

Article

Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement in MRI Imaging

Ka-Hei Cheng ¹, Wen Li ¹ , Francis Kar-Ho Lee ², Tian Li ¹ and Jing Cai ^{1,3,*} 

¹ Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China; ka-hei.cheng@connect.polyu.hk (K.-H.C.); polyu-wen.li@polyu.edu.hk (W.L.); litian.li@polyu.edu.hk (T.L.)

² Department of Clinical Oncology, Queen Elizabeth Hospital, Hong Kong SAR, China; leekh4@ha.org.hk

³ The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518000, China

* Correspondence: jing.cai@polyu.edu.hk; Tel.: +852-3400-8645

Simple Summary: This paper presents a novel approach to produce virtual contrast enhanced (VCE) images for nasopharyngeal cancer (NPC) without the use of contrast agents, which carry certain risks. This model uses pixelwise gradient term to capture the shape and a GAN terms to capture the texture of the real contrast enhanced T1C images. With similar accuracy to existing models, our method shows an advantage in reproducing texture closer to the realistic contrast-enhanced images. This results are tested by various measures, including mean absolute error (MAE), mean square error (MSE) and structural similarity index (SSIM) for similarity accuracy; total mean square variation per mean intensity (TMSVPMI), the total absolute variation per mean intensity (TAVPMI), Tenengrad function per mean intensity (TFPMI) and variance function per mean intensity (VFPMI). Various variations of the model, including fine-tuning of the hyperparameters, normalization methods on the images and using single modality, have also been investigated to test the optimal performance.

Abstract: Background: The development of advanced computational models for medical imaging is crucial for improving diagnostic accuracy in healthcare. This paper introduces a novel approach for virtual contrast enhancement (VCE) in magnetic resonance imaging (MRI), particularly focusing on nasopharyngeal cancer (NPC). **Methods:** The proposed model, Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE), makes use of pixelwise gradient methods with Generative Adversarial Networks (GANs) to enhance T1-weighted (T1-w) and T2-weighted (T2-w) MRI images. This approach combines the benefits of both modalities to simulate the effects of gadolinium-based contrast agents, thereby reducing associated risks. Various modifications of PGMGVCE, including changing hyperparameters, using normalization methods (z-score, Sigmoid and Tanh) and training the model with T1-w or T2-w images only, were tested to optimize the model's performance. **Results:** PGMGVCE demonstrated a similar accuracy to the existing model in terms of mean absolute error (MAE) (8.56 ± 0.45 for Li's model; 8.72 ± 0.48 for PGMGVCE), mean square error (MSE) (12.43 ± 0.67 for Li's model; 12.81 ± 0.73 for PGMGVCE) and structural similarity index (SSIM) (0.71 ± 0.08 for Li's model; 0.73 ± 0.12 for PGMGVCE). However, it showed improvements in texture representation, as indicated by total mean square variation per mean intensity (TMSVPMI) (0.124 ± 0.022 for ground truth; 0.079 ± 0.024 for Li's model; 0.120 ± 0.027 for PGMGVCE), total absolute variation per mean intensity (TAVPMI) (0.159 ± 0.031 for ground truth; 0.100 ± 0.032 for Li's model; 0.153 ± 0.029 for PGMGVCE), Tenengrad function per mean intensity (TFPMI) (1.222 ± 0.241 for ground truth; 0.981 ± 0.213 for Li's model; 1.194 ± 0.223 for PGMGVCE) and variance function per mean intensity (VFPMI) (0.0811 ± 0.005 for ground truth; 0.0667 ± 0.006 for Li's model; 0.0761 ± 0.006 for PGMGVCE). **Conclusions:** PGMGVCE presents an innovative and safe approach to VCE in MRI, demonstrating the power of deep learning in enhancing medical imaging. This model paves the way for more accurate and risk-free diagnostic tools in medical imaging.



Citation: Cheng, K.-H.; Li, W.; Lee, F.K.-H.; Li, T.; Cai, J. Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement in MRI Imaging. *Cancers* **2024**, *16*, 999. <https://doi.org/10.3390/cancers16050999>

Academic Editor: Edward J. Pavlik

Received: 29 January 2024

Revised: 24 February 2024

Accepted: 27 February 2024

Published: 29 February 2024



Copyright: © 2024 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/>).

Keywords: virtual contrast enhancement; tumor contrast; MR-guided radiotherapy; nasopharyngeal carcinoma

1. Introduction

In the rapidly advancing field of medical imaging, the refinement of sophisticated models for image analysis plays a pivotal role in boosting diagnostic precision and patient care. Magnetic resonance imaging (MRI), among various imaging techniques, demonstrates its prowess in delivering high-resolution imagery of soft tissues, notably without employing ionizing radiation. T1-weighted (T1-w) and T2-weighted (T2-w) images, in particular, stand as cornerstones in the diagnosis of diverse medical conditions [1,2].

Virtual-contrast-enhanced T1 images (VCE T1C images) represent a novel technique designed to emulate the visibility of certain tissues and fluids typically accentuated by contrast agents. In clinical practice, these agents are frequently utilized to amplify the distinction between normal and abnormal tissues, especially in the brain, aiding in the delineation of tumors, inflammation and other pathologies. For instance, in cases of nasopharyngeal cancer (NPC), gadolinium-based contrast agents are administered to enhance tumor visibility. However, the usage of these agents is not without associated risks [3–13]. Virtual contrast enhancement, thus, emerges as a safer alternative, mimicking the effects of contrast agents through deep learning (DL) applications. The incorporation of DL in image synthesis has recently attracted considerable attention in the field of medical imaging. Its potential in discerning complex tumor characteristics [14–16] has spurred research into the synthesis of virtual-contrast-enhanced MRI (VCE-MRI) from non-contrast scans, particularly for brain cancer patients.

