




## Article

# Within-Modality Synthesis and Novel Radiomic Evaluation of Brain MRI Scans

Seyed Masoud Rezaei<sup>1</sup> , Nahid Chegeni<sup>1</sup>, Fariborz Baghaei Naeini<sup>2</sup>, Dimitrios Makris<sup>2</sup>   
and Spyridon Bakas<sup>2,3,\*</sup> 

<sup>1</sup> Department of Medical Physics, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran; masoudrezayi1398@gmail.com (S.M.R.)

<sup>2</sup> Faculty of Engineering, Computing and the Environment, Kingston University, Penrhyn Road Campus, Kingston upon Thames, London KT1 2EE, UK; f.baghaeinaeini@kingston.ac.uk (F.B.N.); d.makris@kingston.ac.uk (D.M.)

<sup>3</sup> Richards Medical Research Laboratories, Center for Biomedical Image Computing and Analytics (CBICA), University of Pennsylvania, Floor 7, 3700 Hamilton Walk, Philadelphia, PA 19104, USA

\* Correspondence: spyridon.bakas@pennteam.upenn.edu

**Simple Summary:** Brain MRI scans often require different imaging sequences based on tissue types, posing a common challenge. In our research, we propose a method that utilizes Generative Adversarial Networks (GAN) to translate T2-weighted-Fluid-attenuated-Inversion-Recovery (FLAIR) MRI volumes into T2-Weighted (T2W) volumes, and vice versa. To evaluate the effectiveness of our approach, we introduce a novel evaluation schema that incorporates radiomic features. We train two distinct GAN-based architectures, namely Cycle GAN and Dual Cycle-Consistent Adversarial network (DC<sup>2</sup>Anet), using 510 pair-slices from 102 patients. Our findings indicate that the generative methods can produce results similar to the original sequence without significant changes in radiometric features. This method has the potential to assist clinicians in making informed decisions based on generated images when alternative sequences are unavailable, or time constraints prevent re-scanning MRI patients.

**Abstract:** One of the most common challenges in brain MRI scans is to perform different MRI sequences depending on the type and properties of tissues. In this paper, we propose a generative method to translate T2-Weighted (T2W) Magnetic Resonance Imaging (MRI) volume from T2-weight-Fluid-attenuated-Inversion-Recovery (FLAIR) and vice versa using Generative Adversarial Networks (GAN). To evaluate the proposed method, we propose a novel evaluation schema for generative and synthetic approaches based on radiomic features. For the evaluation purpose, we consider 510 pair-slices from 102 patients to train two different GAN-based architectures Cycle GAN and Dual Cycle-Consistent Adversarial network (DC<sup>2</sup>Anet). The results indicate that generative methods can produce similar results to the original sequence without significant change in the radiometric feature. Therefore, such a method can assist clinics to make decisions based on the generated image when different sequences are not available or there is not enough time to re-perform the MRI scans.

**Keywords:** MRI synthesis; T2W; FLAIR; radiomic; CycleGAN; DC<sup>2</sup>Anet



**Citation:** Rezaei, S.M.; Chegeni, N.; Baghaei Naeini, F.; Makris, D.; Bakas, S. Within-Modality Synthesis and Novel Radiomic Evaluation of Brain MRI Scans. *Cancers* **2023**, *15*, 3565. <https://doi.org/10.3390/cancers15143565>

Academic Editor: Sam Payabvash

Received: 20 May 2023

Revised: 27 June 2023

Accepted: 5 July 2023

Published: 10 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Medical imaging scans, including Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), are routinely acquired and used clinically to macroscopically assess, diagnose, and monitor patients with brain abnormalities. MRI in particular can depict normal anatomy and apparent pathologies while providing data relating to the anatomical structure, tissue density, and microstructure, as well as tissue vascularization, depending on the acquired sequence [1–3]. Structural MRI sequences represent the basic scans acquired across comprehensive centers and community-based healthcare sites, comprising

native T1-weighted (T1W), post-contrast T1-weighted (T1Gd), T2-weighted (T2W), and T2-weighted Fluid-Attenuated Inversion Recovery (FLAIR). T1W scans facilitate observation and analysis of the brain anatomy, with the T1Gd scans particularly being able to easily identify the boundaries of an active tumor, while T2-weighted scans (T2w and FLAIR) help in identifying brain abnormalities, both those related to vascular lesions (e.g., stroke) and vasogenic edema [4]. The simultaneous assessment of multiple varying MRI scans (also known as multi-parametric MRI—mpMRI) from the same patient is the standard clinical practice for the evaluation of patients suspected of stroke or diffuse glioma, as it offers the maximal available medical diagnostic information.

Acquisition of mpMRI might not be possible at all times, due to numerous reasons, including but not limited to the patient's cooperation during a scanning session that could result in motion-degraded scans, thereby hindering further diagnostic usage [5–7]. Towards this end, the artificial synthesis of specific MRI scans has been an active area of research [8,9], with the intention of either substituting specific MRI scans corrupted by various artifacts, or generating scans that were not acquired at all. Although such synthetic scans have been successfully used in many applications, and Generative Adversarial Networks (GANs) have significantly improved their realism, the final result may not always look as realistic and/or may contain information that adversely affects downstream quantitative analyses [10,11]. Cross-domain synthesis of medical images has drawn significant interest in the medical imaging community and describes the artificial generation of a target-modality scan by learning the relationship between paired given source-modality scans and their associated target modality scans [12,13]. Of note, the data is here described as paired when it arises from the same individual at different points in time.

In recent years, deep-learning procedures and particularly Convolutional Neural Networks (CNNs) [14] and GANs have rapidly dominated the domain of medical image synthesis [15,16]. GANs use two competing CNNs: one that generates new images and another that discriminates the generated images as either real or fake. To address the problem of unpaired cross-domain data, which is common in healthcare, the Cycle Generative Adversarial Network (CycleGAN) [17] is typically chosen to obtain high-quality information translatable across images. In CycleGAN, based on the image of a subject  $b_1$  in the source domain, the purpose is to estimate the relevant image of the same subject  $b_2$  in the target domain. In theory, the CycleGAN model entails two mapping functions, i.e.,  $G_1: X \rightarrow Y$  and  $G_2: Y \rightarrow X$ , and associated adversarial discriminators  $D_Y$  and  $D_X$ .  $D_Y$  encourages  $G_1$  to translate  $X$  into outputs indistinguishable from domain  $Y$  and contrariwise for  $D_X$  and  $G_2$ . Nie et al. [18] trained a fully convolutional network (FCN) to generate CT scans from corresponding MRI scans. They specifically used the adversarial training method to train their FCN. Welander et al. [19] evaluated two models, i.e., CycleGAN and UNIT [20], for image-to-image translation of T1- and T2W MRI slices by comparing synthetic MRI scans to real ones. They used paired T1W and T2W images from 1113 axial images (only slice 120). The scans were registered to a standard anatomical template, so they were in the same coordinate space and of the same size. Two models were compared using quantitative metrics, including mean absolute error (MAE), mutual information (MI), and peak signal-to-noise ratio (PSNR). It was shown that the executed GAN models can synthesize visually realistic MRI slices. Dar et al. [21] proposed a method for multi-contrast MRI synthesis based on Conditional GANs (CGANs). They demonstrated how CGANs could generate a T2W scan from a T1W. Theis et al. [22] found that GANs can generate more realistic training data to improve the classification performance of machine learning methods. Also, they showed that models creating more visually realistic synthetic images do not certainly have better quantitative error measurements when compared to real images. Despite the mounting promise of GANs for healthcare, both optimal model selection and quantitative evaluation remain challenging tasks, and solutions produced so far are use-specific and not generalizable. Specifically, for their quantitative performance evaluation, several metrics (such as the MAE, Mean Squared Error (MSE), and PSNR) have been proposed in the

literature [23], albeit no consensus has been reached on an optimal evaluation metric for a particular domain.

Radiomics describe a novel and rapidly advancing area in medical imaging. In contrast to the traditional clinical assessment of considering medical images as pictures intended only for visual interpretation, radiomics represent visual and sub-visual quantitative measurements (also called “features”) extracted from acquired radiology scans [24,25], following specific mathematical formulations, and resulting in measurements that are not even perceivable by the naked eye, i.e., sub-visual [26–34]. These features are widely used in both clinical and pre-clinical research studies attempting to identify associations between radiologic scans and clinical outcomes, or even molecular characteristics [35–38]. The hypothesis is that quantitative computational interrogation of medical scans can provide more and better information than the physician’s visual assessment. This is further exacerbated in observations related to the texture analyses of different imaging modalities. Various open-source tools have been developed to facilitate the harmonized extraction of high throughput radiomic features [39–41], contributing to the increasing evidence of their value. The primary purpose of these tools has been to expedite robust quantitative image analyses based on radiomics and standardize both feature definitions and computation strategies, thereby guaranteeing the reproducibility and reliability of radiomic features. In this study, considering the importance of T2-weighted scans (T2W and FLAIR), we focus on generating FLAIR from T2W MRI scans, and vice versa, based on the CycleGAN [17] and dual cycle-consistent adversarial network (DC<sup>2</sup>Anet) [42] architectures. We further consider radiomics as a novel way to quantify the dissimilarity between the distribution of the actual/real and synthesized scans. We think that radiomics can represent the first potential solution for the quantitative performance evaluation of GANs in the domain of radiology. For comparison, we also compare with traditional metrics, including MSE, MAE, and PSNR [43].

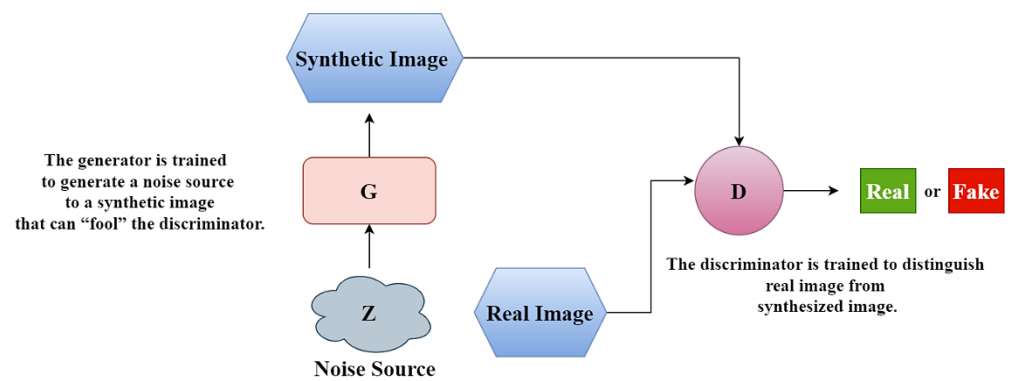
## 2. Methods

### 2.1. Dataset and Registration

The data utilized in our study were obtained from the public data collection ‘ACRIN-DSC-MR-Brain (ACRIN 6677/RTOG 0625)’ [44] at The Cancer Imaging Archive (TCIA) [45]. These data describe brain MRI scans from a multicenter phase-II trial of bevacizumab with temozolomide in recurrent glioblastoma (GBM) patients. We extracted 510 T2W/FLAIR paired slices from 102 patients, which were further divided into 410 pairs from 81 patients for training and 100 pairs from 21 patients to evaluate synthesis results (test set). Due to the limitation of the dataset, the validation set is not considered to train the networks. Each pair include the axial paired T2W/FLAIR slices for the same patient and at the same axial depth. Of note, the testing data are held out of the training process at all times. Our networks take 2D axial-plane slices of the volumes as inputs. In the pre-processing step, T2W 2D scans were first rigidly registered to FLAIR scans using the ITK-SNAP software (version 3.8.0) [46], considering 6 degrees of freedom (i.e., 3 translations and 3 rotations).

