



# OPEN Laplacian filter attention with style transfer GAN for brain tumor MRI imputation

Yoonho Na<sup>1</sup>, Kyuri Kim<sup>1</sup>, Hyungjoo Cho<sup>2</sup>, Sung-Joon Ye<sup>1,3</sup>, Hwiyoung Kim<sup>4</sup>, Sung Soo Ahn<sup>5</sup>, Ji Eun Park<sup>6</sup> & Jimin Lee<sup>7,8</sup>✉

Training deep neural networks with multi-domain data generally gives more robustness and accuracy than training with single domain data, leading to the development of many deep learning-based algorithms using multi-domain data. However, if part of the input data is unavailable due to missing or corrupted data, a significant bias can occur, a problem that may be relatively more critical in medical applications where patients may be negatively affected. In this study, we propose the Laplacian filter attention with style transfer generative adversarial network (LASTGAN) to solve the problem of missing sequences in brain tumor magnetic resonance imaging (MRI). Our method combines image imputation and image-to-image translation to accurately synthesize specific sequences of missing MR images. LASTGAN can accurately synthesize both overall anatomical structures and tumor regions of the brain in MR images by employing a novel attention module that utilizes a Laplacian filter. Additionally, among the other sub-networks, the generator injects a style vector of the missing domain that is subsequently inferred by the style encoder, while the style mapper assists the generator in synthesizing domain-specific images. We show that the proposed model, LASTGAN, synthesizes high quality MR images with respect to other existing GAN-based methods. Furthermore, we validate the use of LASTGAN for data imputation or augmentation through segmentation experiments.

**Keywords** Generative adversarial networks, Data imputation, Image-to-image translation, MRI, Multi-domain

Many deep learning-based methods are rapidly developing in various fields, among which the medical field is no exception. Research using medical data often encounters difficulties, though, in the collection of data essential for training deep neural networks (DNNs). These difficulties are threefold: (1) a lack of rare disease data<sup>1,2</sup>, (2) patients' privacy concerns<sup>3</sup>, and (3) a complicated approval process for obtaining medical data due to legal issues<sup>4</sup>. In addition, some tasks arise when multi-modal data is needed to boost model performance. The fusion of many data from different domains can provide supplementary information and assist the network in producing more accurate predictions<sup>5</sup>. For example, several studies using various contrasts of magnetic resonance imaging (MRI) sequences for the segmentation of target lesions have been reported<sup>6–8</sup>. However, in the case when a multi-modal dataset is needed, it becomes more difficult to obtain a complete dataset of medical data in multiple domains considering the above three issues. Furthermore, even if the data can be obtained, there is still the possibility it contains defective data such as data loss in a specific domain, data corruption, or bad quality due to scanning artifacts. These defects in training data can cause a large bias when training DNNs and lead to a decline in model performance<sup>9</sup>. Consequently, when training DNNs that require multiple pairs of data, normal data may not be

<sup>1</sup>Department of Applied Bioengineering, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, Republic of Korea. <sup>2</sup>Department of Transdisciplinary Studies, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, Republic of Korea. <sup>3</sup>Research Institute of Convergence Science, Seoul National University, Seoul 08826, Republic of Korea. <sup>4</sup>Department of Biomedical Systems Informatics and Center for Clinical Imaging Data Science, Yonsei University College of Medicine, Seoul 03722, Republic of Korea. <sup>5</sup>Department of Radiology and the Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul 03722, Republic of Korea. <sup>6</sup>Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, Republic of Korea. <sup>7</sup>Department of Nuclear Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea. <sup>8</sup>Graduate School of Artificial Intelligence, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea. ✉email: jimilee@unist.ac.kr

included in the training data due to the presence of defective data. In order to effectively utilize all data, defective data must be replaced with other data such as synthesized samples from generative models.

Several approaches have been reported to deal with defective data using generative adversarial networks (GANs) to synthesize realistic medical image data<sup>10–12</sup>. Conte et al.<sup>13</sup> synthesized T1-weighted and fluid-attenuated inversion recovery (FLAIR) MR images for a brain lesion segmentation model that requires multiple MR sequences. Dai et al.<sup>14</sup> proposed a unified GAN that utilizes a single generator to synthesize T1-weighted, contrast-enhanced T1 (T1c), T2, and FLAIR MR images from single MR sequences. Xin et al.<sup>15</sup> synthesized T1, T1c, and FLAIR images from T2 contrast images. Zhang et al.<sup>16</sup> conducted multi-contrast MRI image synthesis using switchable CycleGAN. Jiang et al.<sup>17</sup> proposed RAGAN that uses reconstruction consistency loss to complete the missing T1, T1c, and FLAIR data from the given T2 modal data in real brain MRI. However, 1-to-1 or 1-to-n translation methods like the above generally have limitations in that they cannot fully reflect domain-specific representations, such as the visualization of specific organs or lesions in certain sequences.

Other image-to-image translation methods such as n-to-1 or n-to-n translation using multiple input images have been employed to alleviate this problem. Lee et al.<sup>18</sup> proposed CollaGAN, which can synthesize target contrast brain MR images utilizing complementary images. Joyce et al.<sup>19</sup> presented multiple-input encoder model and spatial transformer module to correct misalignment in the multiple-input data. Shen et al.<sup>5</sup> proposed ReMIC, an n-to-n image translation framework, that can synthesize all missing domain data from a set of random domains by disentangling images into shared content code and domain-specific style code. Hi-Net, developed by Zhou et al.<sup>20</sup>, encodes two MR sequences and combines latent representations of two modalities to synthesize the remaining MR image sequence. Dalmaz et al.<sup>21</sup> proposed a transformer-based generator to translate between multiple sequences of MR images. Using multiple inputs like above generally produces better image quality than 1-to-n translation methods. Zhang et al.<sup>22</sup> proposed a commonality- and discrepancy-sensitive encoder that can exploit both modality-invariant and specific information contained in input modalities for addressing missing modality imputation. Wang et al.<sup>23</sup> introduced SIFormer, a framework that simultaneously performs super-resolution of low-quality MR images and imputes missing sequences using available low-resolution contrast configurations in a single forward pass. Cho et al.<sup>24</sup> presented HF-GAN, which synthesizes multisequence MR images by disentangling complementary and modality-specific information. Recent advancements, such as transformer-based or diffusion model architectures, have further improved image translation quality<sup>25–28</sup>. Despite such improvements, though, relatively less attention has been paid to synthesizing specific areas in brain MR images such as tumor regions.

In this study, we propose a GAN-based image imputation framework called LASTGAN, or Laplacian filter attention with style transfer generative adversarial network. The proposed model synthesizes missing MR image sequences from other available MR sequences in high quality. Specifically, we used three MR sequences out of T1, gadolinium enhanced T1 (T1Gd), T2, and FLAIR sequences to synthesize the remaining MR sequence. In order to synthesize high quality brain tumor MR images while alleviating the limitations of previous studies, we applied a new technique to our proposed model. LASTGAN includes a novel attention mechanism in the generator called Laplacian filter attention that can assist the synthesis of domain-specific tumor representations while preserving the other brain structures through the use of a Laplacian filter, which is a type of edge detector. To further support the network in synthesizing specific MR image sequences, we also adopted a style transfer method.

Our contributions are summarized as follows.

- We propose a GAN-based image imputation framework, LASTGAN, for imputing missing brain tumor MRI sequences.
- We introduce a novel attention mechanism called Laplacian filter attention for the accurate synthesis of brain tumor regions while preserving the anatomical structures of the brain.
- We illustrate the effectiveness of the Laplacian filter attention and style transfer methods that assist the network to synthesize high quality brain tumor MR images.
- We show that the samples synthesized by LASTGAN are feasible for data augmentation and imputation by validation through a brain tumor segmentation experiment.

