentation produces unsatisfactory results and has not been evaluated on clinic datasets.

**Methodology:** In this work, we propose a new metastasis segmentation method DRAU-Net, which integrates a new attention mechanism multi-branch weighted attention module and DResConv module, making the extraction of tumor boundaries more complete. To enhance the evaluation of both the segmentation quality and the number of targets, we propose a novel medical image segmentation evaluation metric: multi-objective segmentation integrity metric, which effectively improves the evaluation results on multiple brain metastases with small size.

**Results:** Experimental results evaluated on the BraTS2023 dataset and collected clinical data show that the proposed method has achieved excellent performance with an average dice coefficient of 0.6858 and multi-objective segmentation integrity metric of 0.5582.

**Conclusion:** Compared with other methods, our proposed method achieved the best performance in the task of segmenting metastatic tumors.

## KEYWORDS

brain metastases, precise treatment, deep learning, medical image segmentation, multi-objective segmentation integrity metric

## 1 Introduction

Brain metastases (BM) represent the predominant intracranial malignancies, emanating from primary sources like breast cancer, melanoma, and other cancers (1). As a distinct pathological entity, the therapeutic approach to managing BM encompasses a multitude of options, including whole-brain radiation therapy, stereotactic radiosurgery, surgical resection, targeted therapy, and immunotherapy (2). Precise identification of BM assumes paramount

significance for clinicians, facilitating the initial screening for intracranial lesions, formulating timely and tailored treatment strategies, and prognosticating follow-up responses to avert unfavorable clinical outcomes.

Owing to the diverse and complex nature of metastatic tumors, magnetic resonance imaging (MRI) technology emerges as a pivotal tool for elucidating the comprehensive landscape of these malignancies. Serving as a non-invasive imaging technique, MRI not only furnishes essential intracranial functional information but also enables clinicians and researchers to attain a more holistic understanding of tumor tissue characteristics and lesion nature, leveraging its high spatial resolution and multimodal advantages (3). The segmentation of metastatic tumors yields extensive three-dimensional data, enriching pathological research. Through this segmentation process, insights into the tumor's shape, size, and distribution are garnered, providing crucial information for the formulation of personalized treatment plans. Evaluation of the impact of treatment on the tumor, achieved by comparing segmentation results at different time points, facilitates timely adjustments to treatment plans, enhancing clinical efficacy. Nevertheless, manual delineation of segmentation results by experts proves inefficient, and the inherent variability in outcomes due to differing subjective opinions among doctors necessitates a more standardized approach (4). Exploring automated segmentation methods not only streamlines the workload for radiologists but also mitigates result discrepancies arising from subjective interpretations (5, 6).

Deep learning (DL), leveraging its data-driven and end-to-end capabilities, has found extensive applications in medical image analysis (7–9). Capitalizing on highly adaptive feature learning and multimodal fusion, deep learning-based frameworks exhibit a commendable ability to accurately delineate tumor boundaries (10–12). Numerous models have been previously proposed for quantitative analysis of BM but there still exist several challenges that hamper the clinical applicability of automatic detection (13). The first common challenge is boosting the detection of the small volume of BM and collaboratively decreasing false-positive (FP) rate (14). For experienced radiologists, detecting minuscule lesions presents a significant challenge, and any lesions that go unnoticed can substantially hinder the accuracy of patient diagnoses. The trade-off between the sensitivity and FP rate often puzzles the researchers in the Deep learning model design and selection (15). Models with high sensitivity would be inclined to identify and preselect the subtle lesions, whereas high FP would impede the accuracy of diagnosis. Yoo (16) proposed a DL model with a 2.5D overlapping patch technique to isolate a BM of less than  $0.04\text{cm}^2$  in CE-MRI. Their model could detect relatively small tumors compared to previous studies, but the overall dice accuracy of the model is not satisfactory. Dikici (17) used a dual-stage framework to enhance the precision of isolating small lesions with an average volume of only  $159.6\text{mm}^3$ . The framework, consisting of the candidate-selection stage and detection stage with a custom-built 3D CNN, achieves a high sensitivity on their BM database. However, due to model parameter limitations, it cannot recognize lesions exceeding 15 mm. Furthermore, the accuracy of BM detection and segmentation is limited by the characteristics and the quality of MRI images (18). In addition, different MRI imaging equipment and sequence parameters pose considerable challenges to the generalization ability of segmentation models. Zhou (19) trained a

DL single-shot detector based on T1-weighted gradient-echo MRI, and the sensitivity of the testing group was 81% but only validated in single-center data. Grøvik (20) used four different MRI sequences for segmentation using the DeepLab V3 network, achieving a dice accuracy of 0.79. Although validated in two central data sets, the overall sample size is only 100 cases. In summary, the accuracy of existing segmentation methods for metastatic tumors is low, and their performance in clinical applications is inadequate. Therefore, developing a new segmentation technique for brain tumors that can precisely segment small metastatic tumors and deliver improved results, even with limited resolution, remains a challenging problem (21, 22).

In the context of medical image segmentation, metrics such as dice similarity coefficient metric and intersection over union not only gauge the accuracy of the segmentation model but are also frequently employed as loss functions for the model. However, current segmentation evaluation metrics often assess segmentation at a global level, which presents certain limitations. For instance, the dice similarity coefficient metric is sensitive to larger segmentation regions when assessing the presence of multiple segmentation regions within an image. Consequently, this approach may not provide an objective evaluation of smaller targets.

To enhance segmentation model generalization, we introduce an encoder-decoder framework incorporating deep convolution and attention mechanisms. Using publicly available brain imaging data for model training, and evaluate and compare performance metrics against various existing segmentation models in this study. The results show that our proposed model has excellent segmentation ability.

The main contributions of this article are as follows:

A new BM segmentation method for effectively extracting tumor boundaries and features: We propose DRAU-Net, a novel medical image segmentation method incorporating a multi-branch weighted attention module and multiple dilated residual convolution modules. This method achieved accurate segmentation results, demonstrating robust performance across a range of clinical medical settings, including those involving low-quality images and datasets with limited layers.

It is the first time an indicator that focuses on the global situation is proposed: In order to solve the problem of quantitatively calculate the lots of BMs with small sizes. This article first proposes a new medical image segmentation evaluation metric: multi-objective segmentation integrity metric (MSIM), which evaluates the integrity of multiple segmentation targets, a metric overlooked by most existing indicators.

We are validating the effectiveness of DRAU-Net on multiple datasets: To assess the generalization capability of our proposed segmentation method across diverse datasets, we have obtained favourable results from both publicly available and collected clinical metastasis datasets.

The structure of this article is organized as follows: The introduction, presented in the first section, highlights the clinical significance of metastasis segmentation and delineates the challenges currently faced in this field. The Materials and Methods section of the second section, details the brain tumor segmentation dataset utilized, outlines the data preprocessing procedures, introduces the novel DRAU-Net segmentation approach, and describes the experimental details. In section three, Experiment and Results, we present and analyse the segmentation evaluation indicators, comparative

experiments, and ablation experiments of this article. The fourth section of the discussion, deliberates on the methodologies proposed within this work and offers insights into potential future directions. Finally, we summarized this article in section five.

## 2 Materials and methods

### 2.1 Data introduction

To assess the model generalization ability across diverse datasets, we employed the BraTS2023 Brain Metastases dataset, which encompasses data on brain metastases acquired from various institutions under different standard clinical conditions (23, 24). As the BraTS2023 test dataset details were not disclosed, we introduced randomness by shuffling the remaining data. Subsequently, 210 samples were designated for the training set, while 28 samples were set aside for the testing set. The BraTS2023 comprises multi-parameter MRI scans, including pre-contrast T1-weighted (t1w), post-contrast T1-weighted (t1c), T2-weighted (t2w), and T2-weighted Fluid Attenuated Inversion Recovery (t2f) images. All MRI images underwent standardization, co-registration to the common analytical template (SRI24), and skip striping. For segmentation purposes, the BraTS2023 Brain Metastases dataset utilizes three labels: No-enhancing tumor Core (NETC; Label 1), Surrounding non-enhancing FLAIR hyperintensity (SNFH; Label 2), and Enhancing Tumor (ET; Label 3).