The development of VCE-MRI involves training DL algorithms with large datasets of MRI scans, both with and without contrast agents. By learning the patterns and characteristics of gadolinium-enhanced images, these algorithms can generate virtual-contrast-enhanced images from standard T1-w and T2-w scans. This process not only obviates the need for contrast agents but also has the potential to reduce scan time and costs associated with the use of these agents.

The medical field has recently spotlighted the advancement of DL in generating synthesized images [17–23]. Deep neural networks' ability to dissect and understand the intricate details of tumor characteristics has led to the innovation of creating VCE-MRI images from non-contrast MRI scans for brain cancer patients [7,24]. Specifically, Gong et al. utilized a U-shaped DL model to merge MRI images without a gadolinium-based contrast agent and with a low dose of a gadolinium-based contrast agent, achieving VCE-MRI images that mimic those produced with a full dose of a gadolinium-based contrast agent. This study showcased the potential of DL in extracting contrast enhancement details from full-dose gadolinium-based contrast agent MRI images and generating vceT1w MRI images of satisfactory quality. Building on this groundwork, a three-dimensional Bayesian neural network that integrates ten different MRI techniques to produce VCE-MRI images was introduced [7]. This confirmed the DL network's capacity to utilize various non-contrast imaging methods for synthesizing images. Despite these encouraging results, the current DL models face challenges in harnessing the full potential of the diverse information available from different imaging inputs. This limitation's impact becomes more significant when diagnosing deeply infiltrative NPC, due to the complex interaction of pixel intensity across different imaging modalities [25].

Our model, Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE), employs pixelwise gradient methods to delineate the shape of VCE images, complemented by the use of a Generative Adversarial Network (GAN) to replicate the image contrast of VCE. Pixelwise gradient originated from image registration [26–28], and, to our knowledge, our work is the first to apply it to image synthesis, in particular, VCE-MRI. The evaluation of our models encompasses not only quantitative accuracy metrics, such

as mean absolute error (MAE), mean square error (MSE) and structural similarity (SSIM), but also qualitative assessments of texture. It was observed that the VCE images generated using the model in [22] exhibit excessively smooth textures compared to actual T1C images. The novelty of PGMGVCE, based on pixelwise gradient techniques, is that it demonstrates a texture in VCE images more akin to realistic T1C images. This is evidenced by various metrics introduced in this paper, such as the total mean square variation per mean intensity (TMSVPMI), total absolute variation per mean intensity (TAVPMI), Tenengrad function per mean intensity (TFPMI) and variance function per mean intensity (VFPMI). Despite similar mean absolute errors between images produced by PGMGVCE and the model in [22] when compared with ground truth, the improved textural fidelity of PGMGVCE images suggests its superiority over the model in [22].

Sections 2.1 and 2.2 introduce the model architectures of the PGMGVCE model. Section 2.3 introduces methods to evaluate the performance of the models. Section 2.4 discusses the data preprocessing steps. Section 3.1 shows the results of the VCE images. Sections 2.5 and 3.2 include comprehensive comparisons of various adaptations of the PGMGVCE model. These comparisons encompass modifications in hyperparameters, the application of different image normalization techniques and training the models exclusively with either T1-w or T2-w images, thereby enriching the study with a thorough analytical perspective. Sections 4 and 5 are the discussion and conclusions, respectively.

2. Methods

2.1. Model Architecture of Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE)

The PGMGVCE architecture is depicted in Figure 1. This model is trained using the pixelwise gradient method in conjunction with a Generative Adversarial Network (GAN), enabling it to accurately replicate the shape and contrast of input images. The gradient method, initially developed for image registration [26–28] in the context of image registration, was adapted for VCE. This adaptation begins by calculating the gradient of an image:

$$\nabla x(i, j) = (x(i + 1, j) - x(i, j), x(i, j + 1) - x(i, j)) \quad (1)$$

where $x(i, j)$ represents the (i, j) pixel of image. To capture the shape of the input images, we consider normalized gradients $n = \frac{\nabla x(i, j)}{\|\nabla x(i, j)\| + \epsilon}$, where ϵ is a small constant to avoid division of zero. If the output images capture the same shape as the input images, the gradient will point to the same or opposite directions, as the gradient of pixel intensity is a geometric quantity. The alignment of the gradient with the input images' shape is measured by the square of the dot product between the input and output normalized gradients, forming the basis of our loss function.

$$\text{Pixelwise Gradient}(y_{\text{output}}, z_{\text{ground truth}}) = -\sum_{\text{pixel}} (n_{\text{output}} \cdot n_{\text{ground truth}})^2 \quad (2)$$

where $y_{\text{output}}, z_{\text{ground truth}}$ are the output and ground-truth real T1C images, respectively.

To learn the image contrast, we employ GAN loss:

$$\text{GAN Loss} = \text{GAN}(\text{output image}, \text{ground truth T1C image}) \quad (3)$$

utilizing a Least Squares Generative Adversarial Network (LSGAN).