## Methods

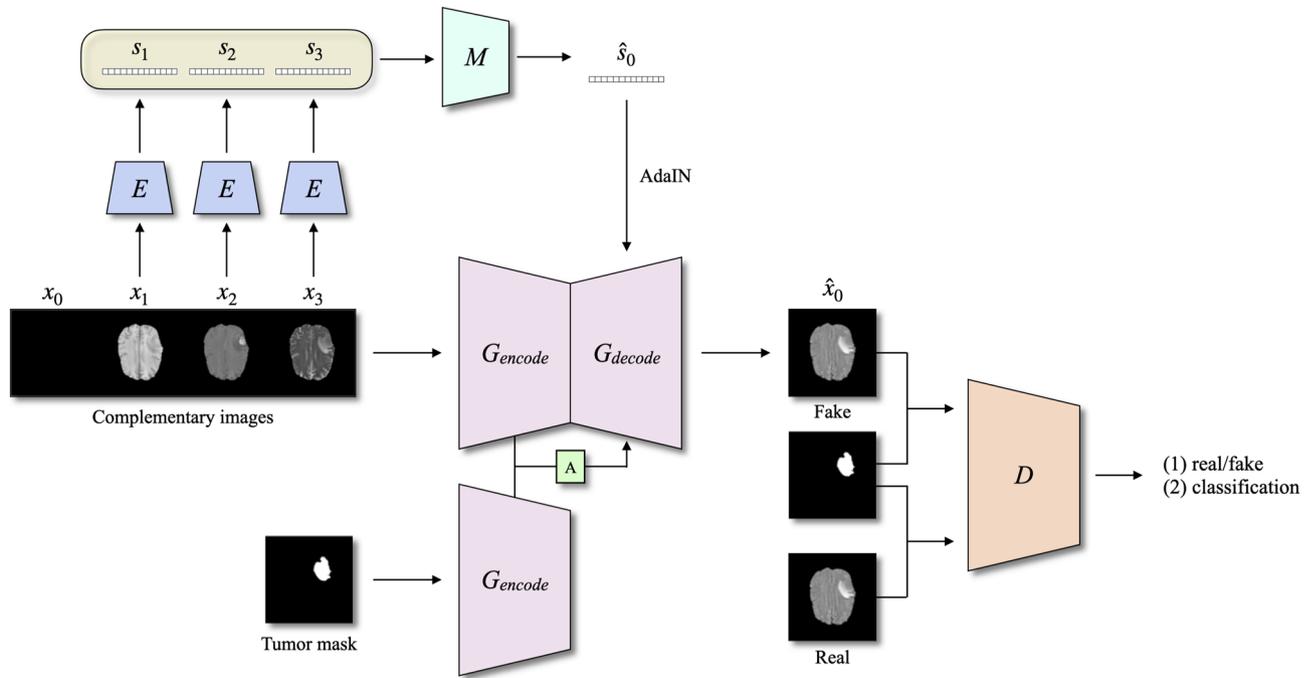
### Network architecture

Figure 1 shows the proposed framework of LASTGAN. The purpose of LASTGAN is to synthesize missing MRI sequences using supplementary information from each domain. To do so, the network comprises four sub-networks: generator  $G$ , discriminator  $D$ , style encoder  $E$ , and style mapper  $M$ . These sub-networks are explained in the following sections.

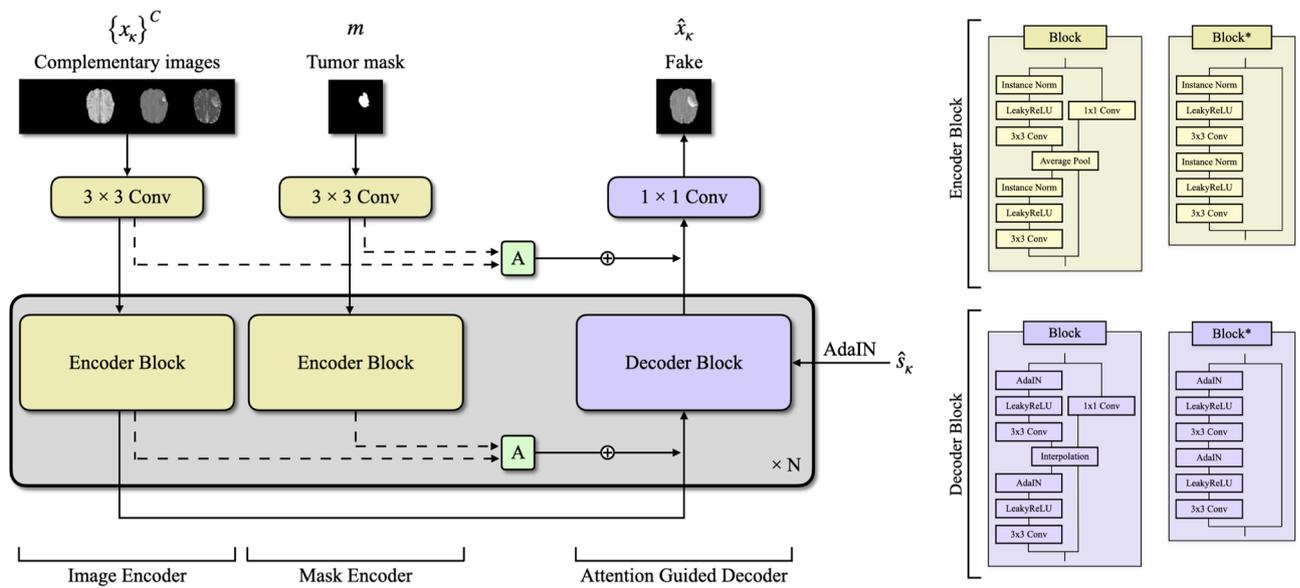
#### Generator

As shown in Fig. 2, our generator  $G$  contains two separate encoders, an image encoder and a mask encoder. The image and mask encoders both extract features from the available MRI sequences  $\{x_\kappa\}^C$  and a tumor mask image  $m$ . From each layer of the encoder block, the extracted features are fed into the Laplacian filter attention module. The general purpose of the Laplacian filter attention module is to preserve the domain-shared data, such as the anatomical brain structures, and also to assist the synthesis of domain-specific data, such as a brain tumor. The output of the attention module is an attention map, which is element-wise added into each decoder layer. Fig. 3 shows a detailed illustration of the Laplacian filter attention module. Following attention map injection, the predicted style vector from style mapper  $M$  is injected to the decoder via AdaIN.

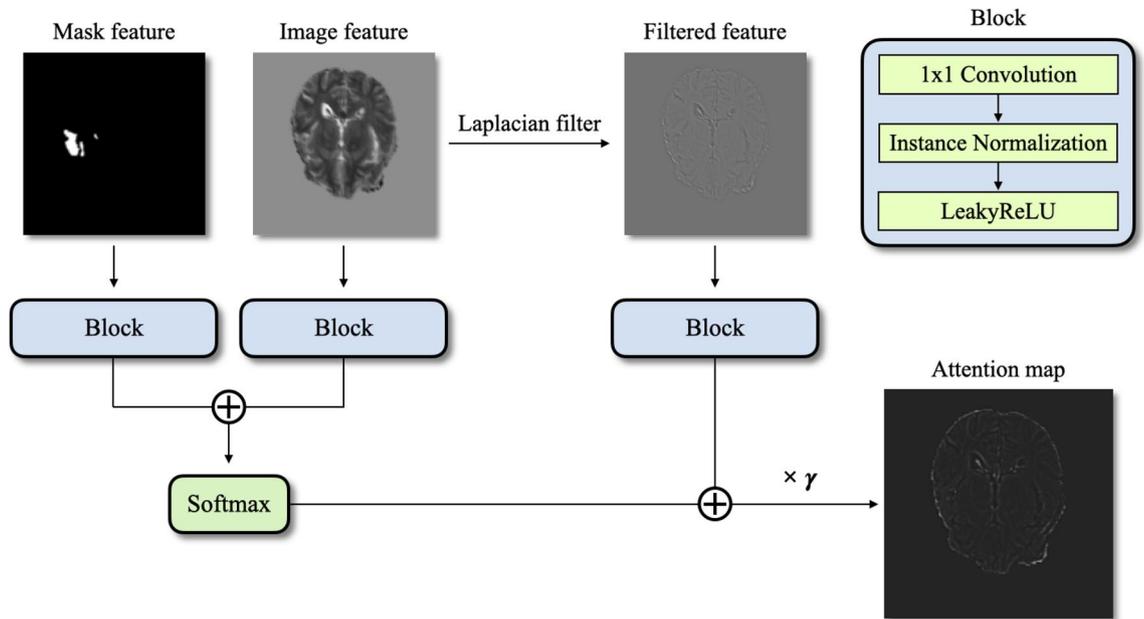
As our generator receives inputs of available MRI sequences  $\{x_\kappa\}^C$ , tumor mask  $m$ , and predicted target domain style vector  $\hat{s}_\kappa$ , the synthesized missing sequence  $\hat{x}_\kappa$  from the generator can be defined as Eq. (1),



**Fig. 1.** Overview of the proposed network, LASTGAN, consisting of four sub-networks: generator  $G$ , discriminator  $D$ , style encoder  $E$ , and style mapper  $M$ . The feature maps from the layers of each generator encoder are extracted to obtain Laplacian attention maps, which are later fed into the generator decoder. The style vector is inferred in a two-stage process: (1) inference of complementary style vectors through  $E$ , and (2) inference of the missing style vector through the complementary style vectors.