Furthermore, this study conducted experiments using the metastasis dataset generously provided by Shanghai Chest Hospital. This dataset encompasses metastatic data from a cohort of 103 patients, acquired through the utilization of the 1.5 T MRI system

(SignA Elite HD; GE Healthcare, Milwaukee, WI, USA). The dataset includes T2 Fluid Attenuated Inversion Recovery (T2 Flair) and post-contrast T1-weighted (T1ce) images. Within the context of brain metastases, segmentation is facilitated by two distinct labels: the whole tumor division label (WT) and the tumor core division label (TC). The visualization results of the dataset are shown in Figure 1.

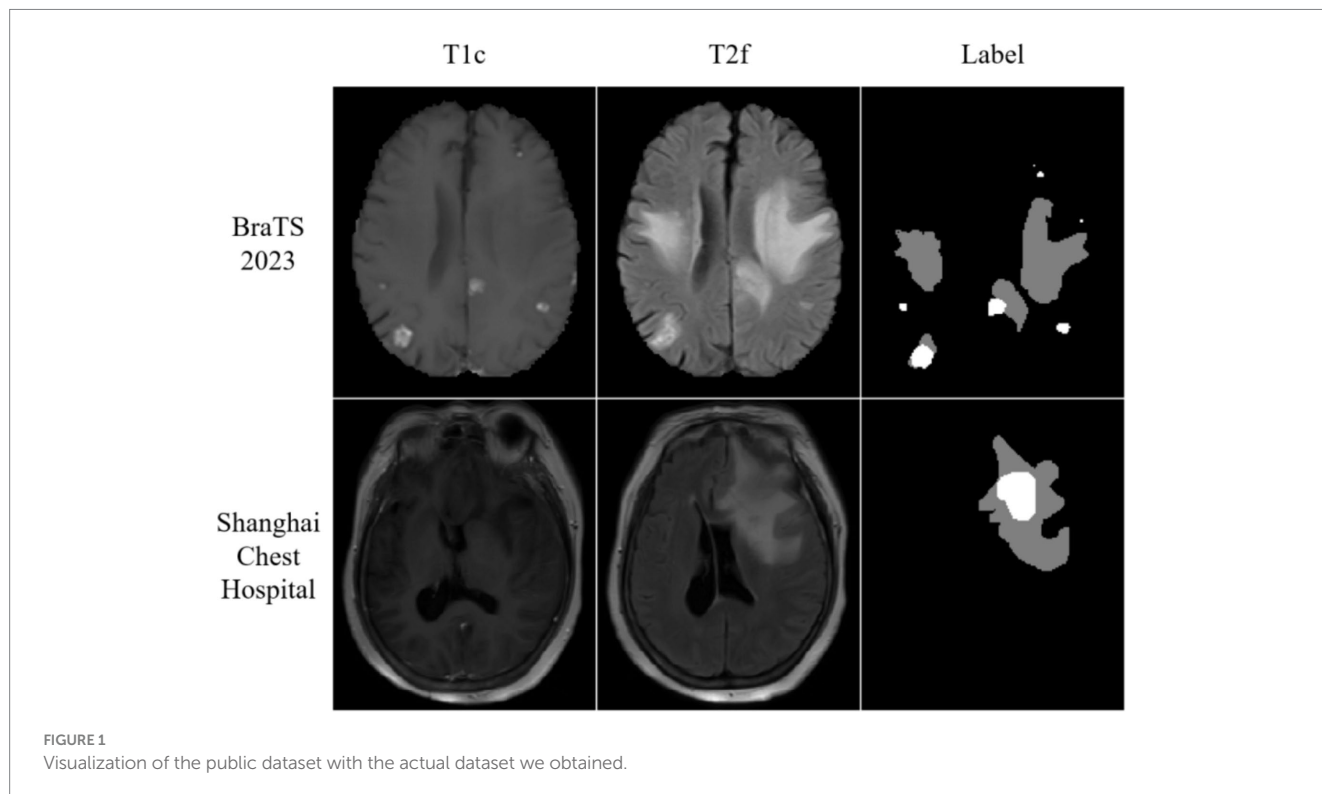
### 2.2 Data preprocessing

In the case of the BraTS2023 Brain Metastases dataset, we opted for post-contrast T1-weighted (t1c) and T2-weighted Fluid Attenuated Inversion Recovery (t2f) as the inputs for our network. To maintain label consistency, we employed the whole tumor division label (Label 2) and the tumor core division label (Label 1 + Label 3).

In accordance with the data supplied by Shanghai Chest Hospital, Flair and T1ce image data underwent regularization using the Z-Score method prior to their integration into the network. Considering that the background in medical images does not provide useful information for segmentation, crop the image to the center region of 160 × 160 and normalize it.

### 2.3 Deep learning network method

To attain precise segmentation of medical images with high accuracy, we consider that segmentation models should amalgamate focalization through convolution and attention mechanisms. Hence, we introduce a coding-decoder structured framework named DRAU-Net. Drawing inspiration from ResU-Net (25), as shown in Figure 2. Our architecture incorporates four convolution blocks on the encoder



