

## Missing Data Imputation via Conditional Generator and Correlation Learning for Multimodal Brain Tumor Segmentation

Tongxue Zhou, Pierre Vera, Stéphane Canu, Su Ruan

## ► To cite this version:

Tongxue Zhou, Pierre Vera, Stéphane Canu, Su Ruan. Missing Data Imputation via Conditional Generator and Correlation Learning for Multimodal Brain Tumor Segmentation. Pattern Recognition Letters, 2022, 158, pp.125-132. 10.1016/j.patrec.2022.04.019. hal-03710281

## HAL Id: hal-03710281 https://hal.science/hal-03710281

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0167865522001155 Manuscript 1fac65329103fc10ff3d12ca9eef7a91

## Missing Data Imputation via Conditional Generator and Correlation Learning for Multimodal Brain Tumor Segmentation

Tongxue Zhou<sup>*a,b,c,d*</sup>, Pierre Vera<sup>*e*</sup>, Stéphane Canu<sup>*c,d*</sup> and Su Ruan<sup>*b,d,\**</sup>

<sup>a</sup> School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China
 <sup>b</sup>LITIS - QuantIF, Université de Rouen Normandie, Rouen 76183, France
 <sup>c</sup>LITIS - Apprentissage, INSA Rouen, Rouen 76800, France
 <sup>d</sup>LITIS, Normandie Univ, INSA Rouen, UNIROUEN, UNIHAVRE, LITIS, Rouen 76800, France
 <sup>e</sup> Centre Henri Becquerel, Rouen 76038, France

### ARTICLE INFO

Keywords: Brain tumor segmentation Conditional generator Correlation learning Missing data Multimodal fusion

## ABSTRACT

Brain tumor is one of the most high-risk cancers which causes the 5-year survival rate of only about 36%. Accurate diagnosis of brain tumor is critical for the treatment planning. However, it's common to missing one modality in clinical scenarios. In this paper, we propose a novel brain tumor segmentation network to impute the missing data. The proposed network consists of a conditional generator, a multi-source correlation network and a segmentation network. To impute the missing data, we propose to use a conditional generator to generate the missing modality under the condition of the available modalities. As the multi MR modalities have a strong relationship in tumor regions, we design a multi-source correlation network to learn the multi-source correlation. On the one hand, the multi-source correlation network can help the conditional generator to generate the missing modality which should keep the consistent correlation with the available modalities. On the other hand, it can guide the segmentation network to learn the correlated feature representations to improve the segmentation performance. The experiments evaluated on BraTS 2018 dataset demonstrate the superior performance of the proposed method when compared with the state-of-the-art methods.

## 1. Introduction

Brain tumor are the abnormal growth of cells inside the brain. Gliomas are one of the most common types of brain tumors, which can primarily or secondarily grow in the Central Nervous System (CNS) and lead to death [1, 2]. Manually segmenting brain tumors requires clinical expertise, it is time-consuming and challenging. Therefore, developing an automatic brain tumor segmentation network is of critical importance to improve diagnosis to perform surgery and to make treatment planning. Nowadays, Magnetic Resonance Imaging (MRI) is an effective technique for radiologist to diagnose the brain tumor and make the treatment planning, which uses a magnetic field and radio frequency waves to give a detailed view of the soft tissues of the brain. In addition, different MR sequences can provide complementary information to improve the diagnosis results [3]. The widely used MR sequences are Fluid Attenuation Inversion Recovery (FLAIR), T1-weighted (T1), contrastenhanced T1-weighted (T1c) and T2-weighted (T2) images, which are presented in Fig. 1. In this work, we consider these MR sequences as different modalities. However, due to the acquisition protocol, image corruption or scanning cost, it usually happens that one modality is missing in clinical practice. This situation makes segmentation more difficult due to incomplete datasets and there is no prior information to indicate which modality will be missing.

In this paper, we propose to combine conditional generator and multi-source correlation to help the brain tumor segmentation in the case of missing data. The preliminary



**Figure 1:** The commonly used four MR modalities: T1, FLAIR, T1c and T2. GT is the corresponding ground truth. Enhancing tumor and necrotic tumor are shown in blue, edema in yellow and enhancing tumor in red

conference version [4] and the journal version [5] simply replaced a missing modalities by an available modality. This was a weak point of the method. This paper extended the previous works by designing an additional conditional generator to generate a representation of the missing modality, which highlights the tumor regions because of the sharing encoders with the segmentation network. Specifically, we use the missing modality index as the condition. The condition imposed on the generator makes it possible to generate the missing modality in a more relevant and supervised way. The whole network architecture is totally different, and more comparison experiments and discussions are addressed. This new work has improved the results of the previous works. The main contributions of our method are:

1) To impute the missing modality, a conditional generator is designed to generate the missing modality with multisource correlation constraint among the modalities.