### 2.2. CycleGAN

A GAN network uses an image generator (G) to synthesize images of a target domain and a discriminator (D) to distinguish between real and synthesized images. A suitable analogy for visual data considers one network as an art forger and the other as an art specialist. The forger (generator G) creates forgeries. The specialist, (discriminator D) receives both forgeries and real images and attempts to tell them apart (Figure 1). Both networks are in competition with each other and are trained simultaneously.



**Figure 1.** Simultaneous training of both models, the generator (G) and the discriminator (D), in competition with each other.

CycleGAN is a framework that allows the unpaired data to generate an image from one domain to another. Therefore, CycleGAN reduces the problem caused by the lack of paired data. A diagram of the CycleGAN model used in this study is presented in Figure 2. Let us assume  $n^A$  images  $x^A \in XA$  (e.g., T2W) and  $n^B$  images  $x^B \in XB$  (e.g., FLAIR). The CycleGAN for T2W and FLAIR images include two mappings  $G_{T2W}: T2W \rightarrow FLAIR$  and  $G_{FLAIR}: FLAIR \rightarrow T2W$ . Therefore, the proposed CycleGAN operates with two generators ( $G_{T2W}, G_{FLAIR}$ ) and two discriminators ( $D_{T2W}, D_{FLAIR}$ ). Given a T2W image,  $G_{T2W}$  learns to generate the respective FLAIR image of the same anatomy that is indistinguishable from real FLAIR images, whereas  $D_{T2W}$  learns to discriminate between synthetic and real FLAIR images. The architecture of the CycleGAN generator is adapted from [17] with 9 residual blocks after early convolutional layers. Similarly, given a FLAIR image,  $G_{FLAIR}$  learns to generate the respective T2W image of the same anatomy that is indistinguishable from real T2W images, whereas  $D_{FLAIR}$  learns to discriminate between synthetic and real T2W images. We apply adversarial losses to both mapping functions for matching the distribution of generated images to the data distribution in the target domain. For the mapping function  $G_{T2W}: T2W \rightarrow FLAIR$  and its discriminator  $D_{T2W}$ , the objective is expressed as follows:

$$I_{GAN}(G_{T2W}, D_{T2W}, T2W, FLAIR) \quad (1)$$

Similarly, the adversarial loss is presented for the mapping function  $G_{FLAIR}: FLAIR \rightarrow T2W$  and its discriminator  $D_{FLAIR}$  as follows:

$$I_{GAN}(G_{FLAIR}, D_{FLAIR}, FLAIR, T2W) \quad (2)$$

Also, the cycle consistency loss, or  $L_{cyc}$  (Forward cycle-consistency and Backward cycle-consistency), is used to keep the cycle consistency between the two sets of networks as follows:

$$L_{cyc}(G_{T2W}, G_{FLAIR}) = [\|G_{FLAIR}(G_{T2W}(T2W)) - T2W\|_1] + [\|G_{T2W}(G_{FLAIR}(FLAIR)) - FLAIR\|_1] \quad (3)$$

The loss of the whole CycleGAN network is:

$$I_{CycleGAN} = I_{GAN}(G_{T2W}, D_{T2W}, T2W, FLAIR) + I_{GAN}(G_{FLAIR}, D_{FLAIR}, FLAIR, T2W) + L_{cyc}(G_{T2W}, G_{FLAIR}) \quad (4)$$

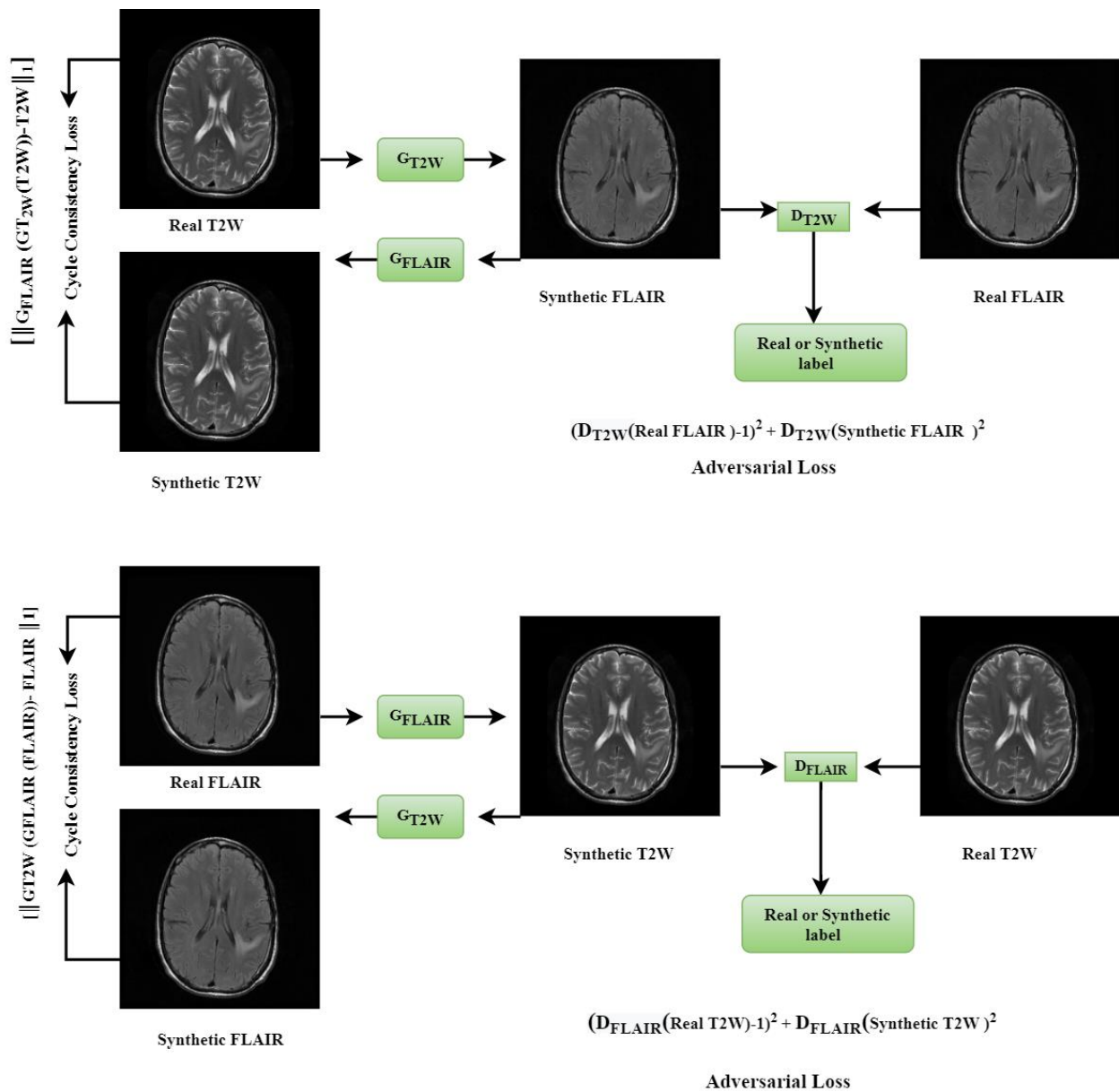


Figure 2. The CycleGAN model based on two mappings  $G_{T2W}: T2W \rightarrow FLAIR$  and  $G_{FLAIR}: FLAIR \rightarrow T2W$ .

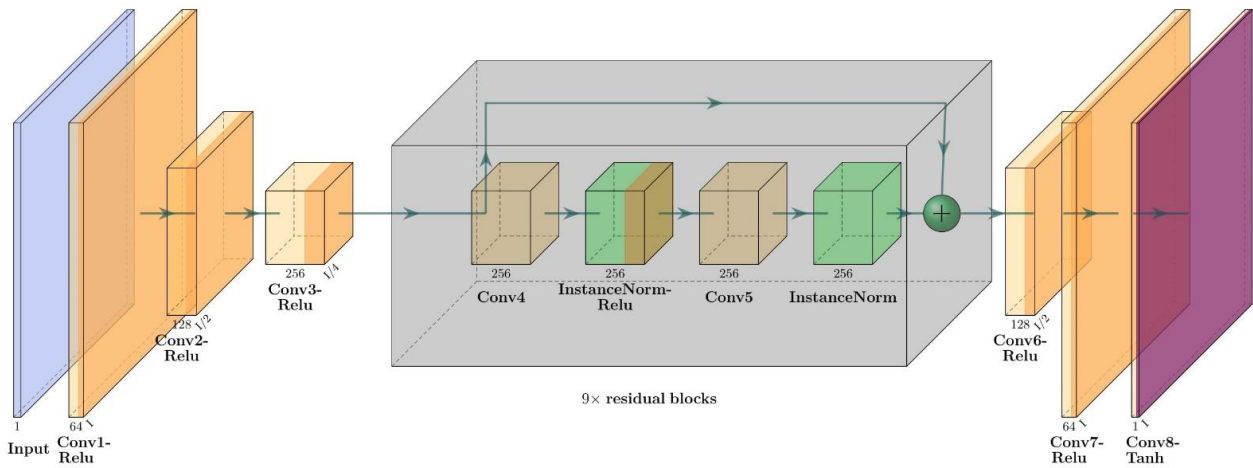
Figure 3 illustrates the architecture of the generator and discriminator networks within the proposed CycleGAN model, depicting the size and type of each layer.