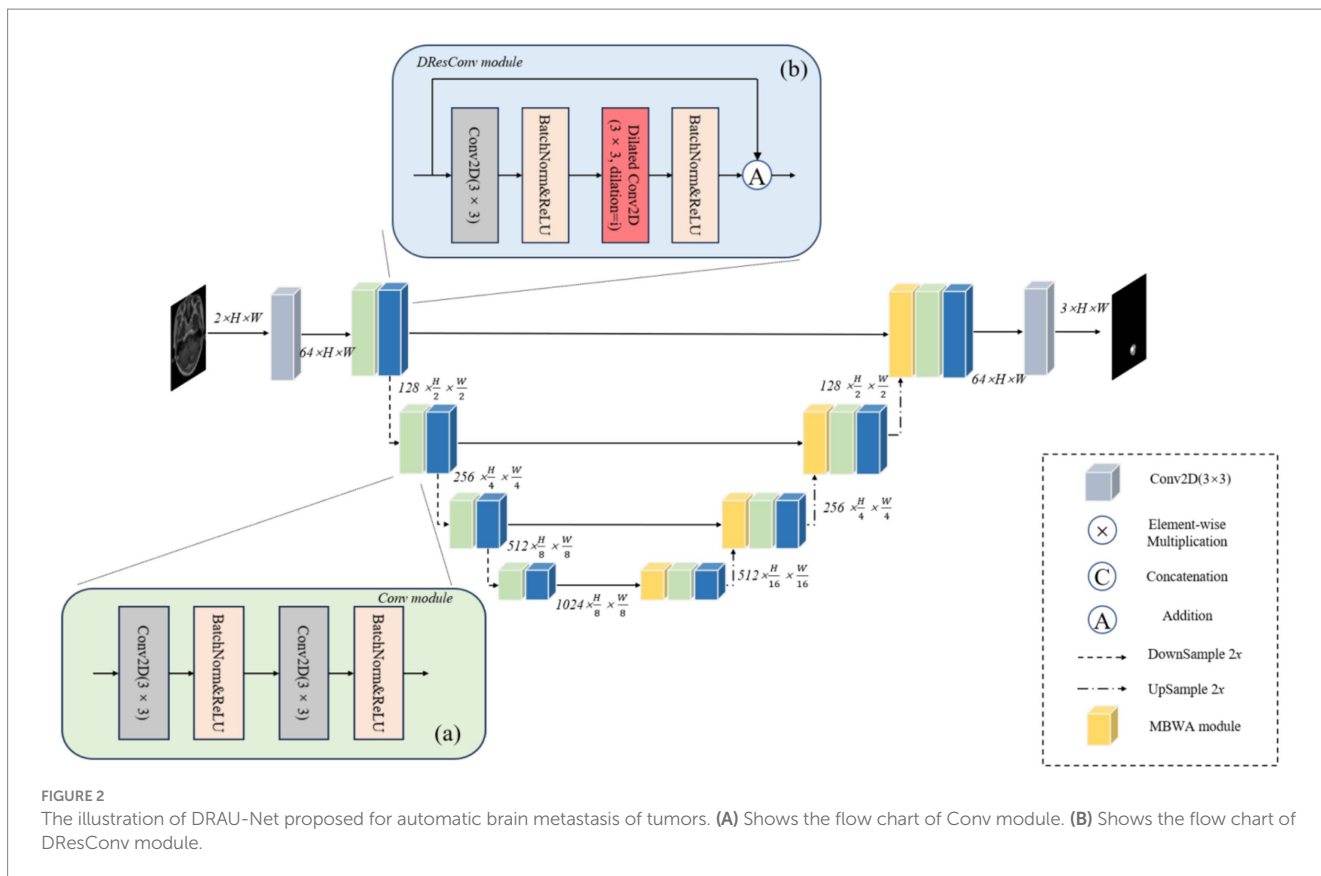


FIGURE 2 The illustration of DRAU-Net proposed for automatic brain metastasis of tumors. (A) Shows the flow chart of Conv module. (B) Shows the flow chart of DResConv module.

path. Post each convolution block, a DResConv module is employed to augment the network expressive capacity, yielding features with varying resolutions through downsample. In the decoder module, we introduce a novel attentional mechanism termed the MBWA module, designed to capture key features across the entire tensor. Subsequently, upsample is achieved through the convolution block and DResConv block, culminating in the final segmentation result. In the subsequent sections, a detailed account of each component’s specific implementation will be provided.

### 2.3.1 Conv module

As shown in Figure 2A. The convolution block comprises two convolution layers, each featuring a 3×3 convolution kernel size and a stride of 1. In the implementation, following each convolutional layer, batch normalization and rectified linear units (ReLU) are applied. Downsample is employed to acquire raw images of diverse sizes, effectively diminishing the computational load of the model, mitigating overfitting, and enhancing the receptive field. This approach not only reduces computational complexity but also promotes a broader sensing field, enabling the subsequent module to effectively capture global information during the learning process.

### 2.3.2 DResConv module

ResNet successfully addresses the challenge of gradient vanishing during deep network training by introducing residual blocks (26). However, comprehending global information without introducing extra parameters remains a critical issue. As illustrated in Figure 2B, we incorporate dilated convolutions with varying dilation rates into the residual block to expand the receptive field without introducing additional parameters, thereby enhancing the model’s ability to

understand global features. Furthermore, convolution layers with distinct dilation rates effectively preserve local details within the image. This facilitates the network in learning a sparser representation of features, thereby capturing the structural information of the image more effectively. The specific implementation process is detailed as follows Equations (1–2):

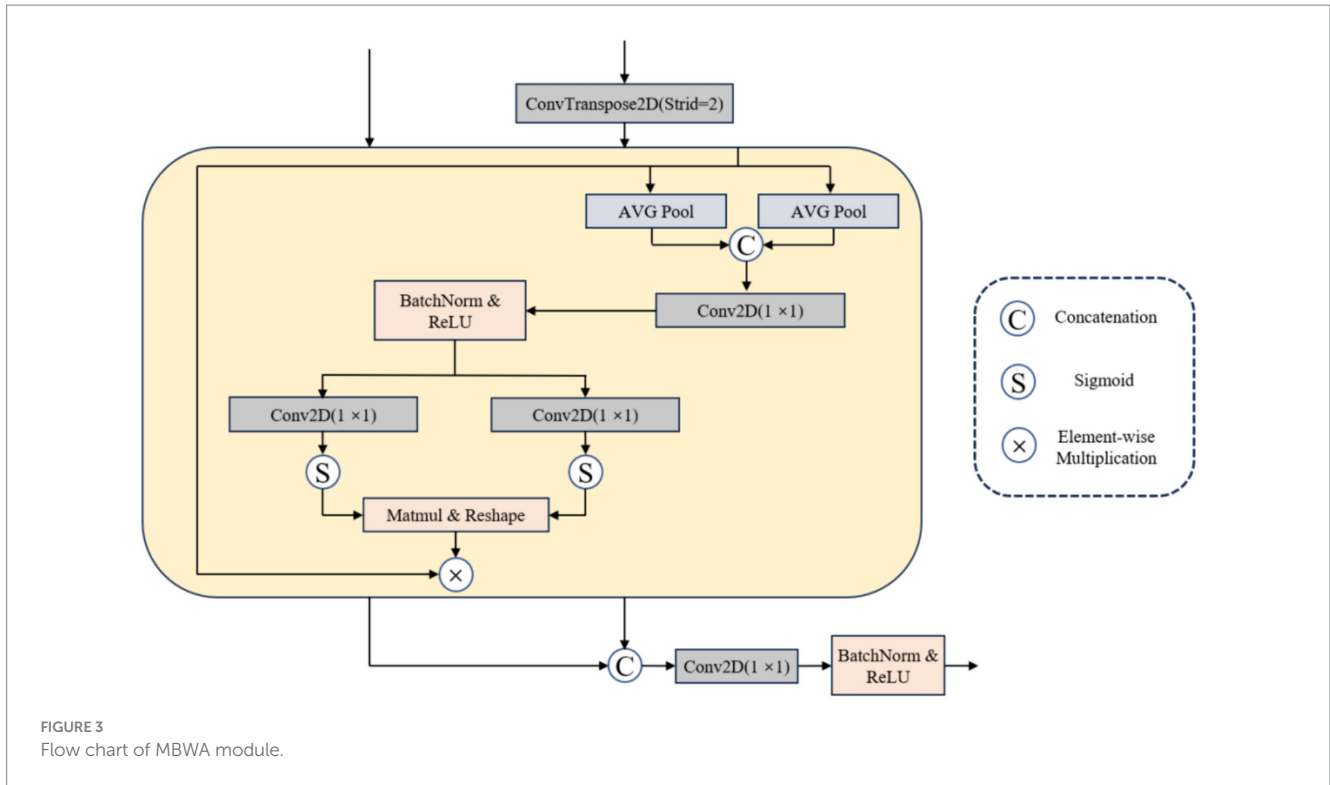
$$F = Conv2D_{kernel=3,dilated=i}(input) \tag{1}$$

$$output = Conv2D_{kernel=3,dilated=i}(F) + input \tag{2}$$

Where *input* and *output* are inputs and output results,  $Conv2D_{kernel=3,dilated=i}$  represents a 2D convolution with a convolution kernel size of 3 and a void rate of *i*, where *i* is the number of layers.

### 2.3.3 MBWA module

Confronted with the challenge of multi-modal metastatic tumor segmentation, while the skip connection in the U-Net network facilitates information flow and transmission (27), it may inadvertently introduce redundant information. Achieving model focus on the target area becomes a significant challenge. In response, this article proposes an effective multi-branch weighted attention (MBWA). As illustrated in Figure 3, the MBWA module incorporates skip connections and utilizes the feature map from the decoder section of the preceding layer as input. Initially, it adjusts the resolution of the decoder features from the previous layer through a 1×1 transposed convolution. Subsequently, these adjusted features undergo weighted attention



coordination within the MBWA module before being concatenated. Finally, a  $1 \times 1$  convolution is applied to adjust the channel dimensions.