2) A multi-source correlation network is introduced to learn the latent multi-source correlation among modalities.

🖹 su.ruan@univ-rouen.fr (S. Ruan)

T. Zhou, P. Vera, S. Canu and S. Ruan: Preprint submitted to Elsevier

<sup>\*</sup>Corresponding author

On the one hand, it can help the conditional generator to emphasize correlated tumor features of the missing modality. On the other hand, it can guide the segmentation network to learn the correlated features to achieve the better performance.

3) The multi-modal MR brain tumor segmentation network utilizing the multi-source correlation to conditionally generate the missing modality is first proposed.

4) The extensive experimental results and discussions demonstrate the effectiveness of our proposed method and the superior performance compared with the state-of-the-art approaches.

The paper is organized as follows: Section 2 introduced the related works. Section 3 details our proposed method. Section 4 presents the experimental settings. Section 5 presents the experimental results. Section 6 concludes this work.

## 2. Related works

Over the past few years, many conventional [6, 7, 8, 9] and deep learning [10, 11, 12, 13, 14, 15] based approaches have been proposed to automatically segment brain tumors in MRI. These methods apply the full modalities to do the segmentation. However, it is not always possible to have the complete modalities, and it is common to missing one modality in clinical scenarios. Therefore, it is necessary to develop an automatic brain tumor segmentation method to overcome this problem.

Currently, a large amount of researches are conducted to segment the brain tumor with missing data [16, 17, 18, 19, 20, 21, 22, 23]. For example, Havaei et al. [16] proposed to first project each available image into a single latent representation space. Then, these single latent representations are merged via computing the mean and variance to achieve the segmentation. Lau et al. [17] proposed to map a variable number of input modalities into a unified representation by calculating the mean for the final segmentation. Chartsias et al. [18] proposed to minimize the L1 or L2 distance of features from different modalities. Recently, Chen et al. [19] employed the feature disentanglement theory to do the brain tumor segmentation with missing data. Shen et al. [20] proposed a domain adaptation method on feature maps to recover the information from the missing modality. Dorent et al. [21] introduced the variational auto-encoders to cope with the missing modalities issue in brain tumor segmentation, which becomes the state-of-the-art method. Grøvik et al. [22] designed a network to handle the missing modalities based on DeepLab V3 architecture with dilated convolution. Conte et al. [23] proposed to synthesize the missing modalities via Generative Adversarial Network, and then use them for brain tumor segmentation.

Recently, conditional Generative Adversarial Network (cGAN) [24, 25] has demonstrated to be a promising approach for image synthesis. However, the training of GAN is highly unstable and difficult to converge. On the contrary, U-Net [26] is easier to implement and has also been widely used in the synthesis task due to its contracting and expanding



**Figure 2:** The overview of our proposed network, consisting of a conditional generator, a multi-source correlation network and a segmentation network, here  $X_4$  represents the missing modality.

paths in the encoder and decoder. Some related works have been proposed to adopt U-Net as image generator [27, 28]. For example, Emami et al. [27] proposed a frequency-aware attention U-net to generate PET (Positron Emission Tomography) images from MRI for Alzheimer's disease. Han et al. [28] introduced a U-Net based deep neural network to achieve CT generation from MR images for brain tumor segmentation.

## 3. Method

The proposed network consists of three sub-networks, a conditional generator, a multi-source correlation network and a segmentation network. The generator network and segmentation network share the same encoders to emphasize the tumor features in the missing modality. Fig. 2 presents the overview of the network. First, the conditional generator takes the available modalities as inputs to generate the missing modality. Then, an additional encoder is used to extract the individual feature representation from the generated modality. Following that, the multi-source correlation network takes the individual feature representations of the new full modalities to discover the multi-source correlation among modalities and learn the correlated feature representations. Finally, a decoder is introduced to do the brain tumor segmentation.

## 3.1. Conditionally generating the missing modality

The proposed conditional generator is a multi-encoder based U-Net. The architecture is depicted in Fig. 3. Each single encoder consists of a CBR (Convolution, Batch normalization, ReLU) layer and a res\_dil block in each layer. Except the first layer, strided-convolution (*stride* = 2) is used to replace the max-pooling to avoid the loss of spatial information. Here, the res dil block is to applied to increase the receptive field to learn more semantic features. In the last second layer of the encoder, the condition (missing modality index) is added as an additional input layer. Here, we propose to encode the missing modality as an index (0, 1, 2, 3 corresponding to T2, T1c, FLAIR and T1 respectively.) using an embedding layer. Then, we flatten and reshape it to the same size of the feature maps in the encoder. Finally, it is concatenated with the feature maps. The condition can constrain the generator to generate the corresponding missing modality in a more targeted way. In the decoder, a up-sampling layer and



**Figure 3:** The architecture of the proposed conditional generator. We take one encoder and the decoder as an example.

a convolution layer are first used. Then the concatenation is used to combine the features from the corresponding level of the encoders. Following that, a CBR layer and a res\_dil block are used to adjust the number of features and enlarge the receptive field, respectively. Finally, the segmentation result is obtained after a sigmoid activation function.