### 2.3. DC<sup>2</sup>Anet

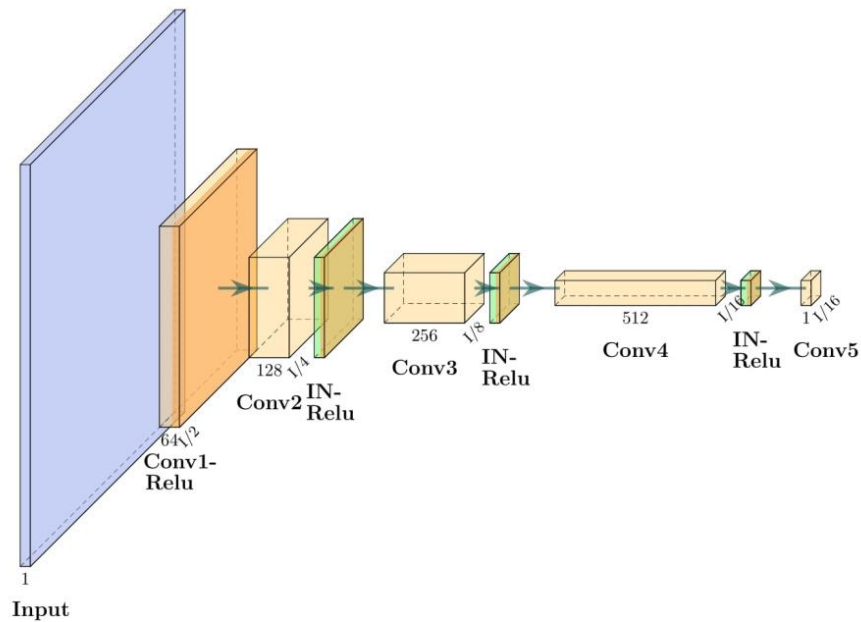
The DC<sup>2</sup>Anet model, introduced by Jin et al. [42], follows a semi-supervised learning approach that alternates between optimizing supervised and unsupervised learning in order to seek a global minimum for the optimal network. The forward and backward mappings are used to generate the T2W image from a FLAIR image, and vice versa. In the forward cycle-consistent adversarial network with aligned learning, the  $G_{FLAIR}$  network generates a synthetic T2W image from a FLAIR image, and this T2W image is then used by the  $G_{T2W}$  network to generate the original FLAIR image in order to learn the domain structures. The input to the discriminators  $Dis_{T2W}$  is either a sample T2W image from the real T2W data or a synthetic T2W image. In the backward cycle-consistent adversarial network with aligned learning, the  $G_{T2W}$  network generates a synthetic FLAIR image from

a T2W image, and this FLAIR image is then used by the  $G_{FLAIR}$  network to generate the original T2W image in order to learn the domain structures. The discriminator's  $Dis_{T2W}$  and  $Dis_{FLAIR}$  are expressed as follows:

$$\begin{aligned}
 L_{sup-adver} (Gen_{T2W}, Dis_{T2W}, Gen_{FLAIR}, Dis_{FLAIR}) = & \mathbb{E}_{FLAIR, T2W \sim p_{data}} (FLAIR, T2W) [\log(Dis_{T2W}(FLAIR, T2W))] \\
 & + \mathbb{E}_{FLAIR \sim p_{data}} (FLAIR) [\log(1 - Dis_{T2W}(FLAIR, Gen_{T2W}(FLAIR)))] \\
 & + \mathbb{E}_{T2W, FLAIR \sim p_{data}} (T2W, FLAIR) [\log(Dis_{FLAIR}(T2W, FLAIR))] \\
 & + \mathbb{E}_{T2W \sim p_{data}} (T2W) [\log(1 - Dis_{FLAIR}(T2W, Gen_{FLAIR}(DWI_{T2W})))]
 \end{aligned}
 \tag{5}$$



(a)



(b)

**Figure 3.** The architecture of proposed CycleGAN model; (a,b) are generator and discriminator networks respectively. Conv and IN are Convolution and Instance normalization, respectively.

In the DC2Anet model, in addition to adversarial and dual cycle consistency used in the CycleGAN model, to achieve accurate and perceptual outputs, four loss functions were measured as follows: voxel-wise, gradient difference, perceptual, and structural similarity losses. We name the loss function of the supervised training of our model as  $L_{sup}$ . Then the value of this function is defined as a weighted sum of  $L_{sup-adversarial}$ ,  $L_{sup-cycle-consistency}$ , and these four losses ( $L_{voxel-wise}$ ,  $L_{gradient}$ ,  $L_{perceptual}$ ,  $L_{structural}$ ). The weights used to calculate the  $L_{sup}$  are hyperparameters of the model, and we set all of them to be one in our experiment. Hence, these four terms are combined, and the relation between them is as follows:  $L_{sup} = L_{sup-adversarial} + L_{sup-cycle-consistency} + L_{voxel-wise} + L_{gradient} + L_{perceptual} + L_{structural}$ . A diagram outlining the forward and backward adversarial losses of the DC<sup>2</sup>Anet model is shown in Figure 4. Of note, the generator network of the DC<sup>2</sup>Anet model is the same as the CycleGAN model (Figure 3a), but the discriminator network is different. The architecture of this discriminator is shown in Figure 5.

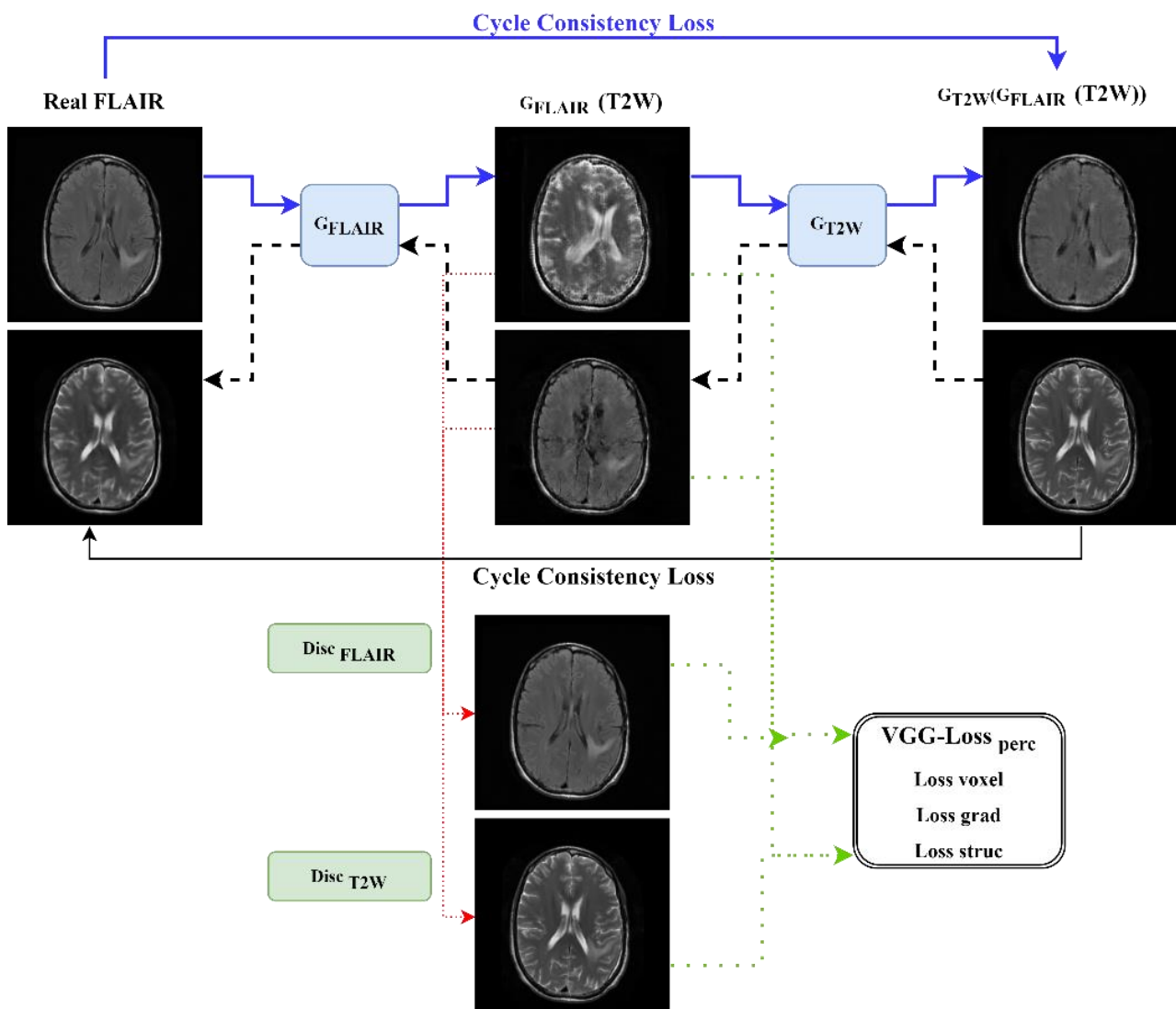
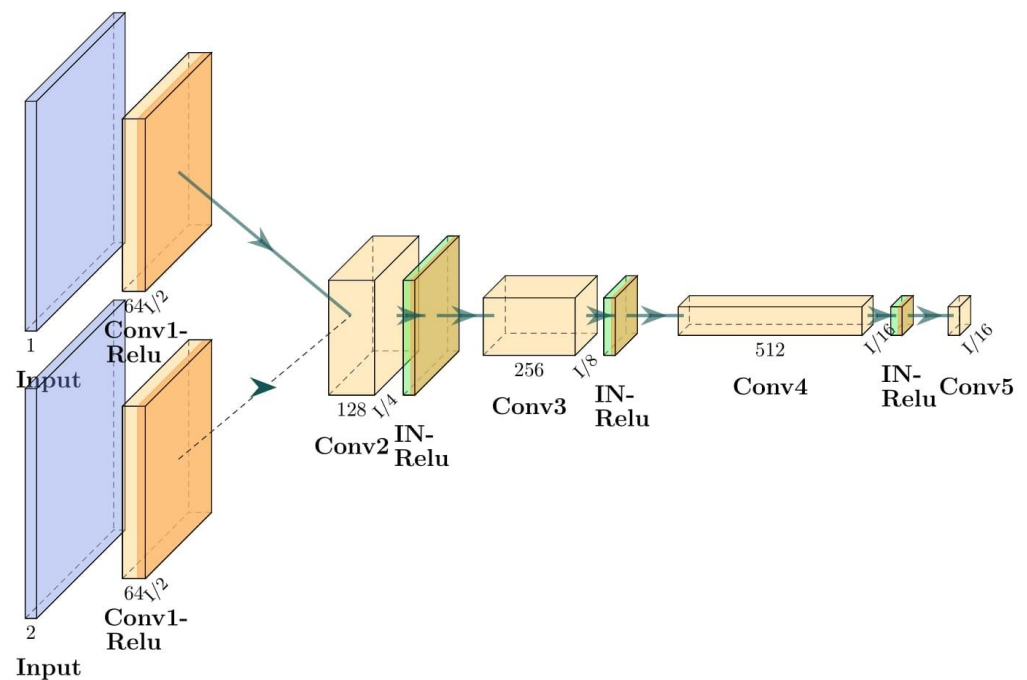


Figure 4. Diagram of the DC<sup>2</sup>Anet model.



**Figure 5.** The architecture of discriminator used for DC<sup>2</sup>Anet model. Conv: Convolution. IN: Instance Normalization.