**Fig. 2.** Architecture of the generator  $G$  of LASTGAN. “A” indicates Laplacian filter attention operations.



**Fig. 3.** Laplacian filter attention module.  $\oplus$  denotes element-wise summation, and  $\gamma$  is a trainable weighting factor.

$$\hat{x}_\kappa = G(\{x_\kappa\}^C, m, \hat{s}_\kappa; \kappa), \tag{1}$$

where  $\{x_\kappa\}^C$  is all sequence data excluding the  $\kappa$  domain sequence. The circumflex is used to distinguish the data generated by the generator  $\hat{x}_\kappa$  from the original target domain data  $x_\kappa$ , and likewise the style vectors generated by the style encoder  $s_\kappa$  and by the style mapper  $\hat{s}_\kappa$ , as described next.

*Style encoder and style mapper*

By the Multiple Cycle Consistency(MCC) loss, proposed by Lee et al.<sup>18</sup>, the generator  $G$  itself learns the missing domain and maps the available domain data to the missing domain data. However, when multiple inputs, especially with zero values in random channels, are fed into the generator, the multiple-domain input data can be mixed up, resulting in data loss inside the network. To solve this problem, the style encoder  $E$  is adopted, which introduces domain-specific style vectors from single domain data. In this case, we need to inject the target domain’s style vector, for which the style mapper  $M$  is designed. The style mapper receives inputs of style vectors from available MR images and maps to the missing domain style vector. As well as in the image, the style vectors from other domains have distinct distributions, and thus the missing target style vector can be easily obtained by solving the regression problem. The predicted missing style vector is later injected into the generator decoder through AdaIN layers.

The generated style vector from style encoder  $E$  and mapper  $M$  can be written as Eq. (2),

$$s_\kappa = E(x_\kappa), \hat{s}_\kappa = M(\{s_\kappa\}^C; \kappa). \tag{2}$$

*Discriminator*

Our model’s discriminator  $D$  receives inputs of real data  $x$  from the training dataset or synthetic data  $\hat{x}$  generated by the generator  $G$ . Discriminating between real and fake data influences the generator to synthesize more realistic data that will eventually fool the discriminator and render it unable to discriminate the data. Since our model works with multi-modal inputs, the discriminator also classifies the input domain to synthesize domain-specific images. In addition, the corresponding tumor mask image is concatenated before feeding the data into the discriminator to make the model focus on the tumor region when discriminating real and fake images.

**Objective functions**

The total objective function can be separated into three sub-objectives, namely for the image domain, for the style vector domain, and for adversarial loss. The notations used in this section are explained below for better understanding.

Suppose there are four domain data denoted by  $\{x_0, x_1, x_2, x_3\}$ , and a single domain data  $x_\kappa$  is randomly missing. If the missing target domain index is 1, the available domain data can be written as  $\{x_1\}^C = \{x_0, x_2, x_3\}$ .

### Image objective function

The MCC was proposed with the main purpose to implement the cycle loss in CycleGAN<sup>29</sup> for multiple inputs. The role of cycle consistency is to encourage strong mapping between forward and reverse processes, leading to a preservation of the structural information in the image. MCC in this case can be written as follows:

$$L_{mcc-imag,\kappa} = \sum_{\kappa' \neq \kappa} \|x_{\kappa'} - \tilde{x}_{\kappa'|\kappa}\|_1 \quad (3)$$

where

$$\tilde{x}_{\kappa'|\kappa} = G(\{\hat{x}_{\kappa}\}^C, m, \tilde{s}_{\kappa'}; \kappa'). \quad (4)$$

The tilde is to indicate that the synthesized data  $\tilde{x}_{\kappa}$  is made from previously synthesized data  $\hat{x}_{\kappa}$ . For example, if we synthesized  $\hat{x}_0$  from  $\{x_1, x_2, x_3\}$ , where  $\kappa = 0$  in this case, the secondly synthesized data  $\tilde{x}_{1|0}$ ,  $\tilde{x}_{2|0}$ , and  $\tilde{x}_{3|0}$  can be written as follows:

$$\tilde{x}_{1|0} = G(\{\hat{x}_0, x_2, x_3\}, m, \tilde{s}_1; 1), \quad (5)$$

$$\tilde{x}_{2|0} = G(\{\hat{x}_0, x_1, x_3\}, m, \tilde{s}_2; 2), \quad (6)$$

$$\tilde{x}_{3|0} = G(\{\hat{x}_0, x_1, x_2\}, m, \tilde{s}_3; 3). \quad (7)$$

Here, the structural similarity index measure (SSIM) loss is applied in the same way that  $L1$  loss is calculated in MCC to prevent blurry artifacts and improve model performance<sup>30</sup>,

$$L_{mcc-SSIM,\kappa} = \sum_{\kappa' \neq \kappa} L_{SSIM}(x_{\kappa'}, \tilde{x}_{\kappa'|\kappa}), \quad (8)$$

where

$$L_{SSIM}(x, y) = 1 - \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

and  $c_1 = (0.01 \cdot L)^2$ ,  $c_2 = (0.03 \cdot L)^2$ , and  $L$  is a dynamic range of pixel values.

### Style vector objective function

One of the main ideas of our proposed method is to utilize a style transfer technique while considering multiple inputs. To do so, the generated style vector needs strong mappings among multiple style vector domains. The same MCC was applied to the style vectors to encourage consistency across the multiple domains, as

$$L_{mcc-sty,\kappa} = \sum_{\kappa' \neq \kappa} \|s_{\kappa'} - \tilde{s}_{\kappa'|\kappa}\|_1 \quad (10)$$

where

$$\tilde{s}_{\kappa'|\kappa} = E(\{\hat{x}_{\kappa}\}^C; \kappa'). \quad (11)$$

$L1$  loss is also used to force the style mapper sub-network to accurately predict the missing target style vector from the style vectors derived from the style encoder:

$$L_{sty} = \mathbb{E}_{x \sim p(x)} [\|E(x_{\kappa}) - M(E(\{\hat{x}_{\kappa}\}^C))\|_1]. \quad (12)$$

### Adversarial loss

Training the generator  $G$  and the discriminator  $D$  adversarially induces the generator to synthesize realistic data that fits the data distribution of the training data, as below:

$$\min_G \max_D \mathbb{E}_{x_{\kappa} \sim p(x)} [\log D(x_{\kappa}, m)] + \mathbb{E}_{x_{\kappa}^C \sim p(x)} [1 - \log D(G(x_{\kappa}^C, m, \hat{s}; \kappa), m)]. \quad (13)$$

### Full objective

The full objective of the model can be written as a linear combination of the above objective functions:

$$L_{total} = \lambda_{adv} L_{adv} + \lambda_{mcc-imag} L_{mcc-imag} + \lambda_{mcc-SSIM} L_{mcc-SSIM} + \lambda_{SSIM} L_{SSIM} + \lambda_{img} L_{img} + \lambda_{mcc-sty} L_{mcc-sty} + \lambda_{sty} L_{sty}, \quad (14)$$

where  $\lambda_{adv}$ ,  $\lambda_{mcc-imag}$ ,  $\lambda_{mcc-SSIM}$ ,  $\lambda_{SSIM}$ ,  $\lambda_{img}$ ,  $\lambda_{mcc-sty}$ , and  $\lambda_{sty}$  are the weights for each corresponding objective.

## Experiments

### Datasets

The multi-modal brain tumor image segmentation (BraTS) benchmark<sup>31–33</sup> 2016 and 2017 datasets consist of 750 subjects in four MRI modalities: fluid-attenuated inversion recovery (FLAIR), T1-weighted (T1), Gd-enhanced T1-weighted (T1Gd), and T2-weighted (T2). The original dataset was collected from 19 different institutions, employing diverse equipment and imaging protocols. The detailed imaging parameters of MRI scans were not provided. The dataset includes two types of brain tumors: glioblastoma (GBM) and lower-grade glioma (LGG). GBMs are highly malignant, fast-growing tumors characterized by strong contrast enhancement on T1-weighted images, as well as regions of necrosis and extensive edema. In contrast, LGGs grow more slowly than GBMs and typically exhibit minimal enhancement on T2/FLAIR images. Tumors in the dataset are segmented into three sub-regions: enhancing tumor, non-enhancing tumor, and edema. The enhancing tumor is identified by hyperintense areas on T1Gd images, while the non-enhancing tumor appears as hypointense regions on T1Gd images. Edema is readily identifiable on T2-weighted images as areas of hyperintense abnormal signal distribution, which appear hypointense on T1-weighted images. Since these are open datasets, Institutional Review Board (IRB) approval was not required.