In this work, we focus on learning the effective features for segmentation. Therefore, we design the encoders shared by the conditional generator and the segmentation network for the available modalities. On the one hand, the shared encoders can reduce the complexity the network. On the other hand, the segmentation network can guide the conditional generator to generate an effective missing modality where the tumor regions are highlighted. Besides, the generated modality can benefit the following segmentation.

#### **3.2.** Learning the multi-source correlation

Since the same tumor regions are observed by different MR modalities for each patient. Therefore, there is a strong correlation in intensity distribution in tumor regions in each pair of MR modalities [4]. In this paper, we introduce a Multi-source Correlation (MC) network to discover the multi-source correlation among modalities based on our previous work [5]. The detailed network architecture is depicted in Fig. 4.

The proposed MC network includes three modules: Parameter Estimation Module, Correlation Expression Module and Correlation Loss. First, each input modality  $\{M_i\}$ , where  $i = \{0, 1, 2, 3\}$ , is fed to the individual encoder to extract the individual feature representation  $f_i(M_i|\theta_i)$ . Here, the generated missing modality obtained by conditional generator is denoted as  $X_4$ . It is noted that only the generated modality needs to train an additional encoder to get the individual feature representation, the available modalities can reuse the encoders from the conditional generator. Then, the parameter estimation module is used to produce the correlation parameters,  $\Gamma_i = \{\alpha_i, \beta_i, \gamma_i, \delta_i\}$ , according to the individual feature representation  $f_i(M_i|\theta_i)$ . The parameter estimation module consists of two fully connected networks. Finally, the correlated feature representation  $g_i(M_i|\theta_i)$  is obtained via the correlation expression module. Based on the observation, we utilize the linear correlation expression (Equation 1) to describe the multi-source correlation. To constrain the estimated correlated feature representation and the original feature representation distributions to be as close as possible, we propose to use a simple and widely



**Figure 4:** The architecture of the proposed multi-source correlation network.

used divergence, Kullback–Leibler divergence, to form the correlation loss function (Equation 2). More comparison results on different divergence functions are presented in Section 5.6.

$$g_i(M_i|\theta_i) = \alpha_i \odot f_j(M_j|\theta_j) + \beta_i \odot f_k(M_k|\theta_k) + \gamma_i \odot f_l(M_l|\theta_l) + \delta_i, (i \neq j \neq k \neq l)$$
(1)

where *M* is the input modality, *i*, *j*, *k* and *l* are the indexes of the modality,  $\theta$  is the network parameters, *f* is the individual feature representation, *g* is the correlated feature representation,  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are the correlation parameters.

$$L_c = \sum_{i=1}^{S} P(f_i) log \frac{P(f_i)}{Q(g_i)}$$
<sup>(2)</sup>

where S is the number of modality,  $P(f_i)$  and  $Q(g_i)$  are the original feature representation distributions and correlated feature representation distributions of modality *i*, respectively.

#### 3.3. Multi-encoder based segmentation network

To take advantage of the different feature information from the multi-modalities, we applied our previous network architecture [29], a multi-encoder based U-Net, to do the brain tumor segmentation.

## 4. Experimental settings

#### 4.1. Dataset and implementation details

The public multi-modal brain tumor segmentation datasets BraTS 2018 [30] is used to evaluate our method. It contains 285 cases, each case has four MR modalities including T1, T1c, T2 and FLAIR. There are three segmentation classes: whole tumor (WT), tumor core (TC) and enhancing tumor (ET). We cropped and resized the images to  $128 \times 128 \times 128$ voxels. The N4ITK method and intensity normalization are used. We implement our proposed network in Keras, a single Nvidia Tesla V100 (32G) is used. We used Nadam optimizer to train the model, the initial learning rate is set as 0.0005, it will reduce by half with patience of 10 epochs. Early stopping is used if the validation loss is not improved over 20 epochs. We randomly split the dataset into 80% training and 20% testing. The experiment results are uploaded to the public online evaluation platform<sup>1</sup> to obtain the final evaluations.