The MBWA module is implemented as follows: First, we carry out the maximum average pooling of the output tensor in horizontal dimension and vertical dimension respectively Equations (3–4):

$$x_1 = AVG_{horizontal}(I) \tag{3}$$

$$x_2 = AVG_{vertical}(I) \tag{4}$$

$I$  denote the output tensor, while  $AVG_{horizontal}$  and  $AVG_{vertical}$  symbolize the average pooling of the feature tensors along the horizontal and vertical dimensions, respectively. Employing global pooling along the horizontal dimension enables feature interaction in the spatial domain. This process preserves the positional information  $x_1$  along the horizontal dimension, yielding the spatial attention weight. Meanwhile,  $x_2$  retains long-range dependencies along the vertical dimension and captures positional information in that axis. By concatenating the obtained horizontal and vertical spatial coding information, we subsequently input this combined information into a  $1 \times 1$  convolution layer. This step facilitates the extraction of meaningful spatial features by integrating both horizontal and vertical position information Equation (5).

$$\tilde{x} = ReLU(BN(Conv2D_{kernel=1}(x_1 \times x_2))) \tag{5}$$

$ReLU$  represents the rectified linear unit activation function,  $BN$  denotes the batch normalization operation, and  $Conv2D_{kernel=1}$

signifies a 2D convolution with a convolution kernel size of 1. Following the acquisition of horizontal and vertical spatial coding information  $\tilde{x}$ , a split separation is executed, and the weight map is generated by reinstating the channel count through two  $1 \times 1$  convolutions. Ultimately, the weights are aggregated and applied as weights to the original input Equations (6–9):

$$x_h, x_v = Split(\tilde{x}) \tag{6}$$

$$w_h = \sigma(BN(Conv2D_{kernel=1}(x_h))) \tag{7}$$

$$w_v = \sigma(BN(Conv2D_{kernel=1}(x_v))) \tag{8}$$

$$O = I \times (w_h + w_v) \tag{9}$$

$Split$  denotes separation along spatial dimensions,  $\sigma$  represents the sigmoid activation function,  $BN$  stands for normalization operation, and  $Conv2D_{kernel=1}$  represents a 2D convolution with a convolution kernel size of 1. The MBWA module incorporates positional and spatial information from the input tensor into the output result. The feature coordination across different dimensions within the MBWA module not only tailors the output result to dynamically adjust channel weights but also introduces long-range dependencies in the spatial dimension, thereby enhancing the network’s attention to critical features. This diminishes redundant information and augments the network’s representational capability.

## 2.4 Loss function

Due to the large number of brain metastases and small lesions, we use a combined loss function to constrain the optimization direction of the model and further improve the segmentation results. The loss is given by the following formula Equations (10–12):

$$Loss = \alpha Loss_{Dice} + \beta Loss_{BCE} \quad (10)$$

$$Loss_{BCE} = -[T \log(P) + (1-T) \log(1-P)] \quad (11)$$

$$Loss_{Dice} = 1 - \frac{2|P \cap T|}{|P| + |T|} \quad (12)$$

Where  $T$  represents the ground truth,  $P$  represents the segmentation result,  $\alpha, \beta$  represents the weight of  $Loss_{BCE}$  and  $Loss_{Dice}$ . In this study,  $\alpha, \beta$  are set to 0.7 and 0.3, respectively.

## 2.5 Implementation details

All the experiments from different models were implemented on the server with the following framework: one 12-core Intel 12,700 K CPU, one NVIDIA 3080Ti GPU (12GB), and 32GB RAM. We implement all models on PyTorch. All experimental and comparison models do not use any pre-trained models already trained. The models are trained using the Adam optimizer with an initial learning rate of  $3 \times 10^{-4}$ , the batch size of 8, and the training epoch is set to 150.

## 3 Experiments and result

### 3.1 Evaluation metrics

The model is evaluated using several commonly employed medical image segmentation metrics. The Dice coefficient metric measures the degree of similarity between two samples. Sensitivity measures the proportion of the sample that is correctly segmented. Positive predictive value (PPV) is the proportion of correctly predicted samples among all predicted samples. The lesion number indicator quantifies the proportion of correctly segmented lesions in the entire dataset. The Jaccard index is used to evaluate the intersection-over-union coefficient, as given below Equations (13–16):

$$Dice = \frac{2|p \cap t|}{|p| + |t|} \quad (13)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

$$PPV = \frac{TP}{TP + FP} \quad (15)$$

$$Jaccard = \frac{|p \cap t|}{|p \cup t|} \quad (16)$$

Here,  $p$  represents the ground truth,  $t$  represents the segmentation result.  $TP, FP$  and  $FN$  indicate true-positive, false-positive, and false-negative predictions.

Commonly used medical image segmentation metrics are typically evaluated based solely on the degree of overall image segmentation. However, this evaluation criterion has several limitations. For instance, in the case of the dice, the degree of the larger segmentation region has a greater impact on the overall segmentation metric when multiple segmentation regions in the image are evaluated. As a result, this approach is unable to provide an objective assessment of smaller targets. In the context of metastasis segmentation, the small size of the metastases and the completeness of their segmentation are critical factors that need to be taken into account. To address this issue, we propose a novel segmentation evaluation metric called the multi-objective segmentation integrity metric.

#### ALGORITHM 1 : Multi-objective segmentation integrity metric

**Input:**  $Label_{gt}, Label_{pre}$

**Output:**  $MSIM_{source}$

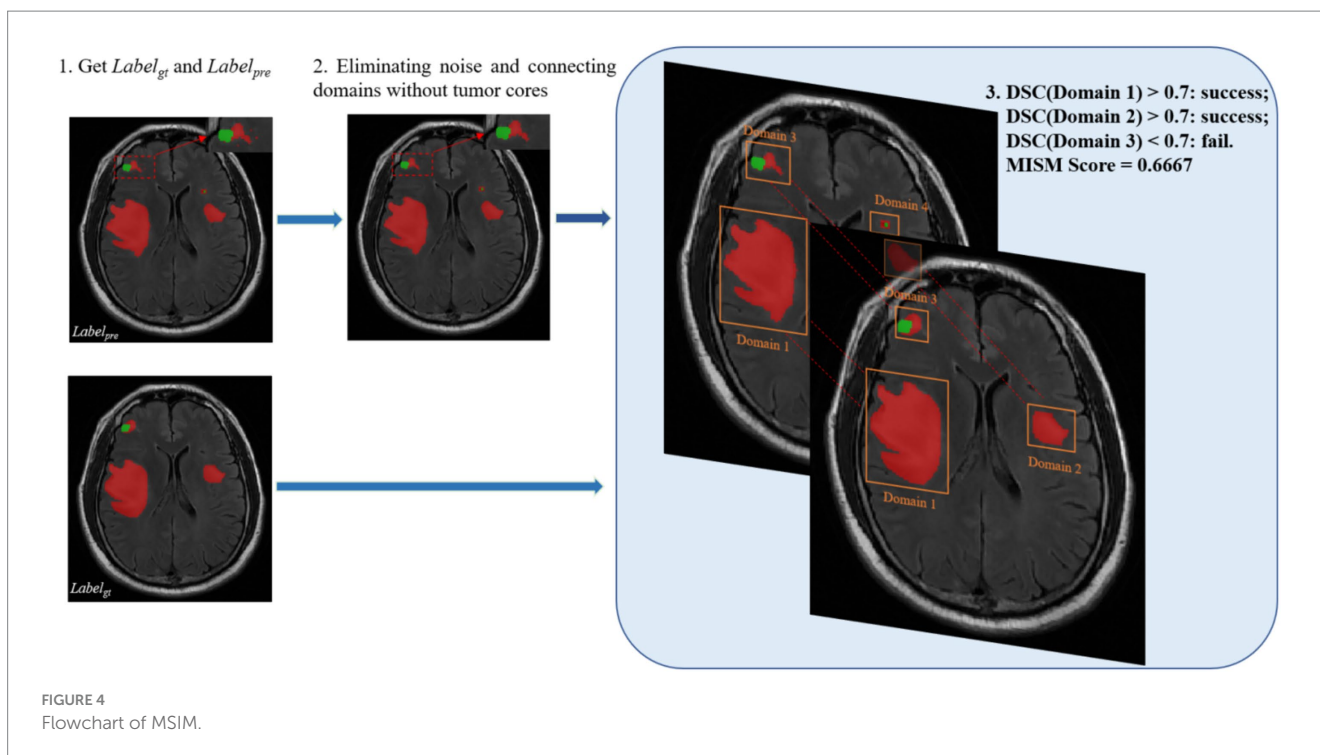
**Define:**  $\delta$ : threshold of area filtering;  $\theta$ : threshold of dice source

```

1: Load  $Label_{gt}$  and  $Label_{pre}$ 
2: Get connected domains:  $g_i \leftarrow Label_{gt}$ 
3: Get connected domains:  $p_i \leftarrow Label_{pre}$ 
4: for  $k = 0 \rightarrow i - 1$  do
5:   Get connected domain size  $s_k \leftarrow g_k$ 
6:   if  $s_k < \delta$  or do not have tumor core label then
7:     Delete  $g_k$ 
8:   end if
9:   Morphological dilation of  $g_k$ 
10:  Morphological erosion of  $g_k$ 
11: end for
12: Sort  $g_i, p_i$  according to the size of the connected domain
13: Get number of connected domains:  $u \leftarrow g_i$ 
14: Get number of connected domains:  $v \leftarrow p_i$ 
15: for  $j = 0 \rightarrow u - 1$  do
16:   $\sigma \leftarrow$  Calculate the dice metric of  $g_i$  and  $p_i$ 
17:  if  $\sigma > \theta$  then
18:     $num + 1$ 
19:  end if
20: end for
21: return  $num / u$ 