### Data pre-processing

In this study, the preprocessing was designed to enhance the model's ability to learn tumor-specific representations. Only MR images containing tumors were selected, ensuring the model concentrated on relevant features. The MR images were normalized to a range of  $[-1, 1]$  to match the intensity values, and tumor masks were binarized to explicitly indicate tumor locations. Each tumor slice, along with its corresponding mask, was resized from 240 240 to 256 256 pixels, and N4 bias field correction was applied to reduce intensity inhomogeneities. Our proposed method utilizes multiple MR sequences as input while excluding one sequence. To improve computational efficiency during training, all possible combinations of inputs were prepared in advance. Specifically, four MR sequences at the same slice level were concatenated in channel dimension, and one sequence was replaced with zero values to simulate exclusion. This approach generated four combinations per slice level, effectively serving as a form of data augmentation. In total, 96,770 samples were generated, with 77,413 used for training and 19,357 reserved for testing.

### Comparative methods and evaluation metrics

In this study, StarGAN v2, CollaGAN, and ResViT were utilized for comparison. StarGAN v2 is a multi-domain image-to-image translation framework capable of both latent-guided and reference-guided synthesis. For reference-guided synthesis, it employs a style encoder to extract domain-specific style codes based on a reference image, enabling diverse image generation across multiple domains. While effective in generating multi-modal images, its 1-to-n translation approach presents the limitation that it is not able to learn the correlations between existing domain. CollaGAN is a framework designed for missing image data imputation by utilizing complementary input images to synthesize the missing modality. It employs a single generator and discriminator, which simplifies the architecture compared to traditional multi-generator models. To ensure bidirectional mappings and preserve structural information, CollaGAN utilizes multiple cycle consistency (MCC) loss. In addition, mask vectors guide the synthesis process by focusing on specific modalities. ResViT combines convolutional neural networks (CNNs) and transformers to tackle multi-modal medical image imputation. Using aggregated residual transformer (ART) blocks, ResViT achieves a balance between global contextual sensitivity and local precision, enabling it to synthesize missing MRI sequences with high realism and structural consistency. However, its reliance on transformer components makes it computationally intensive and data-hungry.

For comparison, StarGAN v2 and CollaGANStarGAN v2, StarGAN v2, CollaGAN, and ResViT were trained on the same BraTS 2016 and 2017 datasets. Although StarGAN v2 can perform two types of image synthesis, latent-guided and reference-guided, only the results of reference-guided synthesis are shown in the present results because this type consistently showed better performance. As evaluation metrics, the structural similarity index measure (SSIM), normalized root mean squared error (NRMSE), peak signal to noise ratio (PSNR), and learned perceptual image patch similarity (LPIPS) were utilized. Also, to validate that our model better reflects the characteristics of each MR sequence in tumor regions compared to the two other models, NRMSE and PSNR were additionally calculated for the tumor regions only. The evaluation metrics are formulated as follows:

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (15)$$

where  $c_1 = (0.01 \cdot L)^2$ ,  $c_2 = (0.03 \cdot L)^2$ , and  $L$  denotes a dynamic range of pixel values,

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2, \quad (16)$$

$$NRMSE = \frac{1}{y_{max} - y_{min}} \sqrt{MSE}, \quad (17)$$

$$PSNR = 20 \cdot \log_{10} \left( \frac{MAX_I}{\sqrt{MSE}} \right), \quad (18)$$

where  $MAX_I$  is the maximum intensity of the image.

$$LPIPS = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w^l \odot (\phi_{h,w}^l(x) - \phi_{h,w}^l(y))\|_2^2 \quad (19)$$

where  $\phi^l$  is the  $l$ th layer activations of pretrained convolutional neural network and  $w$  is the channel-wise scaling factor.

Three features differentiate LASTGAN from other methods. First, our model utilizes the predicted style vector derived from the existing modality for a style transfer method to synthesize data that follows the specific sequence of the target data. Second, the Laplacian filter attention module preserves the brain and tumor structures by feeding the calculated attention map into the generator decoder. Third, the Laplacian filter, which is known as an edge detector, produces more edge information when processing the attention map. Therefore, to see the effectiveness of each aspect of our model, LASTGAN was evaluated with the metrics above in three distinct ablated setups, as follows.

- LASTGAN w/without style transfer (w/o style)
- LASTGAN w/without Laplacian filter attention (w/o attn)
- LASTGAN w/without Laplacian filter (w/o Lap)

Note that the AdaIN layers were substituted with instance normalization layers for LASTGAN w/o style transfer. Also, for LASTGAN w/o Laplacian filter, other operations in the attention module were still utilized except for the Laplacian filter.

#### Experiments on data augmentation and imputation

To show that the synthesized samples from our model are feasible for data augmentation, a brain tumor segmentation model was additionally trained with a U-net<sup>34</sup> backbone using only T2 sequences and corresponding binary tumor masks from the BraTS 2016 and 2017 datasets. This experiment was designed to simulate real-world scenarios in which certain data must be excluded due to missing or corrupted portions, rather than comparing against a baseline that includes complete datasets. In order to investigate performance changes according to the amount of data augmentation, the segmentation model was trained starting from the baseline, which we set as training without data augmentation up to 50% data augmentation, meaning that the amount of augmented data is half of the amount of total training data in the baseline. The augmented data were randomly picked from the samples synthesized by LASTGAN. Note that the samples picked were not synthesized using the data from the training dataset. For the loss function, a combination of binary cross entropy loss and DICE loss was utilized. Training was conducted for 20 epochs with a batch size of 128. Details of the prepared datasets are shown in Table 1.

For the missing data imputation experiment, all modalities (FLAIR, T1, T1Gd, and T2) in the BraTS 2016 and 2017 dataset of MR images were channel-wise concatenated, and the missing sequence was simulated with zero values. The simulated conditions are as follows: (1) a complete set of sequences is given (baseline), (2) one random MRI sequence is missing, (3) the random missing MRI sequence is imputed with a synthesized sample.

To see the tendency of performance degradation due to the amount of missing or imputed data, the probability  $p$  of missing and imputed data was set to 50% and 100%; the former refers to missing/imputed data with a probability of 50% while the other 50% remains in the complete dataset, and the latter refers to missing/imputed data with 100% probability. After training, the models were evaluated with four metrics: DICE, IoU (intersection over union), precision and recall. The evaluation metrics are formulated as follows:

$$DICE = \frac{2TP}{2TP + FP + FN}, \quad (20)$$

$$IoU = \frac{TP}{TP + FP + FN}, \quad (21)$$

$$Precision = \frac{TP}{TP + FP}, \quad (22)$$

$$Recall = \frac{TP}{TP + FN}, \quad (23)$$

where  $TP$ ,  $FP$ , and  $FN$  respectively refer to true positive, false positive, and false negative.