### 4.2. Loss function

In our work, Dice loss (Equation 3) is used as the segmentation loss, L1-Norm (Equation 4) is applied as the generation loss. In addition, Kullback–Leibler divergence is used as the correlation loss (Equation 2). The overall loss function used in our network is defined in Equation 5. To make the loss function of Equation 5 be convex, Nadam (Nesterov-accelerated Adaptive Moment Estimation) optimizer is used to prevent from getting trapped in a local minimum.

$$L_{d} = 1 - 2 \frac{\sum_{i=1}^{C} \sum_{j=1}^{N} p_{ij} g_{ij} + \epsilon}{\sum_{i=1}^{C} \sum_{j=1}^{N} (p_{ij} + g_{ij}) + \epsilon}$$
(3)

where N is the number of the examples, N = 228, C is the number of the classes, C = 3,  $p_{ij}$  is the probability that pixel *i* belongs to the tumor class *j*, the same is true for  $g_{ij}$ , and  $\epsilon$  is a small constant to avoid dividing by 0.

$$L_{g} = \|y_{i} - y_{i}'\|_{1} \tag{4}$$

where y is the real modality, y' is the generated modality.

$$L_{total} = \psi \times L_d + \xi \times L_g + \eta \times L_c \tag{5}$$

where  $\psi = 1, \xi = 0.1$  and  $\eta = 0.1$ .

### 4.3. Performance evaluation

We applied two commonly used evaluation metrics to evaluate our the segmentation performance: Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD).

DSC is a spatial overlap index between the prediction and ground-truth. A larger DSC value denotes a better prediction result.

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{6}$$

where TP, FP and FN denote the number of true positive, false positive and false negative voxels, respectively.

HD is computed between boundaries of the prediction and ground-truth. A smaller HD value indicates a better prediction result.

$$HD = \max\{\max_{s \in S} \min_{r \in R} d(s, r), \max_{r \in R} \min_{s \in S} d(r, s)\}$$
(7)

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

## 5. Experiment results

#### **5.1.** Ablation studies

To demonstrate that the effectiveness of our proposed method. First, we compare our method with two related methods. (1) Replace: The idea is from our previous work [4], which trains the network on full modalities but uses the most correlated available modality to replace the missing one during test. (2) Direct: It directly uses the available modalities to do the segmentation. Then, we gradually add the proposed components (MC: multi-source correlation, G: generator without condition constraint and CG: Conditional Generator) to the 'Direct' method to prove the effectiveness of the proposed components.

From Table 1, we can observe that the 'Replace' method can't have a satisfying result, especially when FLAIR or T1c is missing. The 'Direct' method can outperform the 'Replace' method, but the results are still unsatisfying because one modality is missing and the feature information is not sufficient to obtain a good result. However, when the multi-source correlation is taken into account (Direct + MC), the segmentation results are much improved, especially when FLAIR is missing. The reason is that the multi-source correlation network can guide the network to learn the correlated feature among modalities. In addition, when the generator is integrated to compensate the missing modality, we can observe a significant improvement in terms of average DSC compared with the baseline ('Direct'), especially when FLAIR is missing, a 8.8% improvement in the terms of average DSC is observed. In particular, when the condition is added to the generator, we can achieve the best segmentation results. Compared with 'Direct+MC+G', a 0.6%, 1.3%, 1.4%, 1.4% improvement in the terms of average DSC can be observed when FLAIR, T1, T1c and T2 is missing, respectively. The reason is that the condition can guide the generator to produce the effective feature information of the missing modality in a supervised way than the naive generator. The similar comparison results can be observed with regard to Hausdorff Distance. The proposed method can improve the baseline with 42.4%, 38.1%, 4.0%, 39.3% when FLAIR, T1, T1c, T2 is missing, respectively.

#### 5.2. Compared with the state-of-the-art methods

Furthermore, we compare our method with the state-ofthe-art methods, which have been mentioned in the introduction. Our work focuses on the brain tumor segmentation in the case of missing one modality. Therefore, we compare with other methods in the case of missing one modality. To make it fair, we also compare with a method [20], which addresses only missing one modality issue. The comparison results are illustrated in Table 2. Since the method HeMIS [16] didn't publish the available code, the reported results on HeMIS and U-HeMIS are cited from [21]. It can be observed that our method achieves the best results in the most of

<sup>&</sup>lt;sup>1</sup>https://ipp.cbica.upenn.edu/

Missing Data Imputation via Conditional Generator and Correlation Learning for Multimodal Brain Tumor Segmentation

#### Table 1

Comparison of segmentation performance in terms of DSC and HD. WT, TC, ET denote whole tumor, tumor core and enhancing tumor, respectively. AVG denotes the average results on the three regions, bold results denote the best scores.