#### 2.4. Implementation

To generate a T2W MRI from a FLAIR, and vice versa, the T2W and FLAIR values are converted to  $[0, 1]$  tensor. The resolutions of the FLAIR and T2W images in our dataset are  $256 \times 256$  and  $512 \times 512$  respectively. Therefore, in the first preprocessing step, FLAIR images are registered to T2W images using rigid registration to ensure that all images have a  $256 \times 256$  resolution. Then, the axial T2W/FLAIR pairs were the input of the network with  $256 \times 256$  pixels. Of note, instead of image patches, whole 2D images are used for training. CycleGAN and DC<sup>2</sup>Anet were trained for 400 epochs. The batch size was set to 2, and both the generator and the discriminator used the Adam optimizer [47]. In the first 200 epochs, the learning rate was fixed at  $2 \cdot 10^{-4}$ . For the rest 200 epochs, the learning rate linearly decayed from  $2 \cdot 10^{-4}$  to 0. It was observed that the discriminators found success faster than the generators, therefore the values of iterations for the generator and the discriminator were set to three and one, respectively. The loss function plays a crucial role in training the networks and achieving the desired translation between T2-weighted (T2W) and FLAIR images. In both CycleGAN and DC<sup>2</sup>Anet, adversarial losses are employed to match the distribution of generated images with the target domain data distribution. Specifically, for the mapping function  $G_{T2W}: T2W \rightarrow FLAIR$  and its discriminator  $D_{T2W}$ , we use the adversarial loss defined as  $L_{GAN}(G_{T2W}, D_{T2W}, T2W, \text{and } FLAIR)$ . Similarly, for the mapping function  $G_{FLAIR}: FLAIR \rightarrow T2W$  and its discriminator  $D_{FLAIR}$ , the adversarial loss is expressed as  $L_{GAN}(G_{FLAIR}, D_{FLAIR}, FLAIR, \text{and } T2W)$ . Additionally, to maintain cycle consistency between the generator networks, a cycle consistency loss ( $L_{cyc}$ ) is utilized, which ensures that the translated images can be successfully converted back to the original domain. The complete loss function for the CycleGAN network is given by  $L_{CycleGAN} = L_{GAN}(G_{T2W}, D_{T2W}, T2W, FLAIR) + L_{GAN}(G_{FLAIR}, D_{FLAIR}, FLAIR, T2W) + L_{cyc}(G_{T2W}, G_{FLAIR})$ .

Cyc ( $G_{T2W}, G_{FLAIR}$ ) losses were multiplied by a constant lambda ( $\lambda$ ) based on the importance of the cycle consistency losses (Equation (4)) concerning the adversarial loss; therefore, we set  $\lambda_{cyc}(G_{T2W}, G_{FLAIR}) = 10$  and  $\lambda_{GAN} = 1$ . All experiments including data preprocessing and analysis were performed on the Google Cloud computing service “Google Colab” ([colab.research.google.com](https://colab.research.google.com)) using Python 3.7 and TensorFlow 2.4.1. The parameters of training and hardware configurations are provided in (Table 1).



**Table 1.** Network Hyperparameters, Tuning, and Computing for Image Translation using CycleGAN and DC<sup>2</sup>Anet.

Implementation (Hyperparameter, Tuning, and Computing)	Value
Resolution (FLAIR)	256 × 256
Resolution (T2W)	512 × 512
Preprocessing	Rigid registration to ensure 256 × 256 resolution for FLAIR images
Input Image Size	256 × 256 pixels
Training Epochs	400
Batch Size	2
Optimizer	Adam
Learning Rate	2 · 10 <sup>-4</sup> (fixed for the first 200 epochs, linear decay to 0 for the next 200 epochs)
Generator Iterations	3
Discriminator Iterations	1
Cycle Consistency Loss Weight (λ <sub>cyc</sub> )	10
Adversarial Loss Weight (λ <sub>GAN</sub> )	1
Computing Platform	Google Colab
Programming Language	Python 3.7
Deep Learning Framework	TensorFlow 2.4.1

### 2.5. Evaluation

Several metrics were used to compare the real and synthetic T2W, and FLAIR, images. These metrics including MAE, MSE, and PSNR were used widely in the literature for the same purpose [48–50]. These metrics are defined as:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N | \text{real MRI(T2W/FLAIR)}(i) - \text{synthetic MRI(T2W/FLAIR)}(i) | \quad (6)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\text{real MRI(T2W/FLAIR)}(i) - \text{synthetic MRIT2W/FLAIR}(i))^2 \quad (7)$$

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right) \quad (8)$$

where N is the total number of voxels inside the input image and i is the index of the aligned pixel. MAX denotes the largest pixel value of ground truth T2W and synthetic T2W images, and vice versa, for FLAIR images.

Considering a use-inspired generalizable evaluation approach, beyond just the essential quantification but also considering radiologic appearance, in this study we introduce a novel approach based on radiomic features to compare the real and synthetic images. After running the models and generating all the synthetic images, we segmented the whole brain volume, both in the real and synthetic images, using the module of ITK-SNAP [51,52] within the Cancer Imaging Phenomics Toolkit (CaPTk) [41,53]. CaPTk is a software platform written in C++ to analyze medical images. The package leverages the value of quantitative imaging analytics along with machine learning to derive phenotypic imaging signatures. The specific ITK-SNAP module is based on the geometric active contour model and defines the contours using energy forces and the geometric flow curve. The counter is a collection of points that undergo the interpolation process. In this study, manual delineation of the whole brain and skull is performed. Following this segmentation, radiomic features com-

pliant with the Image Biomarker Standardization Initiative (IBSI) [54] were extracted from both the real and synthetic T2W, as well as from the real and synthetic FLAIR images using the CaPTk 1.9.0 software [53,55]. Twenty-three features were extracted, comprising 8 Gray Level Co-occurrence Matrix features (GLCM) [56,57], 8 Gray Level Size Zone (GLSZM) [58], and 7 Gray Level Run Length Matrix features (GLRLM) [54] (Table 2).

**Table 2.** Extracted Radiomic features for the evaluation of CycleGAN and DC<sup>2</sup>Anet translations.

Radiomic Features Extracted		
GLCM	GLSZM	GLRLM
Auto Correlation	Grey Level Mean	Grey Level Nonuniformity
Cluster Prominence	Grey Level Nonuniformity	High Grey Level Run Emphasis
Cluster Shade	Grey Level Variance	Long Run High Grey Level Emphasis
Contrast	High Grey Level Emphasis	Long Run Low Grey Level Emphasis
Correlation	Large Zone Emphasis	Low Gray Level Run Emphasis
Energy	Large Zone High Grey Level Emphasis	Run Length Nonuniformity
Entropy	Large Zone Low Grey Level Emphasis	Short Run Emphasis
Homogeneity	Low Grey Level Emphasis	

The choice of specific features in our study, namely the 8 Gray Level Co-occurrence Matrix (GLCM) features 8 Gray Level Size Zone (GLSZM) features, and 7 Gray Level Run Length Matrix (GLRLM) features, was based on their established significance in characterizing textural patterns and capturing distinct image characteristics (Table 2). These features have been widely utilized in radiomics studies and have demonstrated their efficacy in quantifying spatial relationships, size variations, and run lengths within an image. By utilizing this comprehensive set of radiomic features, we aimed to capture a wide range of textural characteristics that could potentially distinguish real and synthetic images. Details of radiomic features are included in Supplementary Materials (Supplementary S1).

These features have shown promise in previous studies as reliable indicators of image heterogeneity and structural differences. Their selection was based on their ability to quantitatively represent textural properties and provide discriminative information regarding the underlying tissue or lesion composition. We believe that the inclusion of these 23 selected radiomic features, derived from GLCM, GLSZM, and GLRLM matrices, offers a robust and comprehensive approach for evaluating the differentiation between real and synthetic images. Their relevance lies in their proven capability to capture textural patterns and provide meaningful insights into the image composition, thereby contributing to the assessment and discrimination of real and synthetic images in our study.

The exact parameters used for the feature extraction are: Bins = 20, Radius = 1, and Offset = Combined. Available studies use bin numbers varying from 8 to 1000, as suggested by the IBSI [59], or bin widths from 1 to 75. In this study, textural features were computed using the fixed-bin width approach. Radiomics data were then analyzed using the GraphPad Prism 9.5 software (GraphPad, San Diego, CA, USA). The normality of the extracted radiomic features was evaluated based on the D'Agostino test [60]. When data was described by a normal distribution we used a *t*-test, otherwise we used the Mann-Whitney U test was used. For determining whether the radiomic features between the real and synthetic T2W (as well as between the real and synthetic FLAIR images) are statistically significant, confidence intervals (commonly abbreviated as CI) were calculated.

### 3. Results

We evaluated the proposed CycleGAN and DC<sup>2</sup>Anet architectures on T2W and FLAIR brain tumor images. For the quantitative performance evaluation of the T2W and FLAIR synthesis, in line with current literature we considered the three metrics of MAE, MSE, and PSNR (Table 3).

**Table 3.** Results of three metrics between the real T2W and synthetic T2W and real FLAIR and synthetic FLAIR, including MAE, MSE, and PSNR. Bold text highlights the best result per column.