```

Algorithm 1 outlines the general workflow of our proposed MSIM. As shown in Figure 4, this evaluation metric involves obtaining all the segmented regions and comparing them with the true segmented regions in pairs, which enables the detection of the true segmentation of each region. The dice coefficient metric is used in the MSIM evaluation metric to determine whether each region has been



successfully segmented, we use 0.7 as the success criterion. The final segmentation metric is calculated as the ratio of the number of successfully segmented regions to the total number of regions in the actual image segmentation. The specific calculation process is as follows Equations (17–18):

$$S = \frac{2|V_P \cap V_T|}{|V_P| + |V_T|} > 0.7 \quad (17)$$

$$MISM = \frac{\sum_i S_i}{\sum_j D_j} \quad (18)$$

Where  $V_P$  is the number of pixels in the predicted target domains, and  $V_T$  is the number of pixels in the target domains in the ground truth.  $S_i$  is the number of successfully segmented targets domains, and  $D_j$  is the number of targets domains in the ground truth.

### 3.2 Comparison with other existing segmentation methods

To evaluate the effectiveness of our DRAU-Net, we compare DRAU-Net with other representative segmentation methods, including U-Net, DenseU-Net (28), AttU-Net (29), U-Net++ (30), U-Net3Plus (31), ResU-Net, TransU-Net (32), LMBiS-Net (33) and HSA-Net (34). All models compared in this study were evaluated without any pre-training strategy, model ensembling, or data augmentation techniques. The results in Table 1 show that the pure U-Net model based on CNN achieved a WT dice of 75.39. The DenseU-Net method, which uses dense layers, improves the U-Net

and achieves a WT dice of 73.68. The U-Net++, which improves skip connections, outperformed other models in terms of PPV. Our proposed DRAU-Net achieves a 3.36 increase in dice compared to the most recent U-Net3Plus. In addition, DRAU-Net outperforms HAS-Net in the Sensitivity metric with a TC score of 77.02 and an ET score of 73.90. For the Jaccard metric, our proposed method leads other models with an average score of 65.18. Regarding the MSIM, the proposed evaluation metric for complete segmentation of BM, DRAU-Net performs the best among all models with a score of 48.80.

In Figure 5, the segmentation results of the compared models are presented from multiple dimensions, and the difficult-to-segment regions are highlighted using red dotted lines. It is evident that accurately segmenting the BM, especially the tumor core, remains a significant challenge for existing methods. Both AttU-Net and U-Net3Plus struggle to delineate the tumors boundaries.

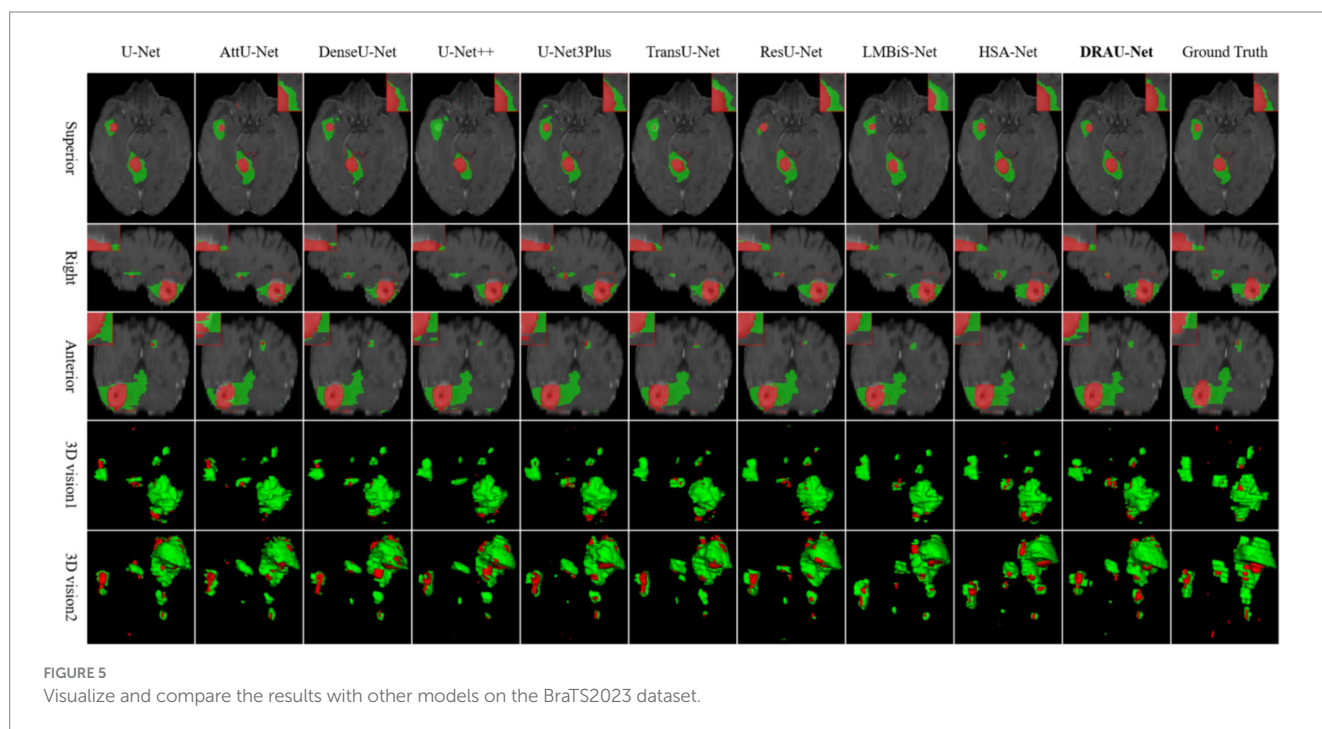
accurately. In contrast, our DRAU-Net demonstrates improved segmentation accuracy. Furthermore, in Figure 5, we show that other methods may miss small tumors when segmenting multiple BM, whereas DRAU-Net accurately segments even small BM.