Method	Train	Validation	Test
Baseline	5,805	580	648
25% Augmentation	5,805+1,451	580	648
50% Augmentation	5,805+2,902	580	648

**Table 1.** Datasets for data augmentation experiments.

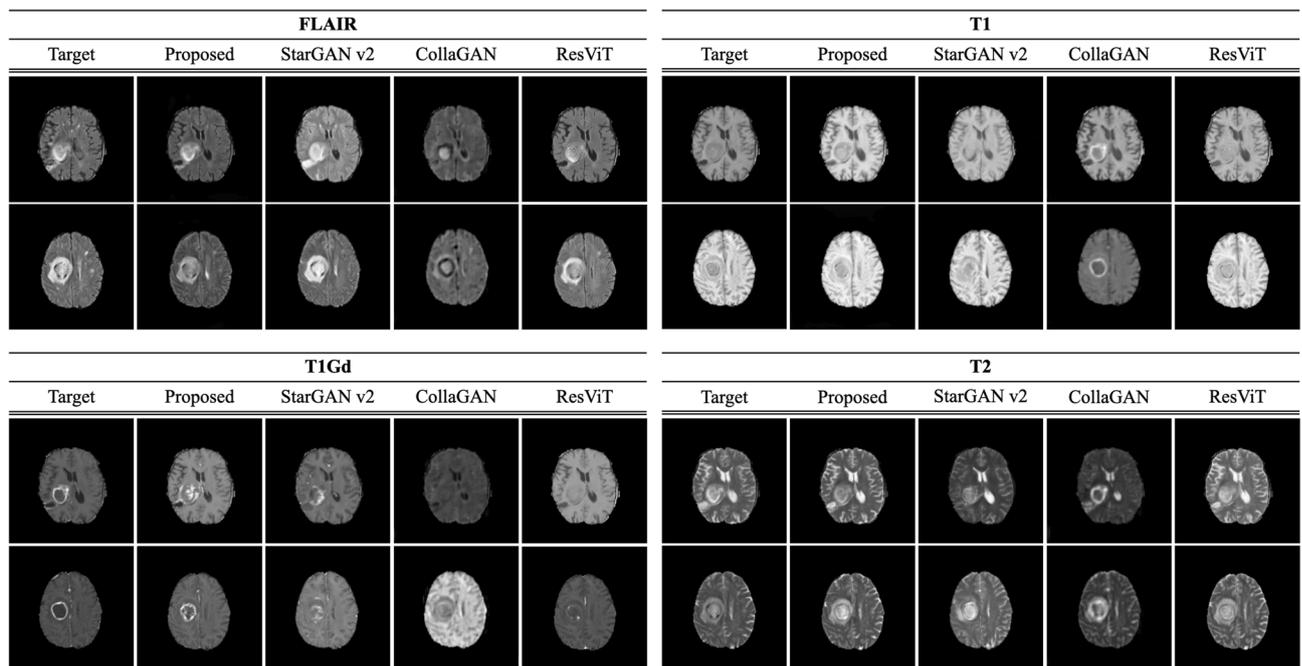
## Results

Figure 4 shows the synthesized samples from LASTGAN, StarGAN v2, CollaGAN, and ResViT. Overall, LASTGAN was able to synthesize more target-like samples, whereas StarGAN v2, CollaGAN, and ResViT results showed less similarity to the target images. In particular, LASTGAN showed better performance in appropriately reflecting the domain characteristics compared to the other methods.

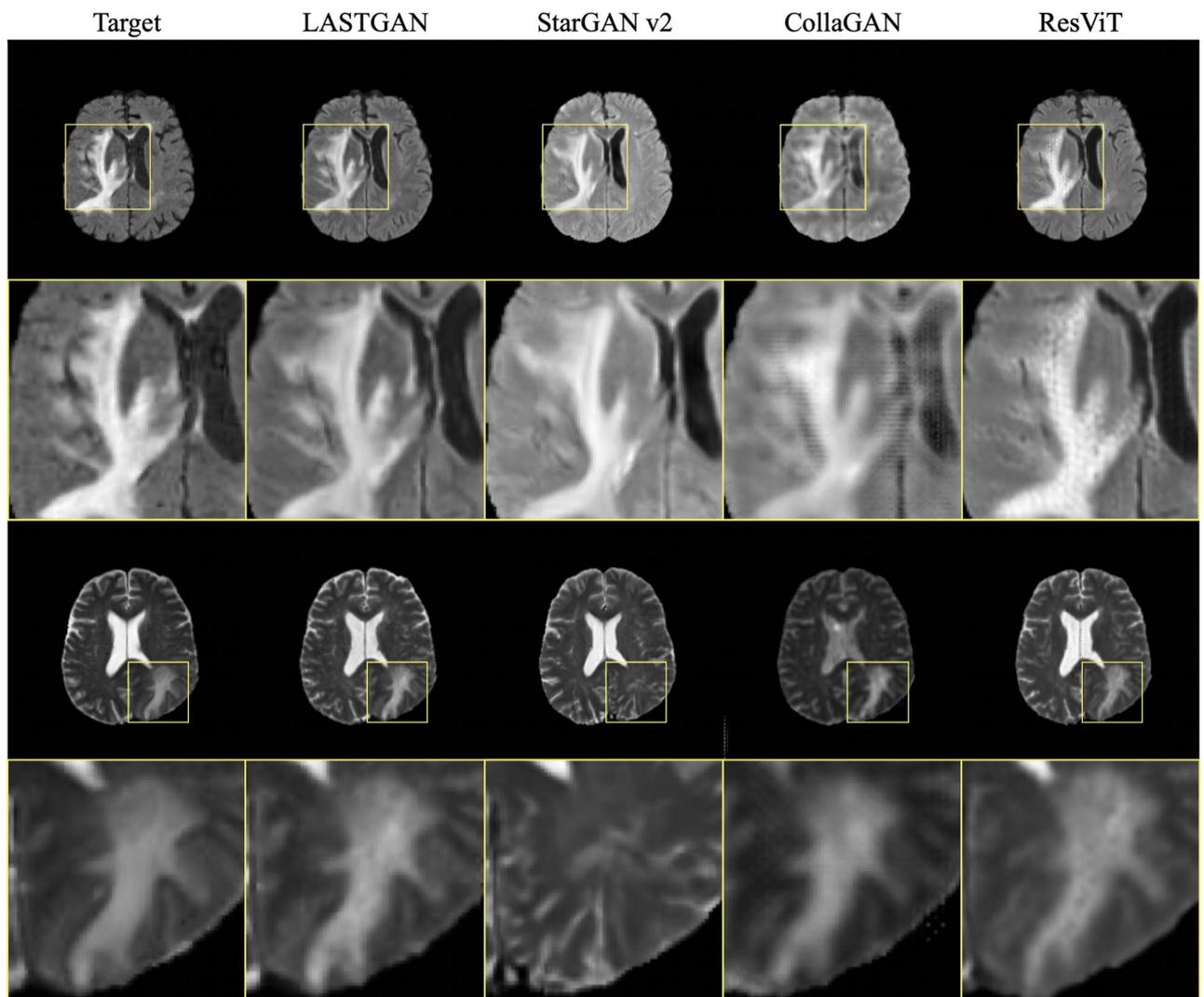
Looking at the results of the other methods, many checkerboard artifacts were frequently observed from the samples of CollaGAN, as seen in Fig. 5. We assume that the mask vectors that are fed with the complementary images degrade the ability to capture the domain features and to generate pixel data, because while the role of the mask vectors in this method is only to indicate the target location, they occupy half of the input portion. These artifacts likely stem from the use of mask vectors fed alongside complementary images, which may degrade the model's ability to effectively capture domain-specific features and generate accurate pixel data. While the mask vectors are intended to indicate target locations, they occupy half of the input space, potentially limiting their utility and degrading synthesis quality. Artifacts were also observed in the samples produced by ResViT; however, there was a notable difference between the two methods. In CollaGAN, the artifacts appeared across the entire image, whereas in ResViT, they were more concentrated in the tumor region. With respect to StarGAN v2, samples in this case showed limitations in synthesizing the domain-specific tumor regions due to its 1-to-n translation method. Further, as shown in Fig. 5, if the tumor region had low visibility in the source image, then the translated image also showed low visibility.

The results of the quantitative evaluation for StarGAN v2, CollaGAN, ResViT, and LASTGAN are given in Table 2. As can be seen in the table, LASTGAN showed better results of SSIM, NRMSE, and PSNR than the other methods. LASTGAN consistently outperformed StarGAN v2 and CollaGAN across all quantitative metrics in all four MRI sequences. Additionally, LASTGAN demonstrated performance that was comparable to ResViT. Table 2 also gives the results of the ablated setups in the bottom three rows to validate our proposed method. The results without Laplacian filter attention showed a 14% decrease in the SSIM metric on average, which indicates that the Laplacian attention module is essential for our proposed method. For the results excluding the AdaIN-based style transfer method, the SSIM result showed about a 4% decrease on average. LASTGAN also demonstrated outstanding results in the tumor region evaluation, as shown in Table 3.