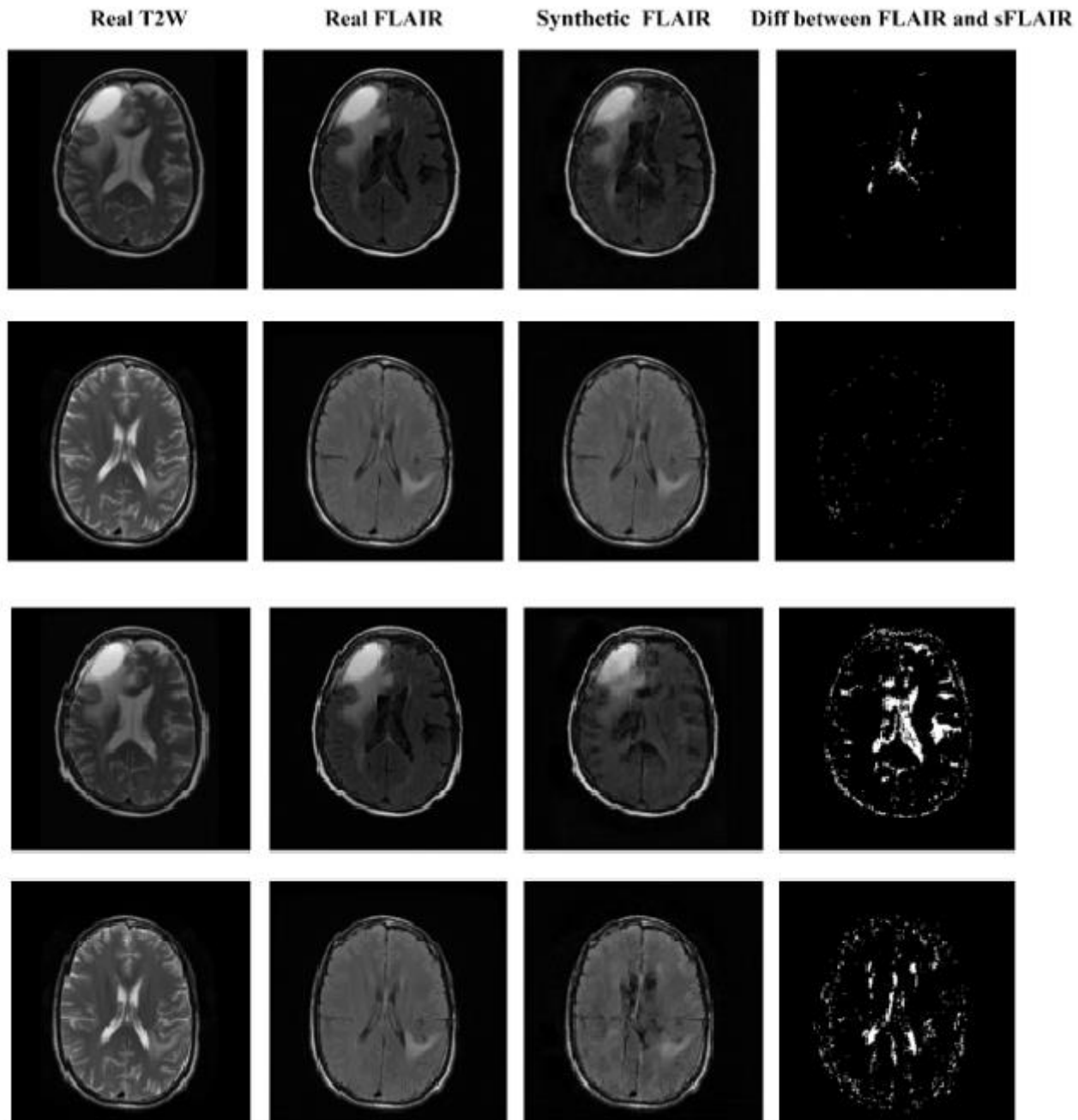
Experiment		MAE	MSE	PSNR
Synthetic and Real T2W	CycleGAN	24.2 ± 8.1	916.6 ± 157	27.5 ± 5.6
	DC <sup>2</sup> Anet	35.2 ± 12	1100 ± 169	<b>20.5 ± 8.2</b>
Synthetic and Real FLAIR	CycleGAN	<b>17.23 ± 8.2</b>	<b>813.8 ± 198</b>	29.7 ± 6.55
	DC <sup>2</sup> Anet	25.6 ± 6.2	915.8 ± 156	21 ± 7.2

The CycleGAN model shows superior performance when compared with the DC<sup>2</sup>Anet model across all the 3 quantitative metrics. In general, we observe a better performance for FLAIR images. Of note, smaller values of the MAE and MSE indicate better results, as opposed to PSNR where larger values are better. The quantitative superiority of CycleGAN and DC<sup>2</sup>Anet for FLAIR images corresponds to the visual realness and the mapping in Figures 6 and 7 which show axial views of five patients' synthetic and real images for CycleGAN and DC<sup>2</sup>Anet, for both the T2W and the FLAIR images. Of note, for both of the T2W and FLAIR images in the perceptual study similar to quantitative analyses, CycleGAN shows the best performance. Hence, the CycleGAN model trained using adversarial and dual cycle consistency generates more realistic synthetic MR images. More examples of CycleGAN and DC<sup>2</sup>Anet for FLAIR and T2W translation are provided in Supplementary Material (Figures S1 and S2).

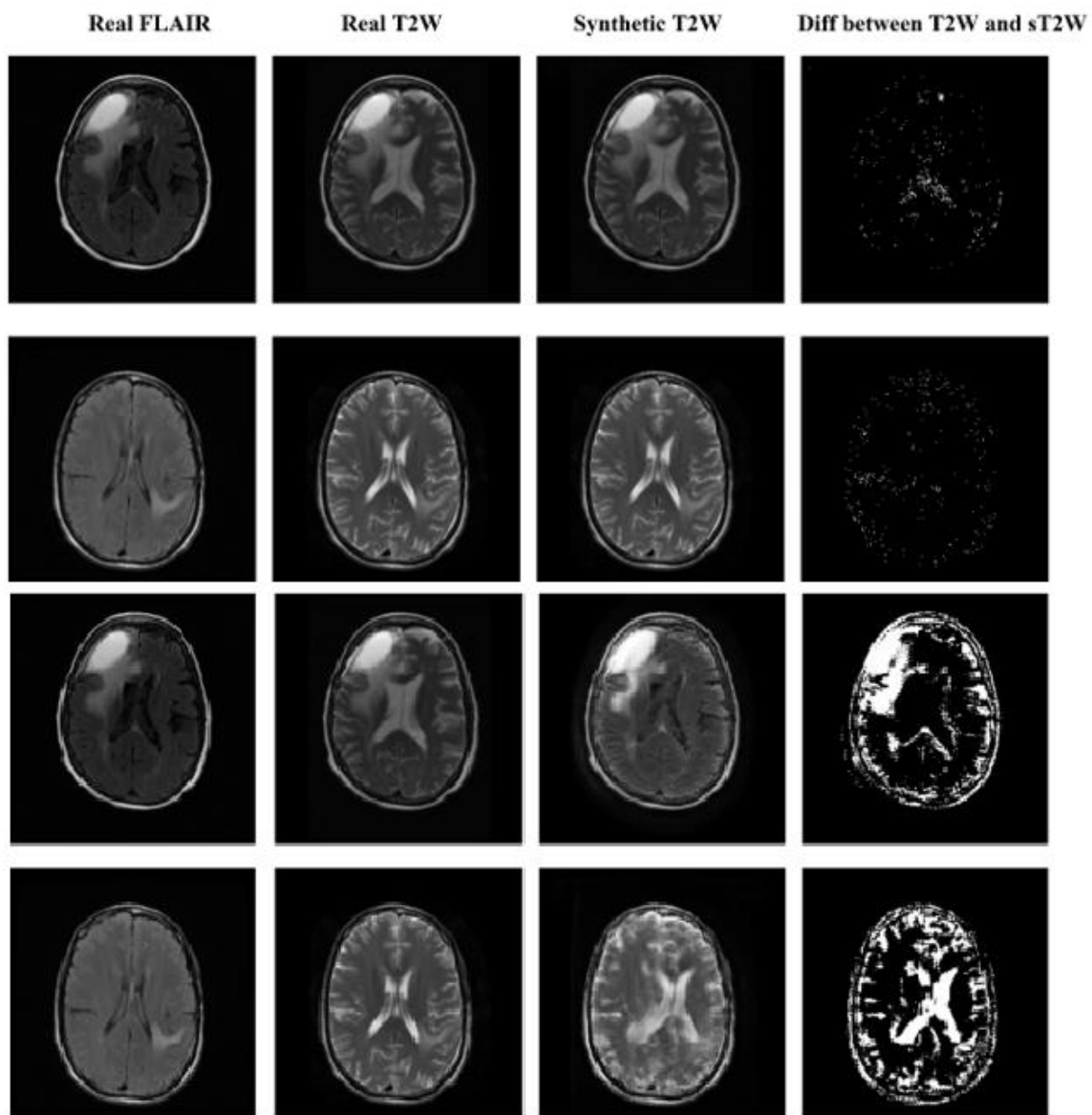
The differences observed between real and synthesis imaging modalities can be attributed to multiple factors. Firstly, synthesis images are computer-generated simulations of original images based on GAN models, while real original images are directly acquired from patients using MRI scanners. Consequently, variations can arise due to the inherent limitations and assumptions of the synthesis process. Furthermore, physiological and technical factors, such as variances in tissue contrast, signal intensity, and image artifacts, can contribute to dissimilarities between synthesis and real images. To further investigate and address these differences, future studies should focus on refining the synthesis algorithms, incorporating more realistic training data, and exploring the impact of various imaging parameters on the synthesis process.

We then performed a secondary quantitative performance evaluation that considers the radiophenotypical properties of the images, by virtue of the extracted radiomic features. Unlike the metrics of MAE, MSE, and PSNR, significance levels and values of radiomic features vary depending on the type of feature and image. Following a comparison of the radiomic features for both the T2W and FLAIR images, our results pointed out that for most radiomic features, there was no significant difference between the real and synthetic T2W, as well as for the real and synthetic FLAIR images. The mean and standard error (SE) of the GLCM features for both T2W and FLAIR images, as well as their statistically significant difference for the CycleGAN and DC<sup>2</sup>Anet models, are shown in Tables 4 and 5. No significant differences were observed for all extracted GLCM features using the CycleGAN model between real T2W images and synthetic T2W images. However, this was not the case for three features (cluster prominence, contrast, and correlation) extracted from the synthetic images of the DC<sup>2</sup>Anet model (Table 4). Notably, there was a significant difference for FLAIR images using the DC<sup>2</sup>Anet model for two extracted features (cluster prominence and correlation), and for the cluster prominence feature of the CycleGAN model (Table 5).

Similarly, with GLCM features, no significant differences were observed for all extracted GLRLM features using the CycleGAN model between real T2W images and synthetic T2W images (Table 6). However, with feature extraction on the DC<sup>2</sup>Anet synthetic images, there was a significant difference for two extracted features (High Grey Level Run Emphasis, and Long Run Low Grey Level Emphasis) between the real and synthetic T2W, as well as between the real and synthetic FLAIR images (Table 7).



**Figure 6.** Qualitative comparison of T2W to FLAIR for CycleGAN (First and second row) and DC<sup>2</sup>Anet (Third and fourth row).



**Figure 7.** Qualitative comparison of FLAIR to T2W for CycleGAN (First and second row) and DC<sup>2</sup>Anet (Third and fourth row).

**Table 4.** GLCM-based radiomic features for real and synthetic T2W for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real T2W	Synthetic T2W-CycleGAN	Synthetic T2W-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE **	Mean ± SE	Mean ± SE	Real T2W vs. Synthetic T2W-CycleGAN	Real T2W vs. Synthetic T2W-DC <sup>2</sup> Anet
Auto Correlation	330 ± 26.40	331 ± 52.22	348 ± 49.5	−176.8 to 173.8	−194.5 to 156.1
Cluster Prominence	17,303 ± 1501	15,367 ± 1009	12,133 ± 455	−2019 to 5892	1214 to 9125 *
Cluster Shade	722.4 ± 41.14	716 ± 18.76	677.2 ± 33.74	−111.2 to 124.0	−72.39 to 162.7
Contrast	43.96 ± 0.9	42.06 ± 1.04	31.67 ± 0.86	−1.487 to 5.296	8.902 to 15.68 *
Correlation	0.005 ± 0.0004	0.0051 ± 0.0007	0.05 ± 0.0098	−0.0212 to 0.0195	−0.0703 to −0.0296 *
Energy	0.43 ± 0.01	0.44 ± 0.01	0.43 ± 0.007	−0.040 to 0.027	−0.029 to 0.037
Entropy	2.47 ± 0.08	2.4 ± 0.03	2.34 ± 0.03	−0.139 to 0.263	−0.074 to 0.328
Homogeneity	0.67 ± 0.01	0.68 ± 0.009	0.63 ± 0.007	−0.044 to 0.021	0.0052 to 0.071

\* Significant difference; \*\* Standard error (SE).