To further validate the effectiveness of our segmentation framework, this study verified the accuracy of WT and ET segmentation on the data of metastatic tumors provided by Shanghai Chest Hospital dataset. As shown in Table 2, DRAU-Net outshines others with its superior dice scores for whole tumor, achieving 69.52 and 68.95%, respectively. These scores not only surpass those of the well-established U-Net and its derivatives such as DenseU-Net, AttU-Net, and U-Net++, but also significantly outperform other latest models like LMBiS-Net and HSA-Net. Of particular note is that DRAU-Net achieved an average sensitivity and average PPV of 65.32 and 71.38, respectively. This indicates that DRAU-Net not only has a powerful ability to detect key tumor regions, but also can accurately segment key tumor regions, which is crucial for effective medical

TABLE 1 The comparison results between the proposed method and other comparative experiments in BraTS2023 dataset.

Methods	Dice (%)↑		PPV (%)↑		Sensitivity (%)↑		Jaccard (%)↑		MSIM (%)↑
	WT	TC	ET	WT	TC	ET	WT	TC	
U-Net	75.39	66.40	<b>85.87</b>	75.21	70.01	67.04	64.56	55.55	47.61
DenseU-Net	73.68	70.73	81.48	75.10	69.72	68.92	63.45	60.00	40.47
AttU-Net	70.61	65.20	85.32	75.01	64.73	65.12	59.82	54.23	44.04
U-Net++	74.93	68.07	87.93	<b>78.07</b>	69.45	68.21	64.55	57.88	47.61
U-Net3Plus	75.79	68.96	84.14	74.82	70.84	68.87	65.69	58.41	45.23
ResU-Net	73.77	66.75	80.46	69.56	69.82	69.48	63.12	56.31	47.61
Res2U-Net	70.90	64.84	78.80	69.33	67.49	66.63	60.01	54.36	36.90
TransU-Net	70.18	64.35	80.69	71.04	64.97	64.46	59.38	53.44	36.90
LMBIS-Net	73.08	66.42	84.04	72.46	67.25	66.78	62.43	55.78	40.47
HSA-Net	78.38	69.46	84.70	74.67	75.13	70.10	68.33	59.12	47.61
<b>DRAU-Net</b>	<b>79.15</b>	<b>71.47</b>	84.13	74.42	<b>77.02</b>	<b>73.90</b>	<b>68.93</b>	<b>61.43</b>	<b>48.80</b>

Bold represents the optimal value.



diagnosis and treatment planning. In addition, DRAU-Net achieved the best performance under the MSIM metric, indicating its robustness in overall segmentation of metastatic tumors. Compared to other segmentation methods, DRAU-Net also produces more stable segmentation results.

In Figure 6, the segmentation results are visualized. Notably, DRAU-Net has the clearest boundary segmentation for WT and TC among all networks, as can be clearly seen from the enlarged red dotted line. Furthermore, this segmentation approach has the least amount of noise at the segmentation edge. Figure 6 presents the segmentation results from multiple dimensions. From the 2D slices, DRAU-Net achieves more accurate segmentation closer to the ground truth. In the two 3D views, DRAU-Net produces smoother boundaries, fewer surrounding noises, and more detailed segmentation results.

### 3.3 Ablation study

This research proposed DRAU-Net framework elaborates two learnable modules, including DResConv, MBWA. To verify their contributions in performance of the segmentation model, we conducted a series of ablation experiments. The arrangement of each module in the ablation experiments is shown in Table 3.

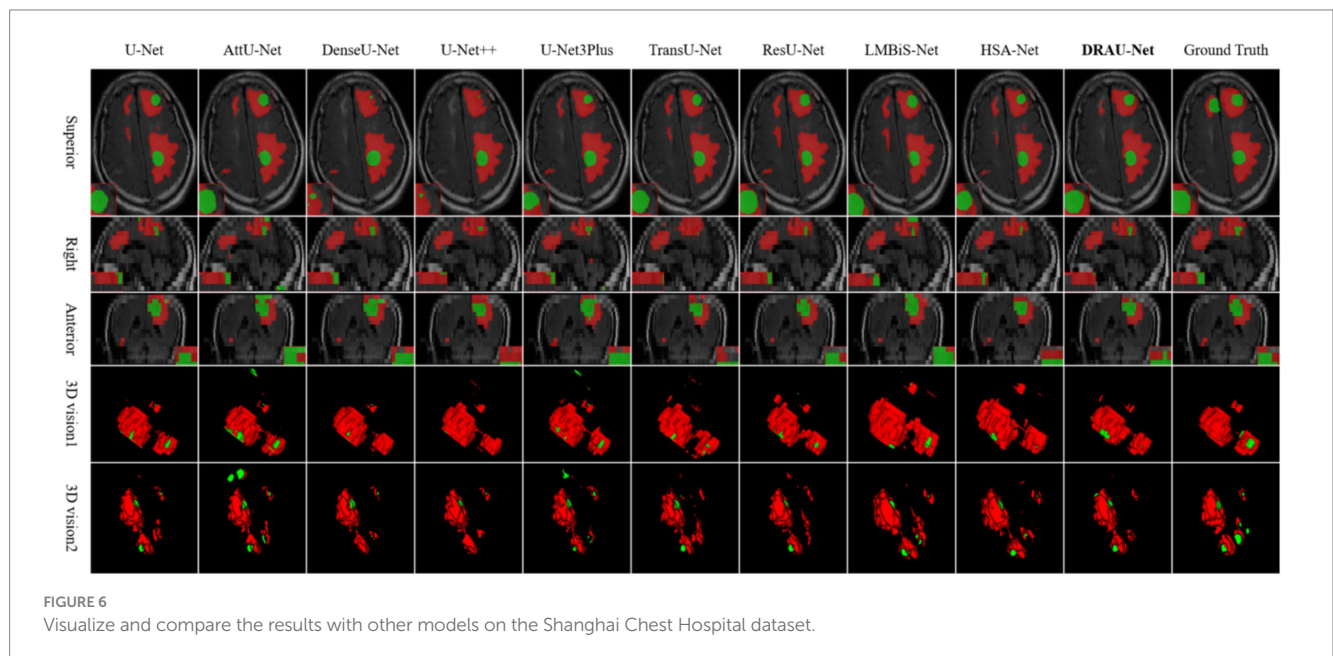
In the context of multi-modal transfer tumors segmentation, attention mechanisms are crucial, as there are typically numerous transfer tumors that are relatively small compared to other tumors. The MBWA module enhances the model feature extraction ability by effectively integrating shallow convolutional features with deep features from the encoder, while simultaneously strengthening the model’s regions of interest. The



TABLE 2 The comparison results between the proposed method and other comparative experiments in Shanghai Chest Hospital dataset.

Methods	Dice (%)↑		PPV (%)↑		Sensitivity (%)↑		Jaccard (%)↑		MSIM (%)↑
	WT	TC	ET	WT	TC	ET	WT	TC	
U-Net	61.17	54.36	56.80	62.85	78.17	64.53	48.74	41.81	51.42
DenseU-Net	63.41	66.76	62.81	84.97	77.88	60.52	52.33	55.30	60.00
AttU-Net	61.21	<b>67.82</b>	57.46	83.58	77.00	62.19	48.35	<b>56.16</b>	54.28
U-Net++	65.90	62.95	64.06	<b>88.37</b>	74.41	54.53	53.19	50.99	51.42
U-Net3Plus	66.85	63.45	66.52	78.45	75.37	57.40	53.87	51.45	54.28
ResU-Net	59.52	67.03	54.38	88.71	74.66	59.75	47.01	56.06	54.28
TransU-Net	55.64	34.28	53.54	46.21	69.35	45.46	42.30	22.80	51.42
LMBIS-Net	64.65	56.63	61.54	66.16	<b>79.53</b>	<b>65.89</b>	51.69	43.66	51.42
HSA-Net	60.61	47.02	57.68	57.12	76.25	65.09	47.62	34.98	48.57
<b>DRAU-Net</b>	<b>69.52</b>	54.19	<b>68.95</b>	61.70	77.71	65.65	<b>56.48</b>	40.73	<b>62.85</b>

Bold represents the optimal value.