2.2. Implementation of the PGMGVCE

The model accepts T1-weighted and T2-weighted MRI slices as input. A series of convolutional layers are used for initial feature extraction. These layers progressively down-sample the image while increasing the feature map depth. Then, a module is designed to effectively integrate information from T1-w and T2-w images. It extracts and combines features from each modality, leveraging their complementary nature to enhance contrast and detail. Then, the model employs a set of trainable weights that adjust dynamically

during training, optimizing the contribution of each modality based on the fusion target. This guides the fusion process by evaluating and weighting the importance of features from different modalities. This system ensures that the most relevant features for contrast enhancement are prioritized. Then, the output is forwarded to a module that utilizes separate convolutional pathways for each modality, followed by feature fusion layers that intelligently merge the extracted features. This approach ensures that the unique characteristics of each modality are preserved and effectively utilized. This module enhances the model's ability to focus on salient features within the medical images, such as areas indicating pathological changes.

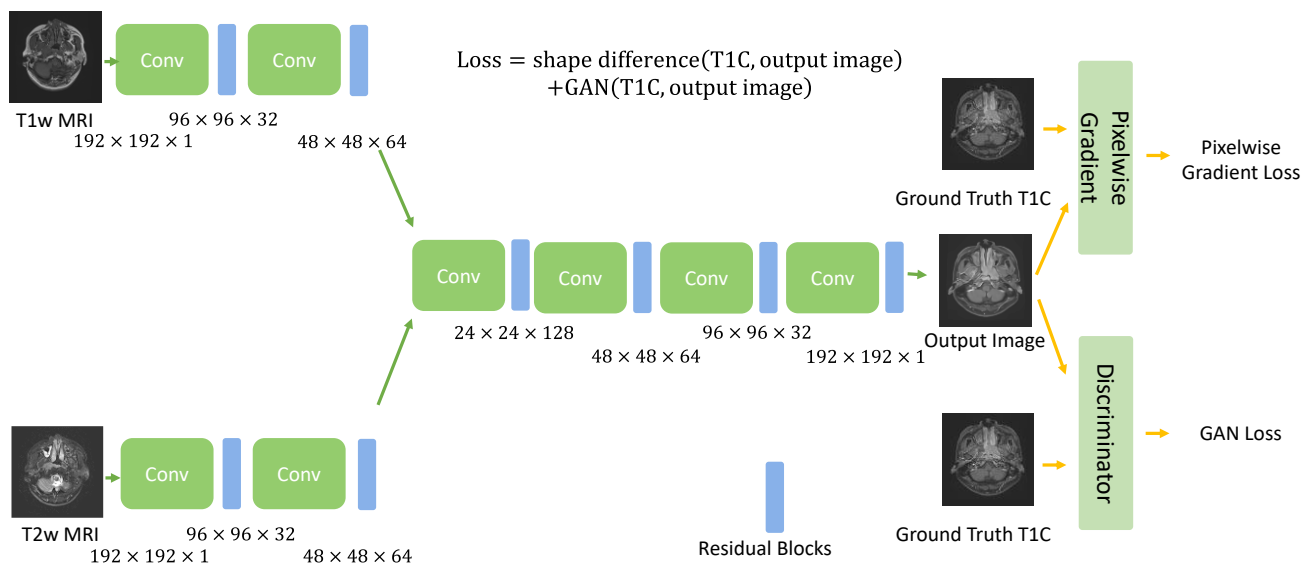


Figure 1. Architecture of the deep learning model for the proposed model, Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE). The multiplications in the figure indicate the dimension and number of channels in that layer.

After that is the discriminator, which incorporates a patch-based approach, where it evaluates the authenticity of different regions of the image separately. This localized assessment enables a more detailed and accurate evaluation of the image quality. This is to improve the realism and diagnostic quality of the fused images.

Each of these components is designed to work in synergy, with the outputs of one module feeding into or influencing others. This integration ensures a cohesive and effective processing pipeline, resulting in high-quality, contrast-enhanced medical images.

The model is trained using an Adam optimizer with a learning rate of 0.0002 and a beta1 value of 0.5. Training is conducted in mini-batches, with each batch containing an equal mix of T1-weighted and T2-weighted images. To ensure stable training, a gradient penalty is applied to the discriminator, encouraging smoother gradients in the generated images. We performed 14,000 iterations for training.

2.3. Evaluation of the Models

The accuracy of the PGMGVCE and the model in [22] are evaluated using mean absolute error (MAE), mean square error (MSE) and structural similarity index (SSIM). The metrics are expressed as follows:

$$\text{AE} = \frac{1}{N} |y(x) - g(x)| \quad (4)$$

$$\text{MSE} = \frac{1}{N} (y(x) - g(x))^2 \quad (5)$$

$$\text{SSIM} = \frac{(2\mu_{y(x)}\mu_{g(x)} + c_1)(2\sigma_{y(x)g(x)} + c_2)}{(\mu_{y(x)}^2 + \mu_{g(x)}^2 + c_1)(\sigma_{y(x)}^2 + \sigma_{g(x)}^2 + c_2)} \quad (6)$$

where N is the number of pixels in each image slice, and $y(x)$ and $g(x)$ denote the synthetic VCE T1C images and the ground truth, respectively. $\mu_{y(x)}$, $\mu_{g(x)}$ and $\sigma_{y(x)}$, $\sigma_{g(x)}$ are the means and variances of the synthetic images and the ground truth, whereas $\sigma_{y(x)g(x)}$ is the covariance of $y(x)$ and $g(x)$. $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ are 2 variables used to stabilize the division by the weak denominator, and L is the dynamic range of the pixel values. Here, $L = 4095$, $k_1 = 0.01$ and $k_2 = 0.03$ were set by default.