Methods			Missing	; FLAIR	1		Missi	ng T1			Missir	ng T1c		Missing T2				
		WT	тс	ET	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	
Replace	DSC	55.1	53.6	67.5	58.7	84.2	77.6	69.7	77.2	85.7	45.8	2.1	44.5	81.8	69.8	72.6	74.7	
	HD	37.9	52.8	14.4	34.9	7.0	6.9	4.9	6.3	6.4	16.0	24.8	15.7	9.0	11.8	3.7	8.2	
Direct	DSC	75.5	76.5	71.9	74.6	84.6	83.7	76.4	81.6	84.8	62.1	40.1	62.3	84.9	84.2	76.8	82.0	
	HD	10.4	9.0	6.2	8.5	7.6	6.9	4.4	6.3	6.8	12.4	10.5	9.9	6.4	6.4	4.0	5.6	
Discourse	DSC	79.9	78.8	72.2	77.0	85.1	83.4	77.1	81.9	85.7	62.8	40.1	62.9	85.4	85.1	77.3	82.6	
Direct+MC	HD	7.9	10.3	8.1	8.8	6.5	5.3	3.0	4.9	6.7	11.8	10.9	9.8	6.8	6.0	3.8	5.5	
Direct+MC+G	DSC	82.9	84.8	75.9	81.2	86.5	84.8	76.3	82.5	86.6	62.6	37.7	62.3	86.5	85.2	76.8	82.8	
	HD	5.9	5.2	3.7	4.9	5.0	4.5	3.4	4.3	4.8	11.6	11.1	9.2	4.8	4.0	2.9	3.9	
Direct+MC+CG (Ours)	DSC	83.5	85.0	76.5	81.7	87.0	86.2	77.7	83.6	86.5	64.0	39.2	63.2	87.1	86.8	78.2	84.0	
	HD	6.4	4.8	3.5	4.9	4.5	4.2	2.9	3.9	5.3	12.0	11.1	9.5	4.4	3.7	2.7	3.6	

#### Table 2

Comparison results among different methods in terms of Dice Similarity Coefficient on BraTS 2018 dataset. WT, TC, ET denote whole tumor, tumor core and enhancing tumor, respectively. AVG denotes the average results on the three regions, bold results denote the best scores.

	Dice Similarity Coefficient (%) (↑)																
Methods		Missing	; FLAIR			Missi	ng T1			Missir	ng T1c		Missing T2				
	WT	тс	ET	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	
HeMIS [16]	44.2	46.6	55.1	48.6	75.6	54.9	60.5	63.7	75.2	18.7	1.0	31.6	70.2	48.8	60.9	60.0	
U-HeMIS [21]	82.1	70.7	69.7	74.2	87.0	72.2	69.7	76.3	87.0	61.0	33.4	60.5	85.1	70.7	69.9	75.2	
URN [17]	81.1	69.5	68.5	73.0	86.5	72.2	69.8	76.2	86.1	52.5	25.8	54.8	85.6	72.0	71.0	76.2	
HVED [21]	83.3	75.3	71.1	76.6	88.6	75.6	71.2	78.5	88.0	61.5	34.1	61.2	86.2	74.2	71.1	77.2	
[20]	61.6	68.0	55.2	61.6	89.0	77.8	64.2	77.0	87.9	57.0	48.4	64.4	89.3	77.5	64.3	77.0	
Ours	83.5	85.0	76.5	81.7	87.0	86.2	77.7	83.6	86.5	64.0	39.2	63.2	87.1	86.6	78.2	84.0	

the cases. In addition, compared to the current state-of-theart method [21] which is based on the variational encoderdecode, our method is superior than it with a large margin. For example, we can improve it with 6.7%, 6.5%, 3.3%, 8.8% in the terms of average DSC when FLAIR, T1, T1c or T2 is missing, respectively. Compared with [20], our method has the clear advantages in the most of the cases.

# **5.3.** Visualization of the segmentation and generation results

To further demonstrate the performance of our method, we visualize an example from BraTS 2018 dataset in Fig. 5. We can observe that the 'Replace' method produces many false predictions (highlighted by green arrows), especially when FLAIR and T1c are missing. 'Direct' method can achieve the better results than 'Replace'. Furthermore, the results can be refined progressively using the proposed components. Our proposed method can detect effectively the tumor regions and its sub-regions. In addition, from the generation results, we can observe that our proposed method can generate the tumor-highlighted images, which can benefit the final segmentation.

## 5.4. Visualization of the feature maps

In addition, we visualize the feature maps from different methods in all the cases of missing one modality in Fig. 6. We can observe that using 'Direct' method can't learn a very effective feature maps because of the incomplete data input. While when the Correlation Constraint (MC) network is integrated, the features in tumor regions become clear and obvious. Moreover, with the help of the conditional generator, a specific feature information from the missing modality is obtained. In this way, the network can leverage the complete and rich feature representations to achieve a better segmentation result.