DResConv module enhances the ability to extract global information by expanding the receptive field while extracting features. The combination of both modules strengthens the model’s feature extraction ability for small targets. The results of the ablation experiment are presented in Table 4, indicating that the MBWA and DResConv modules independently improve the WT dice index by 1.26 and 0.32, respectively. Furthermore, the combination of both modules improves the WT and TC dice coefficients by 2.14 and 1.48, respectively.

### 4 Discussion

Accurate and effective segmentation of BM lesions is essential for clinical diagnosis and prognosis evaluation. This study proposes a method that aims to segment multiple lesions in clinical data without

TABLE 3 Design scheme of ablation study.

Methods	DResConv	MBWA
Ablation 1		
Ablation 2		√
Ablation 3	√	
Ablation 4	√	√

relying on any pre-trained models. In the dataset of Shanghai Chest Hospital, the collected MRI data is highly heterogeneous. Thus, all images were resampled to 16 layers during preprocessing. However, the small number of layers may result in the loss of features of brain metastases on MRI, and some smaller metastases may be missed. Furthermore, due to time limitations, only two radiologists performed the ground truth labelling.

TABLE 4 Quantitative comparison of ablation results using various modules on BraTS2023 dataset.

Methods	Dice (%)↑		PPV (%)↑		Sensitivity (%)↑		Jaccard (%)↑		MSIM (%)↑
	WT	TC	ET	WT	TC	ET	WT	TC	
Ablation 1	77.01	69.99	83.96	75.15	73.50	70.79	66.65	59.13	73.37
Ablation 2	78.27	70.00	83.67	73.48	75.79	72.05	67.93	59.79	75.16
Ablation 3	77.33	70.00	83.87	73.48	75.79	72.05	67.93	59.79	75.00
Ablation 4	79.15	71.47	84.13	74.42	77.02	73.90	68.93	64.43	48.80

The performance of the proposed model was compared with other existing U-Net-based models on the BraTS2023 and Shanghai Chest Hospital datasets. In addition, we also studied and compared other deep-learning methods related to BM. Our proposed method and data are both yield good performance among similar methods.

In the ablation experiments, compared to the proposed model, Ablation 3 without MBWA and Ablation 2 without DResConv both show a significant decrease in the dice and PPV index of the segmentation task, indicating that both modules play an important role in the accuracy of model segmentation. In the experiments, we explore the application of Transformer-based attention mechanisms; however, the results are not satisfactory. Transformers require dividing the sequence into multiple subspaces, which may lead to different heads capturing similar or redundant information. The presence of redundant information can hinder the model's ability to learn crucial features, resulting in decreased performance when dealing with clinical thin-layer BMs. DRAU-Net achieves the most balanced results on three indicators: dice, PPV, and sensitivity, confirming that the combination of the proposed attention mechanism and convolution is more helpful in segmenting brain metastases.

The large number and small size of brain metastases may mislead the segmentation result evaluation. As each segmentation target is critical, we propose MSIM to evaluate the complete segmentation of tumors. Based on the experimental results, the previous typical deep learning models such as U-Net also achieved a good result in our dataset based on dice. However, the MSIM of U-Net3Plus is 45.23, which is far below the result of our DRAU-Net with an MSIM of 48.80. More importantly, the disparity between the best and worst segmentation results, as measured by the dice, is merely 8.26. However, the difference between MSIM is 11.9. In comparison to the dice, the MSIM metric makes up for the deficiency of evaluation of multiple small lesions and results in superior evaluation performance.

However, the model has limitations due to the differences in image quality from different clinical centers. Firstly, due to the difficulty of data acquisition and cleaning, the data collected from Shanghai Chest Hospital in this study only included two modalities: T2 Flair and T1ce. This limitation resulted in slightly lower segmentation accuracy compared to public datasets. Moreover, since the format of multi-center datasets is often non-uniform, achieving a uniform size through resampling often leads to a loss of detail in the original images, reducing segmentation accuracy. Additionally, due to the limited availability of doctors and the extremely time-consuming process of labeling metastatic tumors, the clinical data in this study included only whole tumor division labels and tumor core division labels. This restriction has led to limited utility of the segmentation results for auxiliary diagnosis. In future work, the plan is to collect and expand the dataset by inviting more experts to annotate the data to

reduce annotation errors and implement domain adaptation and data augmentation strategies to enhance segmentation accuracy.

## 5 Conclusion

As a secondary malignant tumor, metastatic tumors present significant challenges in clinical identification due to their complex shape, size, and distribution. In this paper, a multimodal automatic segmentation method for brain metastases based on the U-Net structure is proposed, designed to assist doctors in quickly identifying and locating brain metastases. This approach aims to optimize diagnosis and treatment plans, thereby improving patient outcomes. DRAU-Net captures more remote dependency information through the DResConv module, enhancing the feature extraction capability for small targets. The MBWA module integrates positional and spatial information from the images into the segmentation results, reducing redundant information while increasing focus on critical features. DRAU-Net has been validated on several datasets, demonstrating superior segmentation results compared to mainstream segmentation methods. Additionally, this research introduces the multi-objective segmentation integrity metric, which emphasizes the segmentation integrity of small target regions within multi-target tasks, providing a more objective evaluation for complex segmentation challenges such as BM segmentation.

In the future, the plan is to further optimize the DRAU-Net algorithm by exploring more efficient convolutional operations and attention mechanisms to enhance the model's robustness. Additionally, domain adaptation and diffusion models will be incorporated to extend the application of DRAU-Net to other types of tumors and complex lesion segmentation. Finally, multi-center clinical trials will be conducted to verify the performance of DRAU-Net across different clinical settings and devices, ensuring its reliability and applicability in practical applications.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

## Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

BL: Data curation, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing, Investigation. QS: Data curation, Formal analysis, Funding acquisition, Resources, Writing – review & editing, Investigation. XF: Conceptualization, Project administration, Resources, Supervision, Writing – review & editing, Investigation. YY: Investigation, Writing – original draft. XL: Data curation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported in part by the University Synergy Innovation Program of Anhui Province under Grant GXXT-2021-006.