To check the feasibility of data augmentation and imputation with synthetic samples, segmentation experiments were conducted (see Section IV-C). The results of DICE, IoU, precision, and recall metrics for brain tumor segmentation considering data augmentation are listed in Table 4. The results for all methods show a tendency of increasing DICE score up to 50% data augmentation, implying that all models are feasible for data augmentation. However, the segmentation performance significantly increased when utilizing samples from LASTGAN. For the data imputation experiment, the results are shown in Table 5. The DICE score of the baseline was 0.6594, whereas that with one missing sequence was 0.5679. This indicates that when data is missing in a random sequence, the model finds it harder to learn the generalized output. The DICE score of data imputation with the synthesized samples remained between the baseline and the one-missing-sequence models. When the data was imputed at a probability of 50%, we observed increments in both DICE and IoU (Table 5). This result demonstrates that the sample images from StarGAN v2, CollaGAN, ResViT, and LASTGAN cannot fully



**Fig. 4.** Synthesized samples of LASTGAN (proposed), StarGAN v2, CollaGAN, and ResViT. The target images are shown in the left columns for comparison. Each rows are in the same slice level within the same patient..



**Fig. 5.** Enlarged images of the synthesized samples of LASTGAN, StarGAN v2, CollaGAN, and ResViT. The target images are shown in the left column for comparison.

Method	FLAIR			T1			T1Gd			T2			Average		
	SSIM	NRMSE	LPIPS												
StarGAN v2	0.83261	0.02098	0.10374	0.90552	0.02112	0.10757	0.84361	0.02128	0.10886	0.86368	0.02062	0.11903	0.86136	0.02100	0.1098
CollaGAN	0.81037	0.02102	0.1421	0.84033	0.02098	0.09672	0.83646	0.02076	0.14778	0.83993	0.02017	0.11824	0.83177	0.02073	0.12621
ResViT	<b>0.87936</b>	<b>0.01947</b>	0.10031	0.91342	0.02096	0.07805	0.90316	0.02324	0.09474	<b>0.91769</b>	<b>0.01694</b>	0.07629	0.90341	0.02025	0.08732
LASTGAN	0.87911	0.02008	<b>0.08893</b>	<b>0.91359</b>	<b>0.01965</b>	<b>0.06547</b>	<b>0.90454</b>	<b>0.01974</b>	<b>0.08471</b>	0.91718	0.01905	<b>0.06942</b>	<b>0.90361</b>	<b>0.01963</b>	<b>0.07696</b>
w/o style	0.83810	0.02029	0.09487	0.88600	0.02009	0.08357	0.86044	0.02024	0.08846	0.88139	0.01926	0.0728	0.86648	0.01997	0.08463
w/o attn	0.75874	0.02137	0.12992	0.85622	0.02073	0.10902	0.83900	0.02058	0.11608	0.66265	0.02127	0.12704	0.77915	0.02099	0.11815
w/o Lap	0.86941	0.02037	0.10168	0.90791	0.01997	0.08452	0.89483	0.01994	0.09673	0.90616	0.01946	0.08184	0.89458	0.01993	0.08982

**Table 2.** Evaluation results of StarGAN v2, CollaGAN, ResViT, LASTGAN, and ablated setups. Best values in each column are shown in bold.

replace the original data but can still be imputed to replace defective data. Among these methods, samples from LASTGAN showed better performance than the other comparative models in all metrics.

### Discussion

In the present work, a method combining multi-domain data imputation and style transfer was proposed. But from the results of our ablated methods, we noticed that even without style transfer, the generator of LASTGAN

Method	FLAIR		T1		T1Gd		T2		Average	
	NRMSE	PSNR								
StarGAN v2	0.20232	64.26199	0.26587	62.49304	0.28715	63.32991	0.31422	62.41943	0.26739	63.12609
CollaGAN	0.21925	63.18053	0.19237	65.52283	0.25491	63.76232	0.20298	65.49029	0.21738	64.48899
ResViT	0.13627	<b>67.54454</b>	0.17446	66.87798	0.22130	64.82801	<b>0.11984</b>	69.85980	0.15896	66.92758
LASTGAN	<b>0.13342</b>	67.45185	0.15408	68.12270	<b>0.18662</b>	<b>66.36514</b>	0.12265	<b>70.15309</b>	<b>0.14919</b>	<b>68.02320</b>
w/o style	0.14514	67.43036	<b>0.12801</b>	<b>70.80804</b>	0.20901	64.71160	0.13089	68.26018	0.15364	67.81867
w/o attn	0.17428	65.21512	0.15555	68.39355	0.20337	65.10605	0.18438	65.46838	0.17964	66.03072
w/o Lap	0.13558	67.44829	0.13414	70.03314	0.18924	65.50801	0.14219	67.54472	0.14977	67.69447

**Table 3.** Tumor region evaluation results of StarGAN v2, CollaGAN, ResViT, LASTGAN, and ablated setups. Best values in each column are shown in bold.

Method	Dataset	DICE	IoU	Precision	Recall
<b>Baseline</b>	<b>No Augmentation</b>	<b>0.5691</b>	<b>0.5101</b>	<b>0.6707</b>	<b>0.6825</b>
StarGAN v2	25% Augmentation	0.6977	0.5398	0.6441	0.7669
	50% Augmentation	0.7034	0.5457	0.6355	0.7928
CollaGAN	25% Augmentation	0.6817	0.5211	0.6195	0.7616
	50% Augmentation	0.7063	0.5491	0.6767	0.7428
ResViT	25% Augmentation	0.6920	0.5465	0.6451	0.7673
	50% Augmentation	0.7256	0.5704	0.6882	0.7935
LASTGAN	25% Augmentation	0.7110	0.5549	0.6572	0.7839
	50% Augmentation	<b>0.7529</b>	<b>0.6086</b>	<b>0.6985</b>	<b>0.8265</b>

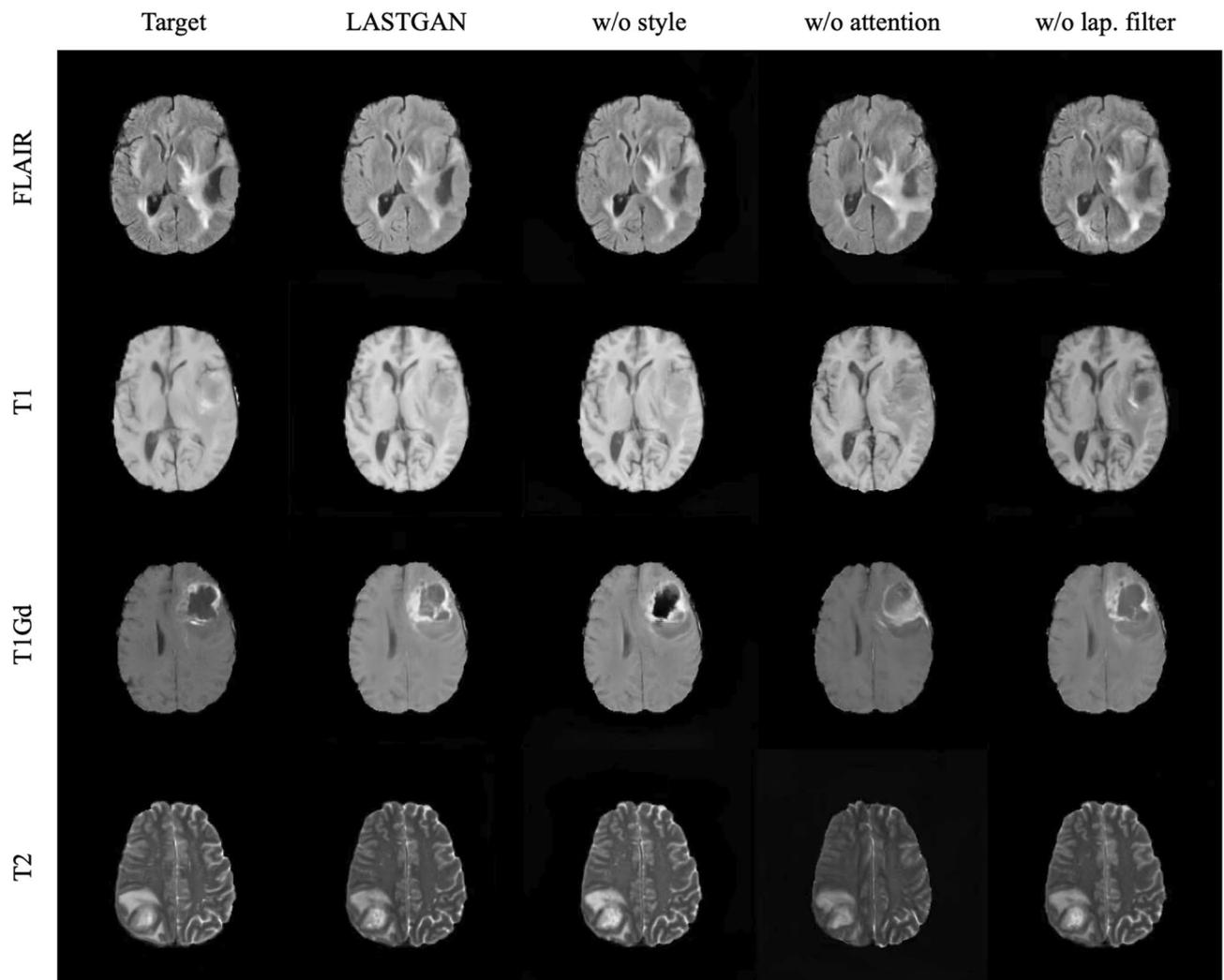
**Table 4.** Segmentation results by data augmentation amount. Best values in each column are shown in bold.