## 5.5. Analysis on the generator architecture

We also compare our method with cGAN [24]. Specifically, the generator is based on 3D U-Net, which has the same architecture with our proposed conditional generator. The discriminator consists of four down-convolutional layers, following by a Dense layer and a Sigmoid function. The condition is embedded in the last second layer for both generator and discriminator. The loss function of the cGAN is defined in Equation 8.

$$L_{cGAN} = E_{x,y}[log D(x, y)] + E_{x,z}[log(1 - D(x, G(x, z)))] + 0.1L_{g}$$
(8)

where G and D denote the generator and discriminator, respectively, x denotes the condition, y denotes the real image and z denotes the input images,  $L_g$  is defined in Equation 4.

The comparison results are presented in Table 3. It can be seen that cU-Net can generally obtain better segmentation results than cGAN in the terms of DSC and HD. The significant improvement can be observed: 2.4%, 1.5%, 2.2% in the terms of average DSC and 14.0%, 17.0%, 21.7% in the terms of average HD when Flair, T1 or T2 is missing. The visualization results are presented in Fig. 7 in which the average DSC is denoted under each segmentation result.



**Figure 5:** Visualization of the segmentation results. Each row presents the segmentation results from different methods in the case of missing one modality. The green arrow points out the segmentation differences among the different methods. The last column presents the generation results of our method.



**Figure 6**: Visualization of the feature maps. The first column presents the four modalities and the ground-truth. The last four columns present the feature maps corresponding to T1, Flair, T1c and T2 from top to bottom obtained by the four methods. The four missing modality situations are separated in color.

Missing Data Imputation via Conditional Generator and Correlation Learning for Multimodal Brain Tumor Segmentation

#### Table 3

Comparison of segmentation performance in terms of DSC and HD between cU-Net and cGAN. WT, TC, ET denote whole tumor, tumor core and enhancing tumor, respectively. AVG denotes the average results on the three regions, bold results denote the best scores.

Methods			Missing	g FLAIR			Miss	ing T1			Missi	ng T1c		Missing T2				
		WT	тс	ΕT	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	WT	тс	ET	AVG	
cGAN H	DSC	81.8	83.2	74.5	79.8	86.2	83.5	77.6	82.4	85.6	66.8	44.6	65.6	86.1	83.4	77.2	82.2	
	HD	7.0	5.8	4.3	5.7	5.7	5.2	3.3	4.7	5.4	10.4	9.2	8.3	5.6	5.0	3.3	4.6	
cU-Net (Ours)	DSC	83.5	85.0	76.5	81.7	87.0	86.2	77.7	83.6	86.5	64.0	39.2	63.2	87.1	86.8	78.2	84.0	
	HD	6.4	4.8	3.5	4.9	4.5	4.2	2.9	3.9	5.3	12.0	11.1	9.5	4.4	3.7	2.7	3.6	

#### Table 4

Comparison of segmentation in terms of DSC and HD using different divergences. WT, TC, ET denote whole tumor, tumor core and enhancing tumor, respectively. AVG denotes the average results on the three regions, bold results denote the best scores.

Divergences			Missing	g FLAIR			Missi	ng T1			Missir	ng T1c		Missing T2				
Divergence	5	WT	ТC	ET	AVG	WT	ТC	ET	AVG	WT	тс	ET	AVG	WT	ТC	ng T2 ET 78.2 2.7 74.8 4.7 4.0 44.1 19.5 24 1	AVG	
KL	DSC	83.5	85.0	76.5	81.7	87.0	86.2	77.7	83.6	86.5	64.0	39.2	63.2	87.1	86.8	78.2	84.0	
	HD	6.4	4.8	3.5	4.9	4.5	4.2	2.9	3.9	5.3	12.0	11.1	9.5	4.4	3.7	2.7	3.6	
Jeffreys	DSC	75.4	62.3	48.8	62.2	83.9	78.4	73.6	78.6	83.6	48.6	30.4	54.2	83.9	77.2	74.8	78.6	
	HD	16.0	10.7	9.1	11.9	7.6	7.4	4.6	6.5	6.9	16.3	15.5	12.9	8.1	7.9	4.7	6.9	
square Hellinger	DSC	32.3	20.1	4.2	18.9	70.7	23.4	2.8	32.3	69.8	12.1	0.1	27.3	70.4	22.1	4.0	32.2	
	HD	61.2	58.6	49.2	56.3	32.8	33.8	39.1	35.2	33.2	41.3	50.7	41.7	33.0	36.8	44.1	38.0	
Exponential	DSC	56.8	10.2	0.8	22.6	77.3	27.6	17.8	40.9	75.4	13.8	3.1	30.8	76.3	29.7	19.5	41.8	
	HD	45.4	34.5	40.3	40.1	14.6	22.4	26.4	21.1	13.8	24.1	27.5	21.8	15.1	23.2	24.1	20.8	



**Figure 7:** Visualization of the generation and segmentation results between cGAN and cU-Net. The average DSC is denoted under each segmentation result. The segmentation differences are highlighted by green circles.