To quantify the smoothness of VCE T1C images, four metrics are introduced. It can be used as a measure to check the difference in texture of the images. All metrics divide the mean intensity since relative pixel intensity variations capture the texture of the images. For example, if an image is multiplied by a constant, the total variations will also be multiplied by that constant, but the texture of the image should be invariant under that multiplication, as it is intrinsic to the image.

The first two metrics are the total mean square variation per mean intensity (TMSVPMI) and the total absolute variation per mean intensity (TAVPMI), which are, respectively, defined as

$$\text{TMSVPMI} = \frac{1}{\mu} \frac{1}{\text{Number of pixels}} \sum_{\text{pixels}} \sqrt{(\nabla_x \text{image})^2 + (\nabla_y \text{image})^2} \quad (7)$$

$$\text{TMAVPMI} = \frac{1}{\mu} \frac{1}{\text{Number of pixels}} \sum_{\text{pixels}} (|\nabla_x \text{image}| + |\nabla_y \text{image}|) \quad (8)$$

where μ is the mean intensity of the image.

The third metric is the Tenengrad function per mean intensity (TFPMI) which is based on [29]:

$$\text{TFPMI} = \frac{1}{\mu} \frac{1}{\text{Number of pixels}} \sum_{\text{pixels}} \sqrt{G_x^2 + G_y^2} \quad (9)$$

where μ is the mean intensity of the image, and G_x and G_y are the Sobel operators which are, respectively, the convolutions of the images with the kernels:

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}.$$

The fourth metric, which is also motivated by [29], is the variance function per mean intensity (VFPMI):

$$\text{VFPMI} = \frac{1}{\mu} \frac{1}{\text{Number of pixels}} \sum_{\text{pixels}(x,y)} (I(x,y) - \mu)^2 \quad (10)$$

where I is the image, and μ is the mean intensity of the image.

The smaller the indices, the smoother the image.

2.4. Data Preprocessing

The dataset, approved by the Research Ethics Committee in Hong Kong (reference number: UW21-412), consisted of 80 NPC patients at stages I to IVb, imaged with T1-w, T2-w and T1C MRI using a 3T-Siemens scanner with TR: 620 ms, TE: 9.8 ms; TR: 2500 ms, TE: 74 ms; and TR: 3.42 ms, TE: 1.11 ms, respectively. The average age of the patients was 57.6 ± 8.6 , including 46 males and 34 females. Each patient's 3D image was converted into 2D slices for training. Image alignment was necessary for contrast enhancement; thus, 3D T1-w and T1C MR images were registered with T2-w MR images using 3dSlicer [30] with a B-Spline transform and mutual information as the similarity metric. Then, different 2D slices were extracted from the 3D images. The dataset was randomly divided into

70 patients for training and 10 for testing, resulting in 3051 and 449 2D image slices for each modality, respectively. Figure 2 shows some sample processed T1-w and T2-w images. These slices were resized to 192×192 , a dimension compatible with the convolutional network structure.