**Table 5.** GLCM-based radiomic features for real and synthetic FLAIR for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real FLAIR	Synthetic FLAIR-CycleGAN	Synthetic FLAIR-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE	Mean ± SE	Mean ± SE	Real FLAIR vs. Synthetic FLAIR—CycleGAN	Real FLAIR vs. Synthetic FLAIR—DC <sup>2</sup> Anet
Auto Correlation	277 ± 5.69	269.6 ± 5.1	263.9 ± 5.84	−12.42 to 27.11	−6.673 to 32.86
Cluster Prominence	17372 ± 16.34	17134 ± 44.08	17036 ± 50.17	96.26 to 379.2 *	194.3 to 477.2 *
Cluster Shade	825.8 ± 42.54	819.7 ± 37.53	807.4 ± 40.61	−137.6 to 149.6	−125.2 to 162.0
Contrast	47.53 ± 1.99	45.03 ± 1.74	42.13 ± 1.78	−4.076 to 9.062	−1.166 to 11.97
Correlation	0.005 ± 0.0005	0.0005 ± 0.0003	0.05 ± 0.003	−0.0077 to 0.0076	−0.0531 to −0.0378 *
Energy	0.45 ± 0.006	0.42 ± 0.01	0.41 ± 0.01	−0.0283 to 0.0777	−0.0093 to 0.0968
Entropy	2.33 ± 0.02	2.41 ± 0.1	2.3 ± 0.11	−0.4098 to 0.2388	−0.2971 to 0.3515
Homogeneity	0.67 ± 0.01	0.65 ± 0.01	0.63 ± 0.01	−0.0287 to 0.0726	−0.0028 to 0.0985

\* Significant difference.

**Table 6.** GLRLM-based radiomic features for real and synthetic T2W for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real T2W	Synthetic T2W-CycleGAN	Synthetic T2W-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE	Mean ± SE	Mean ± SE	Real T2W vs. Synthetic T2W-CycleGAN	Real T2W vs. Synthetic T2W-DC <sup>2</sup> Anet
Grey Level Nonuniformity	73.95 ± 0.58	73.89 ± 0.84	74.66 ± 0.94	−2.821 to 2.937	−3.590 to 2.168
High Grey Level Run Emphasis	60.38 ± 2.3	56.89 ± 2.43	69.63 ± 1.56	−4.241 to 11.21	−16.87 to −1.630 *
Long Run High Grey Level Emphasis	92.88 ± 1.32	92.84 ± 1.95	95.25 ± 1.94	−6.258 to 6.338	−8.670 to 3.926
Long Run Low Grey Level Emphasis	29.69 ± 0.59	28.40 ± 0.60	25.88 ± 0.29	−0.5560 to 3.138	1.973 to 5.667 *
Low Gray Level Run Emphasis	0.62 ± 0.007	0.62 ± 0.006	0.61 ± 0.004	−0.0130 to 0.0300	−0.0068 to 0.0362
Run Length Nonuniformity	53.88 ± 1.57	55.08 ± 1.18	53.63 ± 0.94	−5.690 to 3.296	−4.237 to 4.749
Short Run Emphasis	0.48 ± 0.01	0.50 ± 0.01	0.47 ± 0.01	−0.0579 to 0.0336	−0.0313 to 0.0601

\* Significant difference.

**Table 7.** GLRLM-based radiomic features for real and synthetic FLAIR for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real FLAIR	Synthetic FLAIR-CycleGAN	Synthetic FLAIR-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE	Mean ± SE	Mean ± SE	Real FLAIR vs. Synthetic FLAIR-CycleGAN	Real FLAIR vs. Synthetic FLAIR-DC <sup>2</sup> Anet
Grey Level Nonuniformity	74.05 ± 0.29	71.9 ± 0.48	74.53 ± 0.78	0.1455 to 4.144 *	−2.477 to 1.522
High Grey Level Run Emphasis	69.12 ± 1.5	68.21 ± 1.72	62.75 ± 1.23	−4.442 to 6.259	1.019 to 11.72 *
Long Run High Grey Level Emphasis	118.4 ± 5.58	115.6 ± 5.34	114.3 ± 5.29	−16.41 to 22.17	−15.10 to 23.48
Long Run Low Grey Level Emphasis	31.27 ± 0.67	29.16 ± 1	25.5 ± 0.82	−0.9038 to 5.112	2.757 to 8.773 *
Low Gray Level Run Emphasis	0.62 ± 0.006	0.61 ± 0.009	0.62 ± 0.009	−0.0205 to 0.0412	−0.0301 to 0.0317
Run Length Nonuniformity	53.23 ± 1.15	58.62 ± 3.43	58 ± 3.31	−15.50 to 4.714	−14.87 to 5.338
Short Run Emphasis	0.51 ± 0.006	0.53 ± 0.01	0.52 ± 0.01	−0.0642 to 0.0236	−0.0631 to 0.0247

\* Significant difference.

Extracted features based on GLSZM for the two models used are shown in Tables 8 and 9. By comparing the real and synthetic T2W based on CycleGAN and DC<sup>2</sup>Anet, except for two features (Grey Level Nonuniformity, and Large Zone Low Grey Level Em-

phasis), no significant differences were observed for other extracted GLSZM features (Table 8). However, for FLAIR images using the DC<sup>2</sup>Anet model as shown in Table 8, for two features (High Grey Level Emphasis, and Large Zone Low Grey Level Emphasis), significant differences were observed.

**Table 8.** GLSZM-based radiomic features for real and synthetic T2W for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real T2W	Synthetic T2W-CycleGAN	Synthetic T2W-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE	Mean ± SE	Mean ± SE	Real T2W vs. Synthetic T2W-CycleGAN	Real T2W vs. Synthetic T2W-DC <sup>2</sup> Anet
Grey Level Mean	8.92 ± 0.38	9.49 ± 0.48	9.2 ± 0.48	−2.191 to 1.049	−1.905 to 1.336
Grey Level Nonuniformity	1.26 ± 0.02	1.4 ± 0.04	1.38 ± 0.03	−0.2608 to −0.0051 *	−0.2419 to 0.0137
Grey Level Variance	42.32 ± 0.57	43.40 ± 1.47	42.5 ± 1.372	−5.407 to 3.232	−4.504 to 4.134
High Grey Level Emphasis	136.1 ± 4.6	139.1 ± 10.21	137.6 ± 10.27	−34.25 to 28.30	−32.82 to 29.73
Large Zone Emphasis	121.8 ± 8.08	119.9 ± 7.38	121.6 ± 7.41	−25.29 to 29.12	−27.03 to 27.38
Large Zone High Grey Level Emphasis	240.6 ± 4.04	245.6 ± 3.71	249.1 ± 3.56	−18.50 to 8.430	−22.00 to 4.930
Large Zone Low Grey Level Emphasis	152.2 ± 2.93	145.4 ± 5.44	135.1 ± 4.12	−8.474 to 22.10	1.827 to 32.40 *
Low Grey Level Emphasis	0.18 ± 0.01	0.19 ± 0.009	0.18 ± 0.008	−0.0510 to 0.0237	−0.0451 to 0.0296

\* Significant difference.

**Table 9.** GLSZM-based radiomic features for real and synthetic FLAIR for CycleGAN and DC<sup>2</sup>Anet models.

Features	Real FLAIR	Synthetic FLAIR-CycleGAN	Synthetic FLAIR-DC <sup>2</sup> Anet	95% CI of Diff	
	Mean ± SE	Mean ± SE	Mean ± SE	Real FLAIR vs. Synthetic FLAIR-CycleGAN	Real FLAIR vs. Synthetic FLAIR-DC <sup>2</sup> Anet
Grey Level Mean	10.13 ± 0.26	10.12 ± 0.25	9.6 ± 0.19	−0.860 to 0.874	−0.349 to 1.386
Grey Level Nonuniformity	1.34 ± 0.02	1.3 ± 0.05	1.21 ± 0.06	−0.139 to 0.206	−0.047 to 0.298
Grey Level Variance	45.95 ± 2.34	47.5 ± 1.39	46.13 ± 1.24	−7.729 to 4.629	−6.357 to 6.001
High Grey Level Emphasis	152.5 ± 2.33	152.4 ± 2.66	141.3 ± 3.24	−9.836 to 9.940	1.322 to 21.10 *
Large Zone Emphasis	146.6 ± 9.91	142.8 ± 8	143.5 ± 8.1	−27.19 to 34.78	−27.88 to 34.09
Large Zone High Grey Level Emphasis	333.9 ± 8.78	331 ± 6.32	323.4 ± 7.4	−24.08 to 29.90	−16.49 to 37.49
Large Zone Low Grey Level Emphasis	147.8 ± 9.1	146 ± 8.18	119 ± 5.88	−26.05 to 29.57	0.995 to 56.61 *
Low Grey Level Emphasis	0.17 ± 0.007	0.18 ± 0.006	0.18 ± 0.005	−0.035 to 0.011	−0.032 to 0.014

\* Significant difference.

#### 4. Discussion

A within-modality synthesis strategy was presented for Generating FLAIR images from T2W images and vice versa based on CycleGAN and DC<sup>2</sup>Anet networks. Comprehensive evaluations were conducted for two distinct methods where training images were registered within single subjects. It has been shown, via a perceptual study and in terms of quantitative assessments based on MAE, MSE, PSNR metrics, as well as based on a novel radiomic evaluation, that CycleGAN and DC<sup>2</sup>Anet can be used to generate visually realistic MR images. While our synthesis approaches were primarily evaluated for two specific brain MRI sequences, it has the potential to be applied for image-to-image MRI synthesis, as well as for synthesis across imaging modalities (such as MRI, CT, and PET). The proposed CycleGAN technique uses adversarial loss functions and cycle-consistency loss for learning

to synthesize from registered images for improved synthesis. In the DC<sup>2</sup>Anet model, in addition to used losses in the CycleGAN model, four-loss functions, including voxel-wise, gradient difference, perceptual, and structural similarity losses, were used. In modern medical imaging modalities, generating realistic medical images that can be utterly similar to their real ones remains a challenging objective. Generated synthetic images can ensure a trustworthy diagnosis. Based on the quantitative evaluation, for all metrics, the CycleGAN model was found accurate and outperformed the DC<sup>2</sup>Anet model.