Method	Dataset	DICE	IoU	Precision	Recall
<b>Baseline</b>	<b>Complete dataset</b>	<b>0.6594</b>	<b>0.5615</b>	<b>0.7489</b>	<b>0.6747</b>
Missing	Missing 1 sequence (p = 0.5)	0.5679	0.4768	0.7130	0.5618
	Missing 1 sequence (p = 1.0)	0.4792	0.3990	0.6370	0.4965
StarGAN v2	Missing 1 sequence (p = 0.5)	0.5986	0.5036	0.7120	0.6211
	Missing 1 sequence (p = 1.0)	0.5753	0.4785	0.6885	0.6140
CollaGAN	Missing 1 sequence (p = 0.5)	0.5991	0.5028	0.7524	0.5977
	Missing 1 sequence (p = 1.0)	0.5864	0.4895	0.7671	0.5546
ResViT	Missing 1 sequence (p = 0.5)	0.6004	0.5058	0.7276	0.6021
	Missing 1 sequence (p = 1.0)	0.5987	0.5121	0.7354	0.6276
LASTGAN	Missing 1 sequence (p = 0.5)	0.6039	0.5093	0.7131	0.6212
	Missing 1 sequence (p = 1.0)	0.6141	0.5188	0.7342	0.6403

**Table 5.** Segmentation results by data imputation amount.

could still synthesize target-like images (Fig. 6). This may be because the number of missing domains was fixed to a single domain, and also, the fixed order of the domains in the input made it easier for the generator to learn the domain information. Thus, the generator could self-learn the output domain from the input images through MCC loss. Accordingly, we decided to force the style vector to contain more domain information than the generator could learn by itself.

The easiest way of doing this was to make the generator synthesize other domain images instead of fixing the output to a single missing domain image. In this setting, additional  $L1$  loss was added for each synthesized output and corresponding original MR image in the same domain. With this setting, we were able to obtain domain-shared information (e.g., anatomical structures of the brain that every domain shares) with no domain specificity. However, we did not choose the style-induced setting of LASTGAN in the final model selection because our network utilizes three multiple cyclic loss functions, the calculation of which already consumes a large amount of computational resources. Also, we observed with this model variant slight decreases in SSIM, NRMSE, and PSNR metrics (0.9018, 0.02013, and 25.19213, respectively) as well as inappropriately generated tumors, especially in synthesizing enhanced tumor regions. This outcome aligns with the tumor region evaluation results in Table 3, where the style transfer-ablated setup showed a bias toward the T1 sequence. Through this result, we conclude that disentangling the content and style information for synthesizing brain tumor MR images does



**Fig. 6.** Synthesized samples from the three ablated methods for comparison with LASTGAN.

not always improve the model, particularly when using multiple inputs. Through these experimental results, we draw several key conclusions regarding the style transfer approach. First, while style-content disentanglement is often beneficial in image translation tasks, we found that when using multiple inputs, such strict disentanglement does not consistently improve model performance. Second, the style transfer method helps to preserve domain-specific features, especially in tumor regions where maintaining precise characteristics is crucial for clinical utility. Third, the style vector serves as an explicit guide for domain translation, collaborating with the generator's learned features to produce more accurate results.

For the attention-ablated model, we observed a loss of structural edge information in the middle of the network, whereas our baseline model produced sharper images. Similar to the attention-ablated model, the Laplacian filter-ablated model showed slightly less structural similarity to the target images compared to the baseline. However, in the evaluation results, there were no significant differences between LASTGAN and the Laplacian filter-ablated model. The only fault that can be observed in the baseline results compared to the Laplacian filter-ablated model was that the tumor edge in the enhanced sequence (T1Gd) was visually over-enhanced compared to the target image. Similarly, the Laplacian filter-ablated model exhibited slightly less structural similarity to the target images compared to the baseline. The Laplacian filter operates by calculating edges based on pixel intensity changes within neighboring pixels. While this filter enhances the model's ability to learn high frequency details, such as edges in tumor regions, it is also sensitive to noise in the input data. Nevertheless, our analysis showed that in tumor regions where detailed edge information is crucial, the Laplacian filter consistently improved the model's performance, as evidenced by the superior results in Table 3. This indicates that despite its sensitivity to noise, the Laplacian filter contributes positively to synthesizing high-quality MR images, particularly in regions of clinical importance.

For the data augmentation and imputation experimental result shown in Tables 4 and 5, CollaGAN had slightly better performance than StarGAN v2. This result seems contrary to the evaluation results in Tables 2 and 3, where StarGAN v2 showed better evaluation metrics. This may be due to three reasons. First, in the CollaGAN samples, there was a large number of inappropriately synthesized tumors. For example, the tumor region in the

CollaGAN samples showed visual characteristics of all four modalities mixed together. In Fig. 4, a T1Gd-like enhanced tumor shape is observable in the FLAIR, T1, and T2 sequences. Thus, in the case of the segmentation tasks where edge information is crucial, the CollaGAN samples were better at learning the tumor edges than the StarGAN v2 samples. Second, since the StarGAN v2 is a 1-to-n translation model, its performance is heavily dependent on the source image and reference image. If the tumor region in the source image was not sufficiently visible, the synthesized image was an almost tumor-less normal MR image (see bottom row of Fig. 5). Third, when performing reference-guided synthesis in StarGAN v2, the reference images were selected from other patients' MR images. Therefore, a large variance in image contrast can occur that may lead to over- or under-segmentation<sup>35</sup>. ResViT demonstrated strong performance across both augmentation and imputation tasks due to its hybrid architecture combining convolutional neural networks (CNNs) and vision transformers (ViTs). This design allowed ResViT to effectively balance local precision and global contextual sensitivity, which could be beneficial for synthesizing high-quality tumor regions. As shown in Tables 4 and 5, ResViT achieved higher DICE scores than both StarGAN v2 and CollaGAN for segmentation tasks under augmented and imputed conditions. However, artifacts were frequently observed in synthesized tumor regions (Fig. 5), which were more concentrated compared to CollaGAN's artifacts that appeared across entire images. We assume that this is due to the transformer architecture in ResViT, which generally requires larger datasets to function optimally. In this work, we attempted to maintain consistent training configurations across all models, which may have limited ResViT's ability to fully leverage its transformer-based design.

LASTGAN demonstrates several key advantages in brain tumor MRI imputation compared to StarGAN v2, CollaGAN, and ResViT. Unlike StarGAN v2, which uses a 1-to-n translation approach that can struggle with tumor visibility when source images have low tumor contrast, LASTGAN's multiple inputs with Laplacian filter attention better preserves tumor-specific features. While CollaGAN and ResViT attempts to handle multiple inputs using mask vectors and utilizing transformer architecture, respectively, it often produces artifacts and shows mixed modality characteristics in tumor regions. LASTGAN addresses these limitations through its novel combination of Laplacian filter attention and style transfer, resulting in superior performance across all quantitative metrics. The improvement is particularly pronounced in tumor region evaluation, where LASTGAN achieves significantly better accuracy (NRMSE: 0.14919, PSNR: 68.02320) than both comparative methods. These results demonstrate that LASTGAN effectively address the limitations of previous approaches in maintaining both structural consistency and domain-specific features in synthesized MRI sequences.

Although our LASTGAN model synthesizes MRI samples in high quality, some other limitations remain. Since one of our primary goals when designing the method was to make the network tumor-aware, tumor mask images were necessary to realize the Laplacian filter attention mechanism, which plays a significant role in our model. For this, we fed tumor masks in the discriminator to make the generator focus on tumor regions. Another issue is that our method uses three aforementioned multiple cyclic losses, which consumes a large amount of GPU memory. As a consequence, LASTGAN was trained with an insufficient batch size due to hardware limits. Lastly, the model struggled with synthesizing tumors in the enhanced sequence (T1Gd in this study). This issue is partially because of the dataset, where T1Gd is the only contrast-enhanced image set. Thus, the generator in this work had to synthesize enhanced tumors from the provided unenhanced sequences, which is a hard task compared to synthesizing other unenhanced sequences because the required information, such as enhanced tumor edges, is not fully provided. We expect that using multi-class tumor masks may work better for enhanced sequences because this can induce the model learn the tumors by their types. This issue remains for future work.