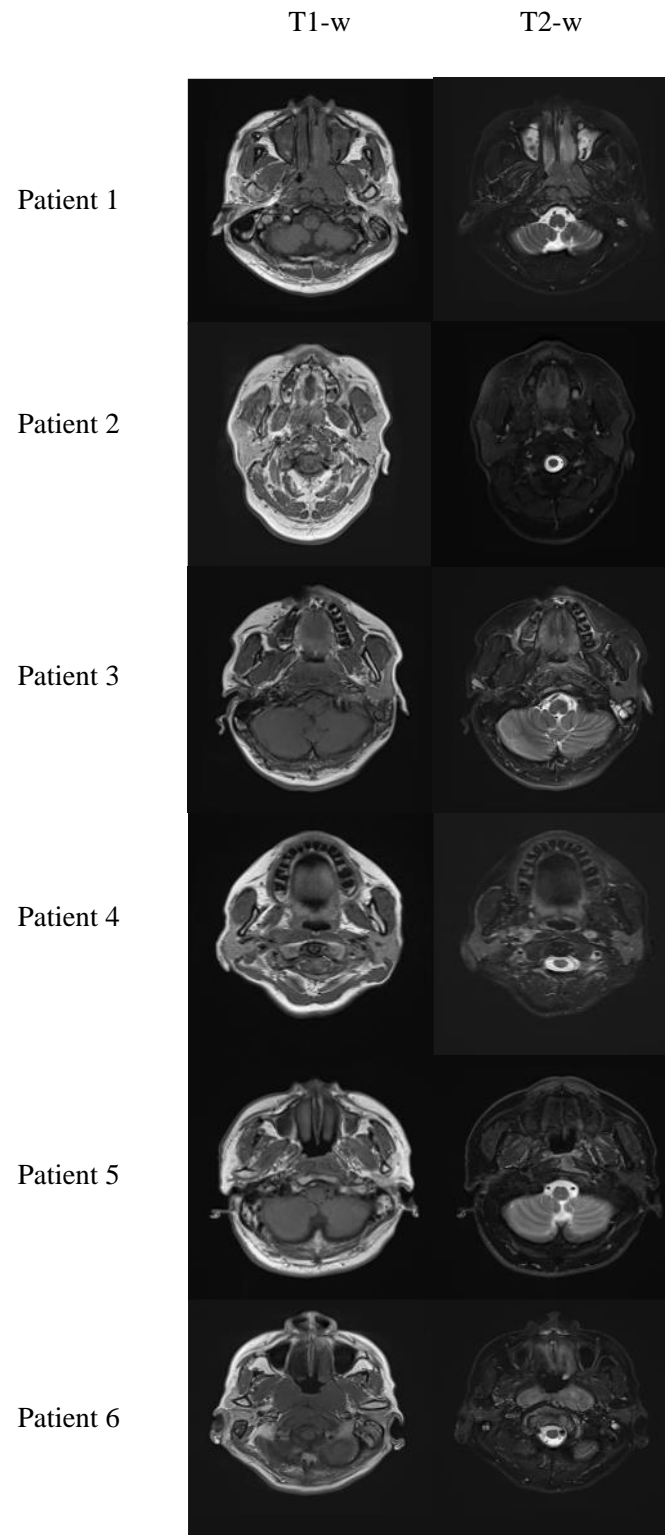


Figure 2. Sample processed T1-w and T2-w images.

2.5. Different Variations of the PGMGVCE

2.5.1. Fine-Tuning the Hyperparameter

We tested various hyperparameter values between pixelwise gradient terms and GAN terms to evaluate the performance. Ratios of 10:1, 1:1 and 1:10 are discussed here.

2.5.2. Different Normalization Methods on Images

Normalization is a crucial preprocessing step to standardize the range of pixel intensity values, thereby enhancing the model's ability to learn and generalize from the data. Different datasets may have different distributions. Some might be Gaussian, while others might have a uniform or skewed distribution. Each normalization method is tailored to work best with a specific type of distribution. Moreover, using the appropriate normalization for a particular data distribution can make the model more robust to variations and outliers, thereby improving its performance and accuracy. Furthermore, when combining features with different scales (e.g., T1-weighted and T2-weighted MRI images), normalization ensures that each feature contributes equally to the analysis and is not dominated by those on larger scales. Normalization can help in emphasizing the importance of smaller-scale features that might be critical for diagnosis in medical images.

Different normalization methods (z-score, Sigmoid and Tanh) [31] are also applied to the images, where

$$x_{z-score} = \frac{x - \mu_x}{\delta_x} \quad (11)$$

$$x_{sigmoid} = \frac{1}{1 + e^{-\frac{x - \mu_x}{\delta_x}}} \quad (12)$$

$$x_{tanh} = \frac{1}{2} \left[\tanh \left(0.01 \left(\frac{x - \mu_x}{\delta_x} \right) \right) + 1 \right] \quad (13)$$

where x represents the intensities of each patient volume, and μ_x and δ_x are the mean value and standard deviation of the patient. $x_{z-score}$, $x_{sigmoid}$ and x_{tanh} represent the corresponding values of patient data after z-normalization, Sigmoid and Tanh normalization methods, respectively.

Z-score normalization, or standard score normalization, involves rescaling the data to have a mean of 0 and a standard deviation of 1. Z-score normalization ensures that each feature contributes equally to the analysis, which is critical when combining features of different scales and units. It enhances the model's sensitivity to outliers, which can be vital for identifying anomalies in medical images. When the underlying data distribution is Gaussian, z-score normalization makes the features more Gaussian-like, which is an assumption in many machine learning models.

Sigmoid normalization transforms data using the Sigmoid function to constrain values within a range of 0 to 1. It bounds the input into a fixed range, which can be beneficial for models that are sensitive to input scale and distribution. The smooth nature of the Sigmoid function provides smooth gradients, which can aid in the convergence during the training of deep learning models. In medical images, this can help preserve the context and relative contrast between different tissue types while standardizing the overall intensity scale.

Tanh normalization is similar to Sigmoid but rescales the data to a range between 0 and 1. Data are centered around 1/2, which can lead to better performance in models where the sign of the data is important. The steeper gradient of Tanh (compared to Sigmoid) around the center can lead to faster learning and convergence in some cases. For medical images, this method can enhance contrast between areas of interest, potentially improving the model's ability to learn and distinguish pathological features.

For the Sigmoid and Tanh normalization, we first transform the image intensity according to (11)–(13) to all the T1-w, T2-w and T1C images. Then, we train and apply the deep learning models to the transformed images. Finally, we perform inverse Sigmoid and inverse Tanh, respectively, to the output synthetic images to obtain images comparable to those of the z-score normalization.

2.5.3. Using Single Modality for Contrast Enhancement

We assessed the performance of using both T1-w and T2-w images versus a single modality for contrast enhancement. The mean absolute error ratio (MAER), mean square error ratio (MSER) and structural similarity ratio (SSIMR) were computed as defined by

$$\text{MAER} = \frac{\text{MAE}_{\text{T1/T2}} - \text{MAE}_{\text{T1 and T2}}}{\text{MAE}_{\text{T1 and T2}}} \quad (14)$$

$$\text{MSER} = \frac{\text{MSE}_{\text{T1/T2}} - \text{MSE}_{\text{T1 and T2}}}{\text{MSE}_{\text{T1 and T2}}} \quad (15)$$

$$\text{SSIMR} = \frac{\text{SSIM}_{\text{T1/T2}} - \text{SSIM}_{\text{T1 and T2}}}{\text{SSIM}_{\text{T1 and T2}}} \quad (16)$$

where $\text{MAE}_{\text{T1/T2}}$, $\text{MSE}_{\text{T1/T2}}$ and $\text{SSIM}_{\text{T1/T2}}$ are, respectively, the MAE, MSE and SSIM between the ground-truth VCE images and the image output of T1/T2 only. $\text{MAE}_{\text{T1 and T2}}$, $\text{MSE}_{\text{T1 and T2}}$ and $\text{SSIM}_{\text{T1 and T2}}$ are, respectively, the MAE, MSE and SSIM between the ground-truth VCE images and the image output of both T1 and T2, comparing the performance of combined modality inputs against single-modality inputs.