CycleGAN and DC<sup>2</sup>Anet models learn the mapping directly from the space of T2W images to the corresponding FLAIR images, and vice versa. Moreover, three metrics including MAE, MSE, and PSNR applied to the data from 510 T2W/FLAIR paired slices from 102 patients, were favorably compared with other reported results in the brain region literature. For example, Krauss et al. [61] compared the assessment results of synthetic and conventional MRIs for patients with Multiple Sclerosis (MS). The images were prospectively acquired for 52 patients with the diagnosed MS. In addition to quantitative evaluations and using the GAN-based approach, the CycleGAN model obtained better results than the study of Krauss et al. Han et al. [62] proposed a GAN-based approach to generate synthetic multi-sequence brain MRI using Deep Convolutional GAN (DCGAN) and Wasserstein GAN (WGAN). Their model was validated by an expert physician who employed Visual Turing test. In agreement with our study, their results revealed that GANs could generate realistic multi-sequence brain MRI images. Nevertheless, this study was different from other research attempts in terms of the employed quantitative evaluation and the proposed models, CycleGAN and DC<sup>2</sup>Anet. In addition to the MAE, MSE, and PSNR metrics, we also conducted a novel evaluation based on radiomic features, to compare the real and synthetic MRI cases. Li et al. [63] proposed a procedure to synthesize brain MRI from CT images. They reported an MAE value of 87.73 and an MSE value of  $1.392 \times 10^4$  for real MRI and synthesized MRI. Although their results differed from our findings, there was a common understanding that the application of the CycleGAN model was subject to much error.

For evaluation, generated images and comparing them with real images, quantitative evaluations can be used. It is clear that even if the model achieves a relatively satisfactory score in quantitative measurements including MAE, MSE, and PSNR metrics, it does not necessarily generate visually realistic images. Although visually CycleGAN produced realistic images there is not much difference between the two models CycleGAN and DC<sup>2</sup>Anet in quantitative measurements including MAE, MSE, and PSNR metrics, as shown in Table 3. It can be implied that the process of determining whether or not an image is visually realistic cannot be done based on the mentioned metrics. However, the current study employed radiomic features as a new evaluation approach to compare the real MRI and their synthetic counterparts. Our results (Tables 4–9) revealed that for the vast majority of radiomic features regarding the two T2W and FLAIR images, no significant difference was observed between real images and images synthesized using CycleGAN. On the other side, some radiomic features indicate significant differences for images synthesized by DC<sup>2</sup>Anet, indicating that the set of radiomic features is more successful in assessing the realism of the generated images than traditional metrics such as MAE, MSE, and PSNR. Therefore, according to the metrics used in this study, it can be concluded that performing evaluations based on radiomic features is a viable option in the GAN models.

In this study, we have used the ACRIN 6677/RTOG 0625 data set which is a multi-center dataset and is one of the strengths of this study. Of note, as a limitation, future studies with a large sample size are suggested. This study considered synthesis for two-contrast brain MRI; hence the proposed models can also be used for other related tasks in medical image analysis such as T1W, CT-PET, and MR-CT. For future research, it is suggested that relevant evaluations be carried out based on radiomic features with larger data sets and other anatomical areas.



Despite the demonstrated effectiveness of our method in generating a T2W MRI from a FLAIR, and vice versa, it is important to acknowledge that the applicability of our approach may have limitations in certain specific cases. The performance of our method may be influenced by factors such as extreme variations in tumor size, irregular tumor shapes, or cases with substantial edema or necrosis. While our methodology has shown promising results in brain tumor patients, further research is needed to investigate its robustness in challenging scenarios and to develop additional techniques to address these limitations. Future studies should also consider expanding the dataset to include a larger cohort of patients with a wider spectrum of brain pathologies to ensure the generalizability of our findings.

## 5. Conclusions

The CycleGAN method can be used to generate realistic synthetic T2W and FLAIR brain scans, supported by both experimental qualitative and quantitative results. Radiomic features, representing quantitative data extracted from radiology images, hold a lot of promise as a novel approach to quantitatively evaluate similarities between synthetic and real MRI scans and make decisions based on the radiologic quality of the synthetic scans. Synthesis of realistic MRI scans can facilitate imaging of uncooperative patients and significantly shorten the image acquisition time thereby contributing to reducing costs to healthcare systems.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15143565/s1>, Supplementary S1: The radiomics features extracted in our study encompassed a total of twenty-three features, which were categorized into three groups: Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone (GLSZM), and Gray Level Run Length Matrix (GLRLM). The GLCM features included Auto Correlation, Cluster Prominence, Cluster Shade, Contrast, Correlation, Energy, Entropy, and Homogeneity. The GLSZM features comprised Grey Level Mean, Grey Level Nonuniformity, Grey Level Variance, Large Zone Emphasis, Large Zone High Grey Level Emphasis, Large Zone Low Grey Level Emphasis, Low Grey Level Emphasis, and Grey Level Nonuniformity. Lastly, the GLRLM features consisted of High Grey Level Run Emphasis, Long Run High Grey Level Emphasis, Long Run Low Grey Level Emphasis, Low Gray Level Run Emphasis, Run Length Nonuniformity, and Short Run Emphasis; Supplementary S2: Figure S1: Qualitative comparison of CycleGAN model. There are five examples of T2W and FLAIR synthesis based on two mappings: Generating FLAIR Images from T2W Images (a), and Generating T2W Images from FLAIR Images (b); Figure S2: Qualitative comparison of the DC<sup>2</sup>Anet model. There are five examples of T2W and FLAIR synthesis based on two mappings: Generating FLAIR Images from T2W Images (a), and Generating T2W Images from FLAIR Images (b).

**Author Contributions:** Conceptualization, S.M.R. and N.C.; methodology, S.M.R. and N.C.; software, S.M.R., N.C. and F.B.N.; validation, S.M.R. and N.C.; formal analysis S.M.R. and N.C.; investigation, S.M.R. and N.C.; resources, S.M.R. and N.C.; data curation, S.M.R. and N.C.; writing—original draft preparation, S.M.R., N.C., F.B.N., D.M. and S.B.; writing—review and editing; supervision, D.M. and S.B.; funding acquisition, S.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research reported in this publication was partly supported by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) under award numbers: NIH/NCI:U24CA189523, NIH/NCI:U01CA242871. The content of this publication is solely the responsibility of the authors and does not represent the official views of the NIH.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Ethical review and approval were waived for this study due to the information is from a public dataset.

**Data Availability Statement:** The data presented in this study are available in this article.

**Acknowledgments:** The completion of this research project would not have been possible without the contributions and support of many individuals and organizations. We are deeply grateful to Jundishapur University of Medical Sciences, Kingston University and the University of Pennsylvania as well as the National Cancer Institute (NCI) of the National Institutes of Health (NIH).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Faro, S.H.; Mohamed, F.B. *Functional MRI: Basic Principles and Clinical Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.
2. Bakas, S.; Reyes, M.; Jakab, A.; Bauer, S.; Rempfler, M.; Crimi, A.; Takeshi Shinohara, R.; Berger, C.; Min Ha, S.; Rozycki, M.; et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv* **2018**, arXiv:181102629.
3. Shukla, G.; Alexander, G.S.; Bakas, S.; Nikam, R.; Talekar, K.; Palmer, J.D.; Shi, W. Advanced magnetic resonance imaging in glioblastoma: A review. *Chin. Clin. Oncol.* **2017**, *6*, 40. [[CrossRef](#)] [[PubMed](#)]
4. Bruzzone, M.G.; D'Incerti, L.; Farina, L.L.; Cuccarini, V.; Finocchiaro, G. CT and MRI of brain tumors. *Q. J. Nucl. Med. Mol. Imaging* **2012**, *56*, 112–137. [[PubMed](#)]
5. Rohde, G.K.; Barnett, A.S.; Basser, P.J.; Marengo, S.; Pierpaoli, C. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magn. Reson. Med.* **2004**, *51*, 103–114. [[CrossRef](#)]
6. Aoki, J.; Kimura, K.; Iguchi, Y.; Shibazaki, K.; Sakai, K.; Iwanaga, T. FLAIR can estimate the onset time in acute ischemic stroke patients. *J. Neurol. Sci.* **2010**, *293*, 39–44. [[CrossRef](#)]
7. Krupa, K.; Bekiesińska-Figatowska, M. Artifacts in magnetic resonance imaging. *Pol. J. Radiol.* **2015**, *80*, 93.
8. Hiasa, Y.; Otake, Y.; Takao, M.; Matsuoka, T.; Takashima, K.; Carass, A.; Prince, J.L.; Sugano, N.; Sato, Y. Cross-modality image synthesis from unpaired data using CycleGAN. In Proceedings of the Simulation and Synthesis in Medical Imaging: Third International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 16 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 31–41.
9. Nie, D.; Trullo, R.; Lian, J.; Wang, L.; Petitjean, C.; Ruan, S.; Wang, Q.; Shen, D. Medical Image Synthesis with Deep Convolutional Adversarial Networks. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 2720–2730. [[CrossRef](#)]
10. Xie, C.; Wang, J.; Zhang, Z.; Zhou, Y.; Xie, L.; Yuille, A. Adversarial Examples for Semantic Segmentation and Object Detection. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 1378–1387. [[CrossRef](#)]
11. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. *arXiv* **2014**, arXiv:14126572.
12. Xiang, L.; Li, Y.; Lin, W.; Wang, Q.; Shen, D. Unpaired deep cross-modality synthesis with fast training. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 155–164.
13. Van, N.H.; Zhou, K.; Vemulapalli, R. Cross-domain synthesis of medical images using efficient location-sensitive deep network. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 677–684.
14. Abdullah, M.A.M.; Alkassar, S.; Jebur, B.; Chambers, J. LBTS-Net: A fast and accurate CNN model for brain tumour segmentation. *Health Technol. Lett.* **2021**, *8*, 31–36. [[CrossRef](#)]
15. Lan, L.; You, L.; Zhang, Z.; Fan, Z.; Zhao, W.; Zeng, N.; Chen, Y.; Zhou, X. Generative Adversarial Networks and Its Applications in Biomedical Informatics. *Front. Public Health* **2020**, *8*, 164. [[CrossRef](#)]
16. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27 (NIPS 2014)*; Neural Information Processing Systems Foundation, Inc.: La Jolla, CA, USA, 2014.
17. 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, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
18. Nie, D.; Trullo, R.; Lian, J.; Petitjean, C.; Ruan, S.; Wang, Q.; Shen, D. Medical image synthesis with context-aware generative adversarial networks. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, 11–13 September 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 417–425.
19. Welander, P.; Karlsson, S.; Eklund, A. Generative adversarial networks for image-to-image translation on multi-contrast MR images—A comparison of CycleGAN and UNIT. *arXiv* **2018**, arXiv:180607777.
20. Richardson, E.; Weiss, Y. The Surprising Effectiveness of Linear Unsupervised Image-to-Image Translation. In Proceedings of the 25th International Conference on Pattern Recognition, Milano, Italy, 10–15 January 2021; pp. 7855–7861. [[CrossRef](#)]
21. Dar, S.U.; Yurt, M.; Karacan, L.; Erdem, A.; Erdem, E.; Cukur, T. Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks. *IEEE Trans. Med. Imaging* **2019**, *38*, 2375–2388. [[CrossRef](#)] [[PubMed](#)]
22. Theis, L.; van den Oord, A.; Bethge, M. A note on the evaluation of generative models. *arXiv* **2015**, arXiv:151101844.
23. Gui, J.; Sun, Z.; Wen, Y.; Tao, D.; Ye, J. A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 3313–3332. [[CrossRef](#)]

24. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)] [[PubMed](#)]
25. McNitt-Gray, M.; Napel, S.; Jaggi, A.; Mattonen, S.; Hadjiiski, L.; Muzi, M.; Goldgof, D.; Balagurunathan, Y.; Pierce, L.; Kinahan, P.; et al. Standardization in Quantitative Imaging: A Multicenter Comparison of Radiomic Features from Different Software Packages on Digital Reference Objects and Patient Data Sets. *Tomography* **2020**, *6*, 118–128. [[CrossRef](#)]
26. Parekh, V.S.; Jacobs, M.A. Deep learning and radiomics in precision medicine. *Expert Rev. Precis. Med. Drug Dev.* **2019**, *4*, 59–72. [[CrossRef](#)]
27. Mayerhoefer, M.E.; Materka, A.; Langs, G.; Häggström, I.; Szczypiński, P.; Gibbs, P.; Cook, G. Introduction to radiomics. *J. Nucl. Med.* **2020**, *61*, 488–495. [[CrossRef](#)]
28. Rizzo, S.; Botta, F.; Raimondi, S.; Origgi, D.; Fanciullo, C.; Morganti, A.G.; Bellomi, M. Radiomics: The facts and the challenges of image analysis. *Eur. Radiol. Exp.* **2018**, *2*, 36. [[CrossRef](#)]
29. Lv, W.; Yuan, Q.; Wang, Q.; Ma, J.; Jiang, J.; Yang, W.; Feng, Q.; Chen, W.; Rahmim, A.; Lu, L. Robustness versus disease differentiation when varying parameter settings in radiomics features: Application to nasopharyngeal PET/CT. *Eur. Radiol.* **2018**, *28*, 3245–3254. [[CrossRef](#)]
30. Lei, M.; Varghese, B.; Hwang, D.; Cen, S.; Lei, X.; Azadikhah, A.; Desai, B.; Oberai, A.; Duddalwar, V. Benchmarking features from different radiomics toolkits/toolboxes using Image Biomarkers Standardization Initiative. *arXiv* **2020**, arXiv:200612761.
31. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.G.P.M.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **2012**, *48*, 441–446. [[CrossRef](#)]
32. Avanzo, M.; Stancanello, J.; El Naqa, I. Beyond imaging: The promise of radiomics. *Phys. Med.* **2017**, *38*, 122–139. [[CrossRef](#)] [[PubMed](#)]
33. Aerts, H.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Data from: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*, 4006. [[CrossRef](#)] [[PubMed](#)]
34. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)] [[PubMed](#)]
35. Binder, Z.A.; Thorne, A.H.; Bakas, S.; Wileyto, E.P.; Bilello, M.; Akbari, H.; Rathore, S.; Ha, S.M.; Zhang, L.; Ferguson, C.J.; et al. Epidermal Growth Factor Receptor Extracellular Domain Mutations in Glioblastoma Present Opportunities for Clinical Imaging and Therapeutic Development. *Cancer Cell* **2018**, *34*, 163–177.e7. [[CrossRef](#)] [[PubMed](#)]
36. Fathi Kazerooni, A.; Bakas, S.; Saligheh Rad, H.; Davatzikos, C. Imaging signatures of glioblastoma molecular characteristics: A radiogenomics review. *J. Magn. Reson. Imaging* **2020**, *52*, 54–69. [[CrossRef](#)]
37. Bakas, S.; Akbari, H.; Pisapia, J.; Martinez-Lage, M.; Rozycki, M.; Rathore, S.; Dahmane, N.; O'Rourke, D.M.; Davatzikos, C. In vivo detection of EGFRvIII in glioblastoma via perfusion magnetic resonance imaging signature consistent with deep peritumoral infiltration: The  $\phi$ -index. *Clin. Cancer Res.* **2017**, *23*, 4724–4734. [[CrossRef](#)]
38. Bakas, S.; Shukla, G.; Akbari, H.; Erus, G.; Sotiras, A.; Rathore, S.; Sako, C.; Ha, S.M.; Rozycki, M.; Shinohara, R.T.; et al. Overall survival prediction in glioblastoma patients using structural magnetic resonance imaging (MRI): Advanced radiomic features may compensate for lack of advanced MRI modalities. *J. Med. Imaging* **2020**, *7*, 031505. [[CrossRef](#)]
39. Van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J.W.L. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)]
40. Prasanna, P.; Tiwari, P.; Madabhushi, A. Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe): A new radiomics descriptor. *Sci. Rep.* **2016**, *6*, 37241. [[CrossRef](#)] [[PubMed](#)]
41. Davatzikos, C.; Rathore, S.; Bakas, S.; Pati, S.; Bergman, M.; Kalarot, R.; Sridharan, P.; Gastounioti, A.; Jahani, N.; Cohen, E.; et al. Cancer imaging phenomics toolkit: Quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome. *J. Med. Imaging* **2018**, *5*, 011018. [[CrossRef](#)] [[PubMed](#)]
42. Jin, C.-B.; Kim, H.; Liu, M.; Han, I.H.; Lee, J.I.; Lee, J.H.; Joo, S.; Park, E.; Ahn, Y.S.; Cui, X. DC<sup>2</sup>Anet: Generating Lumbar Spine MR Images from CT Scan Data Based on Semi-Supervised Learning. *Appl. Sci.* **2019**, *9*, 2521. [[CrossRef](#)]
43. Larue, R.T.H.M.; Defraene, G.; De Ruysscher, D.; Lambin, P.; van Elmpt, W. Quantitative radiomics studies for tissue characterization: A review of technology and methodological procedures. *Br. J. Radiol.* **2017**, *90*, 20160665. [[CrossRef](#)]
44. Kinahan, P.; Muzi, M.; Bialecki, B.; Herman, B.; Coombs, L. Data from the ACRIN 6668 Trial NSCLC-FDG-PET. *Cancer Imaging Arch.* **2019**, *10*. [[CrossRef](#)]
45. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)]
46. Yushkevich, P.A.; Piven, J.; Hazlett, H.C.; Smith, R.G.; Ho, S.; Gee, J.C.; Gerig, G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **2006**, *31*, 1116–1128. [[CrossRef](#)]
47. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.

48. Bi, L.; Kim, J.; Kumar, A.; Feng, D.; Fulham, M. Synthesis of Positron Emission Tomography (PET) Images via Multi-channel Generative Adversarial Networks (GANs). In Proceedings of the Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment: Fifth International Workshop, CMMI 2017, Second International Workshop, RAMBO 2017, and First International Workshop, SWITCH 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, 14 September 2017; pp. 43–51. [\[CrossRef\]](#)
49. Nie, D.; Cao, X.; Gao, Y.; Wang, L.; Shen, D. Estimating CT Image from MRI Data Using 3D Fully Convolutional Networks. In *Design, User Experience, and Usability: Design Thinking and Methods*; Springer: Cham, Switzerland, 2016; pp. 170–178. [\[CrossRef\]](#)
50. Jin, C.-B.; Kim, H.; Liu, M.; Jung, W.; Joo, S.; Park, E.; Ahn, Y.S.; Han, I.H.; Lee, J.I.; Cui, X. Deep CT to MR Synthesis Using Paired and Unpaired Data. *Sensors* **2019**, *19*, 2361. [\[CrossRef\]](#)
51. Hemalatha, R.; Thamizhvani, T.; Dhivya, A.J.A.; Joseph, J.E.; Babu, B.; Chandrasekaran, R. Active Contour Based Segmentation Techniques for Medical Image Analysis. In *Medical and Biological Image Analysis*; IntechOpen: London, UK, 2018; p. 2. [\[CrossRef\]](#)
52. Heydarheydari, S.; Firoozabadi, S.M.; Mirnajafi-Zadeh, J.; Shankayi, Z. Pulsed high magnetic field-induced reversible blood-brain barrier permeability to enhance brain-targeted drug delivery. *Electromagn. Biol. Med.* **2021**, *40*, 361–374. [\[CrossRef\]](#)
53. Pati, S.; Singh, A.; Rathore, S.; Gastounioti, A.; Bergman, M.; Ngo, P.; Ha, S.M.; Bounias, D.; Minock, J.; Murphy, G.; et al. The Cancer Imaging Phenomics Toolkit (CaPTk): Technical Overview. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, 17 October 2019; Volume 11993, pp. 380–394. [\[CrossRef\]](#)
54. Galloway, M.M. Texture analysis using gray level run lengths. *Comput. Graph. Image Process.* **1975**, *4*, 172–179. [\[CrossRef\]](#)
55. Rathore, S.; Bakas, S.; Pati, S.; Akbari, H.; Kalarot, R.; Sridharan, P.; Rozycki, M.; Bergman, M.; Tunc, B.; Verma, R.; et al. Brain cancer imaging phenomics toolkit (brain-CaPTk): An interactive platform for quantitative analysis of glioblastoma. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, 14 September 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 133–145.
56. Mohanaiah, P.; Sathyanarayana, P.; GuruKumar, L. Image texture feature extraction using GLCM approach. *Int. J. Res. Publ.* **2013**, *3*, 1–5.
57. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [\[CrossRef\]](#)
58. Thibault, G.; Angulo, J.; Meyer, F. Advanced statistical matrices for texture characterization: Application to DNA chromatin and microtubule network classification. In Proceedings of the 18th IEEE International Conference on Image Processing, Brussels, Belgium, 11–14 September 2011; pp. 53–56. [\[CrossRef\]](#)
59. Duron, L.; Balvay, D.; Perre, S.V.; Bouchouicha, A.; Savatovsky, J.; Sadik, J.-C.; Thomassin-Naggara, I.; Fournier, L.; Lecler, A. Gray-level discretization impacts reproducible MRI radiomics texture features. *PLoS ONE* **2019**, *14*, e0213459. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Yap, B.W.; Sim, C.H. Comparisons of various types of normality tests. *J. Stat. Comput. Simul.* **2011**, *81*, 2141–2155. [\[CrossRef\]](#)
61. Krauss, W.; Gunnarsson, M.; Nilsson, M.; Thunberg, P. Conventional and synthetic MRI in multiple sclerosis: A comparative study. *Eur. Radiol.* **2017**, *28*, 1692–1700. [\[CrossRef\]](#)
62. Han, C.; Hayashi, H.; Rundo, L.; Araki, R.; Shimoda, W.; Muramatsu, S.; Furukawa, Y.; Mauri, G.; Nakayama, H. GAN-based synthetic brain MR image generation. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 734–738.
63. Li, W.; Li, Y.; Qin, W.; Liang, X.; Xu, J.; Xiong, J.; Xie, Y. Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy. *Quant. Imaging Med. Surg.* **2020**, *10*, 1223–1236. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.