## Conclusion

In this study, we proposed LASTGAN for the imputation of missing sequences in brain tumor MR images. Our method combines image imputation and style transfer methods, by which LASTGAN can more accurately synthesize brain tumor MR images. We also proposed a new attention mechanism called Laplacian filter attention. With the proposed attention module, LASTGAN was able to synthesize domain-specific regions, such as brain tumors, while preserving the domain-shared anatomical structures of the brain. Moreover, segmentation experiments showed that the samples generated by the proposed method are feasible for data augmentation and imputation. We anticipate that our proposed method can be applied for missing data imputation in the near future. As LASTGAN enables the synthesis of missing MRI sequences, we believe that our method can support the training of medical artificial intelligence models, specifically targeting rare diseases.

## Data availability

The brain tumor MRI datasets analyzed in this study are available in the Medical Segmentation Decathlon<sup>31–33</sup> task01 brain tumour, <http://medicaldecathlon.com/>.

Received: 7 July 2024; Accepted: 8 September 2025

Published online: 10 October 2025

## References

1. Liu, Y., Meng, L. & Zhong, J. Magan: mask attention generative adversarial network for liver tumor ct image synthesis. *J. Healthcare Eng.* **2021**, 1–11 (2021).
2. Sun, L. et al. Hierarchical amortized gan for 3d high resolution medical image synthesis. *IEEE J. Biomed. Health Inform.* **26**, 3966–3975 (2022).
3. Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M. & Hammoudeh, M. A. A. Braingan: brain mri image generation and classification framework using gan architectures and cnn models. *Sensors* **22**, 4297 (2022).
4. Hirte, A. U. et al. Realistic generation of diffusion-weighted magnetic resonance brain images with deep generative models. *Magn. Reson. Imaging* **81**, 60–66 (2021).

5. Shen, L. et al. Multi-domain image completion for random missing input data. *IEEE Trans. Med. Imaging* **40**, 1113–1122 (2020).
6. Ma, C., Luo, G. & Wang, K. Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of mr images. *IEEE Trans. Med. Imaging* **37**, 1943–1954 (2018).
7. Huo, Y. et al. Splenomegaly segmentation on multi-modal mri using deep convolutional networks. *IEEE Trans. Med. Imaging* **38**, 1185–1196 (2018).
8. Shin, H.-C. et al. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *Simulation and Synthesis in Medical Imaging: Third International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 3* 1–11 (Springer, 2018).
9. Little, R. J. & Rubin, D. B. *Statistical Analysis With Missing Data*, vol. 793 (Wiley, 2019).
10. Ali, H. et al. The role of generative adversarial networks in brain mri: a scoping review. *Insights Imaging* **13**, 98 (2022).
11. Yi, X., Walia, E. & Babyn, P. Generative adversarial network in medical imaging: a review. *Med. Image Anal.* **58**, 101552 (2019).
12. Azad, R., Khosravi, N., Dehghanmashadi, M., Cohen-Adad, J. & Merhof, D. Medical image segmentation on mri images with missing modalities: a review. *arXiv preprint arXiv:2203.06217* (2022).
13. Conte, G. M. et al. Generative adversarial networks to synthesize missing t1 and flair mri sequences for use in a multisequence brain tumor segmentation model. *Radiology* **299**, 313–323 (2021).
14. Dai, X. et al. Multimodal mri synthesis using unified generative adversarial networks. *Med. Phys.* **47**, 6343–6354 (2020).
15. Xin, B., Hu, Y., Zheng, Y. & Liao, H. Multi-modality generative adversarial networks with tumor consistency loss for brain mr image synthesis. In *2020 IEEE 17th international symposium on biomedical imaging (ISBI)* 1803–1807 (IEEE, 2020).
16. Zhang, H., Li, H., Dillman, J. R., Parikh, N. A. & He, L. Multi-contrast mri image synthesis using switchable cycle-consistent generative adversarial networks. *Diagnostics* **12**, 816 (2022).
17. Jiang, Y., Zhang, S. & Chi, J. Multi-modal brain tumor data completion based on reconstruction consistency loss. *J. Digit. Imaging* **36**, 1794–1807 (2023).
18. Lee, D., Kim, J., Moon, W.-J. & Ye, J. C. Collagan: collaborative gan for missing image data imputation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* 2487–2496 (2019).
19. Joyce, T., Chartsias, A. & Tsafaris, S. A. Robust multi-modal mr image synthesis. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11–13, 2017, Proceedings, Part III* 20 347–355 (Springer, 2017).
20. Zhou, T., Fu, H., Chen, G., Shen, J. & Shao, L. Hi-net: hybrid-fusion network for multi-modal mr image synthesis. *IEEE Trans. Med. Imaging* **39**, 2772–2781 (2020).
21. Dalmaz, O., Yurt, M. & Çukur, T. Resvit: residual vision transformers for multimodal medical image synthesis. *IEEE Trans. Med. Imaging* **41**, 2598–2614 (2022).
22. Zhang, Y. et al. Unified multi-modal image synthesis for missing modality imputation. *IEEE Trans. Med. Imaging* **2024**, 125 (2024).
23. Wang, Y. et al. A unified hybrid transformer for joint mri sequences super-resolution and missing data imputation. *Phys. Med. Biol.* **68**, 135006 (2023).
24. Cho, J., Woo, J. & Park, J. A unified framework for synthesizing multisequence brain mri via hybrid fusion. *arXiv preprint arXiv:2406.14954* (2024).
25. Özbey, M. et al. Unsupervised medical image translation with adversarial diffusion models. *IEEE Trans. Med. Imaging* **2023**, 526 (2023).
26. Zhang, D., Wang, C., Chen, T., Chen, W. & Shen, Y. Scalable swin transformer network for brain tumor segmentation from incomplete mri modalities. *Artif. Intell. Med.* **149**, 102788 (2024).
27. Liu, J. et al. One model to synthesize them all: multi-contrast multi-scale transformer for missing data imputation. *IEEE Trans. Med. Imaging* **42**, 2577–2591 (2023).
28. Meng, X., Sun, K., Xu, J., He, X. & Shen, D. Multi-modal modality-masked diffusion network for brain mri synthesis with random modality missing. *IEEE Trans. Med. Imaging* **2024**, 256 (2024).
29. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* 2223–2232 (2017).
30. Zhao, H., Gallo, O., Frosio, I. & Kautz, J. Loss functions for image restoration with neural networks. *IEEE Trans. Comput. Imaging* **3**, 47–57 (2016).
31. Menze, B. H. et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**, 1993–2024 (2014).
32. Antonelli, M. et al. The medical segmentation decathlon. *Nat. Commun.* **13**, 4128 (2022).
33. Bakas, S. et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629* (2018).
34. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In *booktitleMedical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18 234–241 (Springer, 2015).
35. Liu, M. et al. Style transfer using generative adversarial networks for multi-site mri harmonization. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III* 24 313–322 (Springer, 2021).

## Acknowledgements

This work was supported in part by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI21C1161); in part by an Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Ministry of Science and ICT (MSIT), Republic of Korea (No.RS-2020-II201336, Artificial Intelligence Graduate School Program (UNIST)); and in part by a grant from the National Research Foundation of Korea (NRF) funded by the MSIT, Republic of Korea (No. NRF-2021R1F1A1057818, RS-2024-00344958).

## Author contributions

Conceptualization, Yoonho Na and Jimin Lee; Data curation, Sung Soo Ahn and Ji Eun Park; Funding acquisition, Jimin Lee; Investigation, Yoonho Na, Kyuri Kim, Hyungjoo Cho and Hwiyoung Kim; Methodology, Yoonho Na and Jimin Lee; Project administration, Hwiyoung Kim, Sung-Joon Ye, and Jimin Lee; Software, Yoonho Na; Supervision, Jimin Lee; Validation, Yoonho Na and Kyuri Kim; Visualization, Yoonho Na and Sung-Joon Ye; Writing–original draft, all authors; Writing–review & editing, all authors.

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to J.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025