3. Results

3.1. Comparison between the Model in [22] and the PGMGVCE

We compared the PGMGVCE with the model in (16) using MAE, MSE and SSIM (Table 1). The comparison statistics between the models and the ground truth are very close to each other. As shown in Figure 3, the VCE images qualitatively also look very similar. These indicate that, in terms of accuracy, the PGMGVCE is similar to that in [22]. However, qualitatively, the texture of the model in [22] appears to be smoother than the ground-truth T1C. This can be illustrated by the texture statistics. The TMSVPMI, TAVPMI, TFPMI and VFPMI of the ground-truth T1C, VCE images produced by the PGMGVCE and model in [22] are shown in Table 2. The p-values of TMSVPMI, TAVPMI, TFPMI and VFPMI of the model by [22] are the same as the ground-truth T1C, that is, 0.002, 0.003, 0.004 and 0.0001, respectively, which is statistically significant to claim that the TMSVPMI, TAVPMI, TFPMI and VFPMI of the PGMGVCE are larger than those of the model in [22]. This is an indication that the texture of the result of the PGMGVCE is closer to the realistic ones.

Table 1. Mean absolute error (MAE), mean square error (MSE) and structural similarity index (SSIM) between the ground-truth real-contrast-enhanced T1C and virtual contrast enhanced (VCE) images produced by the Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE) and model in [22].

	Ground-Truth and VCE Images Produced by the PGMGVCE	Ground Truth and Model in [22]
MAE	8.56 ± 0.45	8.72 ± 0.48
MSE	12.43 ± 0.67	12.81 ± 0.73
SSIM	0.71 ± 0.08	0.73 ± 0.12

3.2. Comparison of Different Variations of the PGMGVCE

3.2.1. Fine-Tuning the Hyperparameter

Figure 4 shows sample images trained with the ratio of the hyperparameter for the pixelwise gradient loss term and GAN loss term of 10:1, 1:1 and 1:10. The MAE and MSE between the images were synthesized with the ratio of pixelwise gradient loss to GAN loss of 10:1 and the ground truth of 10.517 ± 0.247 and 14.008 ± 0.284 , respectively; the ratio of pixelwise gradient loss to GAN loss of 1:1 and the ground truth of 9.904 ± 0.178 and 12.866 ± 0.197 , respectively; and the ratio of pixelwise gradient loss to GAN loss of 1:10

and the ground truth of 10.517 ± 0.231 and 13.811 ± 0.273 , respectively. It demonstrated that the ratio of 1:1 between the pixelwise gradient loss and GAN loss shows slightly better performance than the other two ratios.