## References

- Boire A, Brastianos PK, Garzia L, Valiente M. Brain metastasis. *Nat Rev Cancer*. (2020) 20:4–11. doi: 10.1038/s41568-019-0220-y
- Suh JH, Kotecha R, Chao ST, Ahluwalia MS, Sahgal A, Chang EL. Current approaches to the management of brain metastases. *Nat Rev Clin Oncol*. (2020) 17:279–99. doi: 10.1038/s41571-019-0320-3
- Jyothi P, Singh AR. Deep learning models and traditional automated techniques for brain tumor segmentation in MRI: a review. *Artif Intell Rev*. (2023) 56:2923–69. doi: 10.1007/s10462-022-10245-x
- Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging*. (2017) 8:171–82. doi: 10.1007/s13244-016-0534-1
- Zhang L, Liu X, Xu X, Liu W, Jia Y, Chen W, et al. An integrative non-invasive malignant brain tumors classification and Ki-67 labeling index prediction pipeline with radiomics approach. *Eur J Radiol*. (2023) 158:110639. doi: 10.1016/j.ejrad.2022.110639
- Deng Y, Lan L, You L, Chen K, Peng L, Zhao W, et al. Automated CT pancreas segmentation for acute pancreatitis patients by combining a novel object detection approach and U-net. *Biomed Sign Proces Cont*. (2023) 81:104430. doi: 10.1016/j.bspc.2022.104430
- Hsu DG, Ballangrud Å, Shamseddine A, Deasy JO, Veeraraghavan H, Cervino L, et al. Automatic segmentation of brain metastases using T1 magnetic resonance and computed tomography images. *Phys Med Biol*. (2021) 66:175014. doi: 10.1088/1361-6560/ac1835
- Bezdan T, Milosevic S, Venkatachalam K, Zivkovic M, Bacanin N, Strumberger I. (2021). “Optimizing convolutional neural network by hybridized elephant herding optimization algorithm for magnetic resonance image classification of glioma brain tumor grade.” in *Zooming innovation in consumer technologies conference (ZINC)*. IEEE. p. 171–176.
- Aljohani M, Bahgat WM, Balaha HM, AbdulAzeem Y, El-Abd M, Badawy M, et al. An automated metaheuristic-optimized approach for diagnosing and classifying brain tumors based on a convolutional neural network. *Res Eng*. (2024) 23:102459. doi: 10.1016/j.rineng.2024.102459
- Li X, Fang X, Yang G, Su S, Zhu L, Yu Z. TransU<sup>2</sup>-net: an effective medical image segmentation framework based on transformer and U<sup>2</sup>-net. *IEEE J Transl Eng Health Med*. (2023) 11:441–50. doi: 10.1109/JTEHM.2023.3289990
- Qiu Q, Yang Z, Wu S, Qian D, Wei J, Gong G, et al. Automatic segmentation of hippocampus in hippocampal sparing whole brain radiotherapy: a multitask edge-aware learning. *Med Phys*. (2021) 48:1771–80. doi: 10.1002/mp.14760
- Hu H, Guan Q, Chen S, Ji Z, Lin Y. Detection and recognition for life state of cell cancer using two-stage cascade CNNs. *IEEE/ACM Trans Comput Biol Bioinform*. (2017) 17:887–98. doi: 10.1109/TCBB.2017.2780842
- Shirokikh B, Dalechina A, Shevtsov A, Krivov E, Kostjuchenko V, Durgaryan A, et al. Systematic clinical evaluation of a deep learning method for medical image segmentation: radiosurgery application. *IEEE J Biomed Health Inform*. (2022) 26:3037–46. doi: 10.1109/JBHI.2022.3153394
- Pennig L, Shahzad R, Caldeira L, Lennartz S, Thiele F, Goertz L, et al. Automated detection and segmentation of brain metastases in malignant melanoma: evaluation of

## Acknowledgments

Thank you to The International Brain Tumor Segmentation (BraTS) challenge and Shanghai Chest Hospital for providing the data.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

a dedicated deep learning model. *Am J Neuroradiol*. (2021) 42:655–62. doi: 10.3174/ajnr.A6982

15. Jalalifar SA, Soliman H, Sahgal A, Sadeghi-Naini A. Automatic assessment of stereotactic radiation therapy outcome in brain metastasis using longitudinal segmentation on serial MRI. *IEEE J Biomed Health Inform*. (2023) 27:2681–92. doi: 10.1109/JBHI.2023.3235304

16. Yoo SK, Kim TH, Chun J, Choi BS, Kim H, Yang S, et al. Deep-learning-based automatic detection and segmentation of brain metastases with small volume for stereotactic ablative radiotherapy. *Cancers*. (2022) 14:2555. doi: 10.3390/cancers14102555

17. Dikici E, Ryu JL, Demirel M, Bigelow M, White RD, Slone W, et al. Automated brain metastases detection framework for T1-weighted contrast-enhanced 3D MRI. *IEEE J Biomed Health Inform*. (2020) 24:2883–93. doi: 10.1109/JBHI.2020.2982103

18. Xue J, Wang B, Ming Y, Liu X, Jiang Z, Wang C, et al. Deep learning-based detection and segmentation-assisted management of brain metastases. *Neuro-Oncology*. (2020) 22:505–14. doi: 10.1093/neuonc/noz234

19. Zhou Z, Sanders JW, Johnson JM, Gule-Monroe M, Chen M, Briere TM, et al. MetNet: computer-aided segmentation of brain metastases in post-contrast T1-weighted magnetic resonance imaging. *Radiother Oncol*. (2020) 153:189–96. doi: 10.1016/j.radonc.2020.09.016

20. Grøvik E, Yi D, Iv M, Tong E, Nilsen LB, Latysheva A, et al. Handling missing MRI sequences in deep learning segmentation of brain metastases: a multicenter study. *NPJ Dig Med*. (2021) 4:33. doi: 10.1038/s41746-021-00398-4

21. Chartrand G, Emiliani RD, Pawlowski SA, Markel DA, Bahig H, Cengarle-Samak A, et al. Automated detection of brain metastases on T1-weighted MRI using a convolutional neural network: impact of volume aware loss and sampling strategy. *Magn Reson Imaging*. (2022) 56:1885–98. doi: 10.1002/jmri.28274

22. Zhou Z, Sanders JW, Johnson JM, Gule-Monroe MK, Chen MM, Briere TM, et al. Computer-aided detection of brain metastases in T1-weighted MRI for stereotactic radiosurgery using deep learning single-shot detectors. *Radiology*. (2020) 295:407–15. doi: 10.1148/radiol.2020191479

23. Karargyris A, Umerton R, Sheller MJ, Aristizabal A, George J, Wüest A, et al. Federated benchmarking of medical artificial intelligence with MedPerf. *Nat Mach Intel*. (2023) 5:799–810. doi: 10.1038/s42256-023-00652-2

24. Moawad AW, Janas A, Baid U, Ramakrishnan D, Jekel L, Krantchev K, et al. (2023). The brain tumor segmentation (BraTS-METS) challenge 2023: brain metastasis segmentation on pre-treatment MRI. ArXiv. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10312806/>

25. Zhang Z, Liu Q, Wang Y. Road extraction by deep residual u-net. *IEEE Geosci Remote Sens Lett*. (2018) 15:749–53. doi: 10.1109/LGRS.2018.2802944

26. He K, Zhang X, Ren S, Sun J. (2016). “Deep residual learning for image recognition.” in *Proceedings of the IEEE conference on computer vision and pattern recognition IEEE*. p. 770–778.

27. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation In: *Medical image computing and computer-assisted intervention-MICCAI 2015*. 18th international conference, Munich, Germany. Cham: Springer International Publishing (2015). 234–41.

28. Kaku A, Hegde CV, Huang J, Chung S, Wang X, Young M, et al. (2019). DARTS: DenseUnet-based automatic rapid tool for brain segmentation. arXiv preprint. arXiv:191105567. Available at: <https://arxiv.org/abs/1911.05567>
29. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. (2018). Attention U-net: Learning where to look for the pancreas. Available at: <http://arxiv.org/abs/1804.03999>
30. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: a nested U-net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*. (2018) 11045:3–11. doi: 10.1007/978-3-030-00889-5\_1
31. Huang H, Lin L, Tong R, Hu H, Zhang Q, Iwamoto Y, et al. (2020). “Unet 3+: a full-scale connected unet for medical image segmentation.” in *ICASSP 2020–2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE. pp.1055–1059.
32. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, et al. (2021). Transunet: transformers make strong encoders for medical image segmentation. arXiv preprint. arXiv:2102.04306. Available at: <https://arxiv.org/abs/2102.04306>
33. Matloob Abbasi M, Iqbal S, Aurangzeb K, Alhussein M, Khan TM. LMBiS-net: a lightweight bidirectional skip connection based multipath CNN for retinal blood vessel segmentation. *Sci Rep*. (2024) 14:15219. doi: 10.1038/s41598-024-63496-9
34. Yu Z, Li X, Li J, Chen W, Tang Z, Geng D. HSA-net with a novel CAD pipeline boosts both clinical brain tumor MR image classification and segmentation. *Comput Biol Med*. (2024) 170:108039. doi: 10.1016/j.combiomed.2024.108039