We can see that both methods can generate the tumorenhanced missing modality well, while cU-Net can further emphasize tumor features, especially when Flair is missing. Consequently, it leads to a more precise segmentation of tumor regions by cU-Net than by cGAN.

#### 5.6. Analysis on the divergence functions

To measure the similarity of the probability distributions, we analyze four commonly used f-divergences: Kullback–Leibler (KL) divergence, Jeffreys divergence (Equation 9), squared Hellinger divergence (Equation 10), and exponential divergence (Equation 11). The compared segmentation results are presented in Table 4. It can be observed that KL divergence achieves the best segmentation performance.

$$D_J(P \| Q) = \sum_{i=1}^{M} (P(f_i) - Q(g_i))(\ln P(f_i) - \ln Q(g_i))$$
(9)

$$H^{2}(P,Q) = \sum_{i=1}^{M} 2(\sqrt{P(f_{i})} - \sqrt{Q(g_{i})})^{2}$$
(10)

$$D_e(P||Q) = \sum_{i=1}^{M} P(f_i) (\ln P(f_i) - \ln Q(g_i))^2$$
(11)

where M is the number of modality,  $P(f_i)$  and  $Q(g_i)$  are the original feature representation distributions and correlated feature representation distributions of modality i, respectively.

### 6. Conclusion

Accurate brain tumor segmentation plays an critical role in the medical imaging field. However, it is difficult to have a complete set of multi-modalities. It usually happens to missing one modality in the real clinical practice. In this paper, we proposed a novel brain tumor segmentation network to impute the missing data. Motivated by the multi-source correlation among MR multi-modalities, we first proposed a multi-source correlation network to learn the correlation

among modalities. Then, the learned multi-source correlation is used to guide the conditional generator to generate the missing modality through the shared encoders with segmentation task, in which the tumor related features are further enhanced. Following that, a multi-encoder based segmentation network is applied on the new complete modality set to do the final brain tumor segmentation. The experimental results and the analyses demonstrate the effectiveness of our proposed method and superior performance of our method compared with the state-of-the-art methods. The proposed method is capable to impute the missing modality, which can provide more complementary information about brain tumor. In addition, the generated modality can assist the clinical doctors to improve tumor diagnosis accuracy and also help them to design the following treatment planning. Although missing one modality is very common in clinical practice, it will be interesting to ameliorate the architecture of the conditional generator and extend it to cover missing any number of modalities. Besides, we will try different correlation expressions to improve the segmentation performance.

## Acknowledgments

This work was co-financed by European regional development fund (ERDF, 18P03390/18E01750/18P02733), Haute-Normandie Regional Council via M2SINUM project, and supported by China Scholarship Council.

## References

- A. Elazab, C. Wang, S. J. S. Gardezi, H. Bai, Q. Hu, T. Wang, C. Chang, B. Lei, Gp-gan: Brain tumor growth prediction using stacked 3d generative adversarial networks from longitudinal mr images, Neural Networks 132 (2020) 321–332.
- [2] M. L. Goodenberger, R. B. Jenkins, Genetics of adult glioma, Cancer genetics 205 (12) (2012) 613–621.
- [3] T. Zhou, S. Ruan, S. Canu, A review: Deep learning for medical image segmentation using multi-modality fusion, Array 3 (2019) 100004.
- [4] T. Zhou, S. Canu, P. Vera, S. Ruan, Brain tumor segmentation with missing modalities via latent multi-source correlation representation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 533–541.
- [5] T. Zhou, S. Canu, P. Vera, S. Ruan, Latent correlation representation learning for brain tumor segmentation with missing mri modalities, IEEE Transactions on Image Processing 30 (2021) 4263–4274.
- [6] A. Khosravanian, M. Rahmanimanesh, P. Keshavarzi, S. Mozaffari, Fast level set method for glioma brain tumor segmentation based on superpixel fuzzy clustering and lattice boltzmann method, Computer Methods and Programs in Biomedicine 198 (2021) 105809.
- [7] J. Lapuyade-Lahorgue, J.-H. Xue, S. Ruan, Segmenting multi-source images using hidden markov fields with copula-based multivariate statistical distributions, IEEE Transactions on Image Processing 26 (7) (2017) 3187–3195.
- [8] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, Y. Zhu, Kernel feature selection to fuse multi-spectral mri images for brain tumor segmentation, Computer Vision and Image Understanding 115 (2) (2011) 256– 269.
- [9] D. Zikic, B. Glocker, E. Konukoglu, A. Criminisi, C. Demiralp, J. Shotton, O. M. Thomas, T. Das, R. Jena, S. J. Price, Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel mr, in: International Conference on Medical Image

Computing and Computer-Assisted Intervention, Springer, 2012, pp. 369–376.