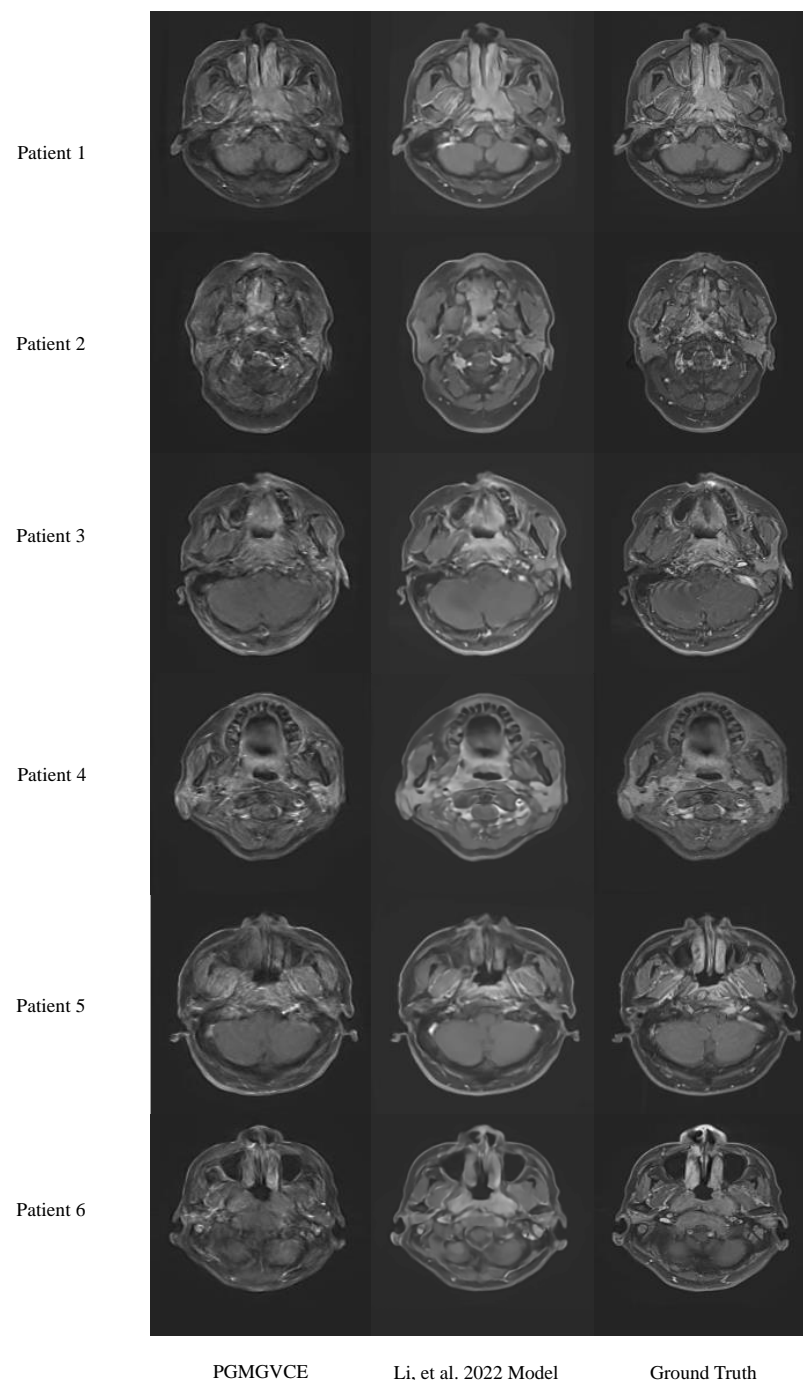


Figure 3. Results of the Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE), the model in [22] and ground truth. It can be seen, qualitatively, that the texture of model in [22] appears to be smoother than the ground truth.

Table 2. The total mean square variation per mean intensity (TMSVPMI), total absolute variation per mean intensity (TAVPMI), Tenengrad function per mean intensity (TFPMI) and variance function per mean intensity (VFPMI) of the ground-truth real-contrast-enhanced T1C and virtual contrast enhanced (VCE) images produced by the Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE) and model in [22].

	Ground Truth	Model in [22]	PGMGVCE
TMSVPMI	0.124 ± 0.022	0.079 ± 0.024	0.120 ± 0.027
TAVPMI	0.159 ± 0.031	0.100 ± 0.032	0.153 ± 0.029
TFPMI	1.222 ± 0.241	0.981 ± 0.213	1.194 ± 0.223
VFPMI	0.0811 ± 0.005	0.0667 ± 0.006	0.0761 ± 0.006

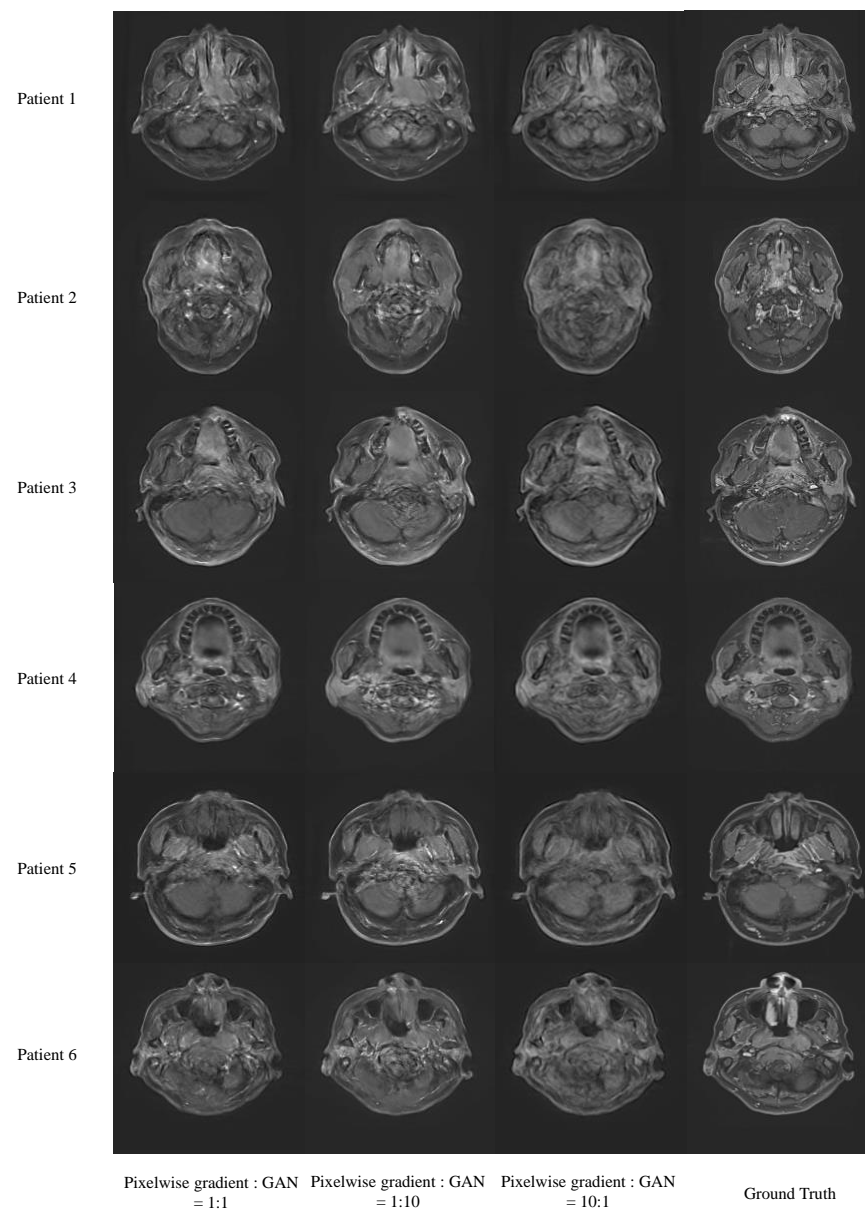


Figure 4. Results of the Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE) with different ratios of hyperparameter of the pixelwise gradient term and the GAN term.