- [10] M. I. Sharif, J. P. Li, M. A. Khan, M. A. Saleem, Active deep neural network features selection for segmentation and recognition of brain tumors using mri images, Pattern Recognition Letters 129 (2020) 181–189.
- [11] Y. Ding, L. Gong, M. Zhang, C. Li, Z. Qin, A multi-path adaptive fusion network for multimodal brain tumor segmentation, Neurocomputing 412 (2020) 19–30.
- [12] A. Tiwari, S. Srivastava, M. Pant, Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019, Pattern Recognition Letters 131 (2020) 244–260.
- [13] Y. Liu, W. Cui, Q. Ha, X. Xiong, X. Zeng, C. Ye, Knowledge transfer between brain lesion segmentation tasks with increased model capacity, Computerized Medical Imaging and Graphics 88 (2021) 101842.
- [14] M. Rahimpour, J. Bertels, D. Vandermeulen, F. Maes, K. Goffin, M. Koole, Improving 11w mri-based brain tumor segmentation using cross-modal distillation, in: Medical Imaging 2021: Image Processing, Vol. 11596, International Society for Optics and Photonics, 2021, p. 115960Z.
- [15] M. U. Rehman, S. Cho, J. Kim, K. T. Chong, Brainseg-net: Brain tumor mr image segmentation via enhanced encoder–decoder network, Diagnostics 11 (2) (2021) 169.
- [16] M. Havaei, N. Guizard, N. Chapados, Y. Bengio, Hemis: Heteromodal image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 469–477.
- [17] K. Lau, J. Adler, J. Sjölund, A unified representation network for segmentation with missing modalities, arXiv preprint arXiv:1908.06683.
- [18] A. Chartsias, T. Joyce, M. V. Giuffrida, S. A. Tsaftaris, Multimodal mr synthesis via modality-invariant latent representation, IEEE transactions on medical imaging 37 (3) (2017) 803–814.
- [19] C. Chen, Q. Dou, Y. Jin, H. Chen, J. Qin, P.-A. Heng, Robust multimodal brain tumor segmentation via feature disentanglement and gated fusion, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 447–456.
- [20] Y. Shen, M. Gao, Brain tumor segmentation on mri with missing modalities, in: International Conference on Information Processing in Medical Imaging, Springer, 2019, pp. 417–428.
- [21] R. Dorent, S. Joutard, M. Modat, S. Ourselin, T. Vercauteren, Heteromodal variational encoder-decoder for joint modality completion and segmentation, in: D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, A. Khan (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 74–82.
- [22] E. Grøvik, D. Yi, M. Iv, E. Tong, L. B. Nilsen, A. Latysheva, C. Saxhaug, K. D. Jacobsen, Å. Helland, K. E. Emblem, et al., Handling missing mri sequences in deep learning segmentation of brain metastases: a multicenter study, NPJ digital medicine 4 (1) (2021) 1–7.
- [23] G. M. Conte, A. D. Weston, D. C. Vogelsang, K. A. Philbrick, J. C. Cai, M. Barbera, F. Sanvito, D. H. Lachance, R. B. Jenkins, W. O. Tobin, et al., Generative adversarial networks to synthesize missing t1 and flair mri sequences for use in a multisequence brain tumor segmentation model, Radiology 299 (2) (2021) 313–323.
- [24] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784.
- [25] B. Yu, L. Zhou, L. Wang, J. Fripp, P. Bourgeat, 3d cgan based crossmodality mr image synthesis for brain tumor segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 626–630.
- [26] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.

- [27] H. Emami, Q. Liu, M. Dong, Frea-unet: Frequency-aware u-net for modality transfer, arXiv preprint arXiv:2012.15397.
- [28] X. Han, Mr-based synthetic ct generation using a deep convolutional neural network method, Medical physics 44 (4) (2017) 1408–1419.
- [29] T. Zhou, S. Ruan, Y. Guo, S. Canu, A multi-modality fusion network based on attention mechanism for brain tumor segmentation, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, 2020, pp. 377–380.
- [30] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (brats), IEEE transactions on medical imaging 34 (10) (2014) 1993–2024.