3.2.2. Different Normalization Methods on Images

Figure 5 shows sample images with z-score, Sigmoid and Tanh normalization methods before training. From qualitative inspection, the differences between different normalization methods do not deviate much from each other and the ground truth. For detailed statistics, the MAE and MSE between the image synthesized with z-score normalization and the ground truth are 4.222 ± 0.843 and 4.662 ± 0.918 , respectively; the MAE and MSE between the image synthesized with Sigmoid normalization and the ground truth are 3.814 ± 0.612 and 4.213 ± 0.726 , respectively; and the MAE and MSE between the image synthesized with Tanh normalization and the ground truth are 3.932 ± 0.672 and 4.175 ± 0.892 , respectively. This shows that Sigmoid normalization slightly outperforms the other two.

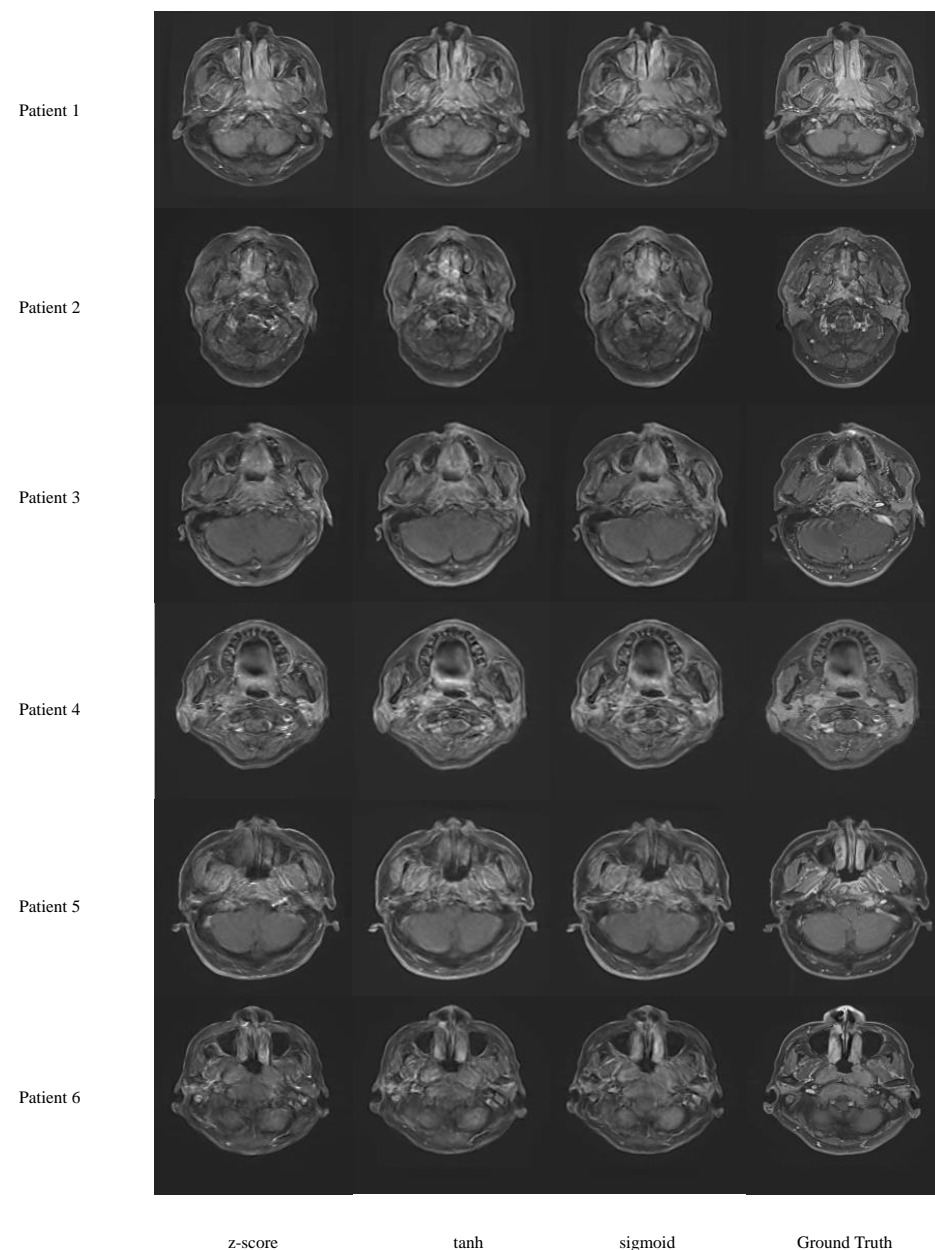


Figure 5. Results of the Pixelwise Gradient Model with GAN for Virtual Contrast Enhancement (PGMGVCE) with different normalization methods of the images.

3.2.3. Using Single Modality for Contrast Enhancement

Figure 6 shows the results of synthesizing VCE images using T1-w, T2-w and both T1-w and T2-w images. The MAER of using T1-w images only is 0.456 ± 0.102 while the MAER of using T2-w images only is 0.389 ± 0.098 . The MSER of using T1-w only is 0.447 ± 0.146 while the MSER of using T2-w images only is 0.413 ± 0.161 . The SSIMR of using T1-w only is 0.283 ± 0.103 while the SSIMR of using T2-w images only is 0.241 ± 0.112 .

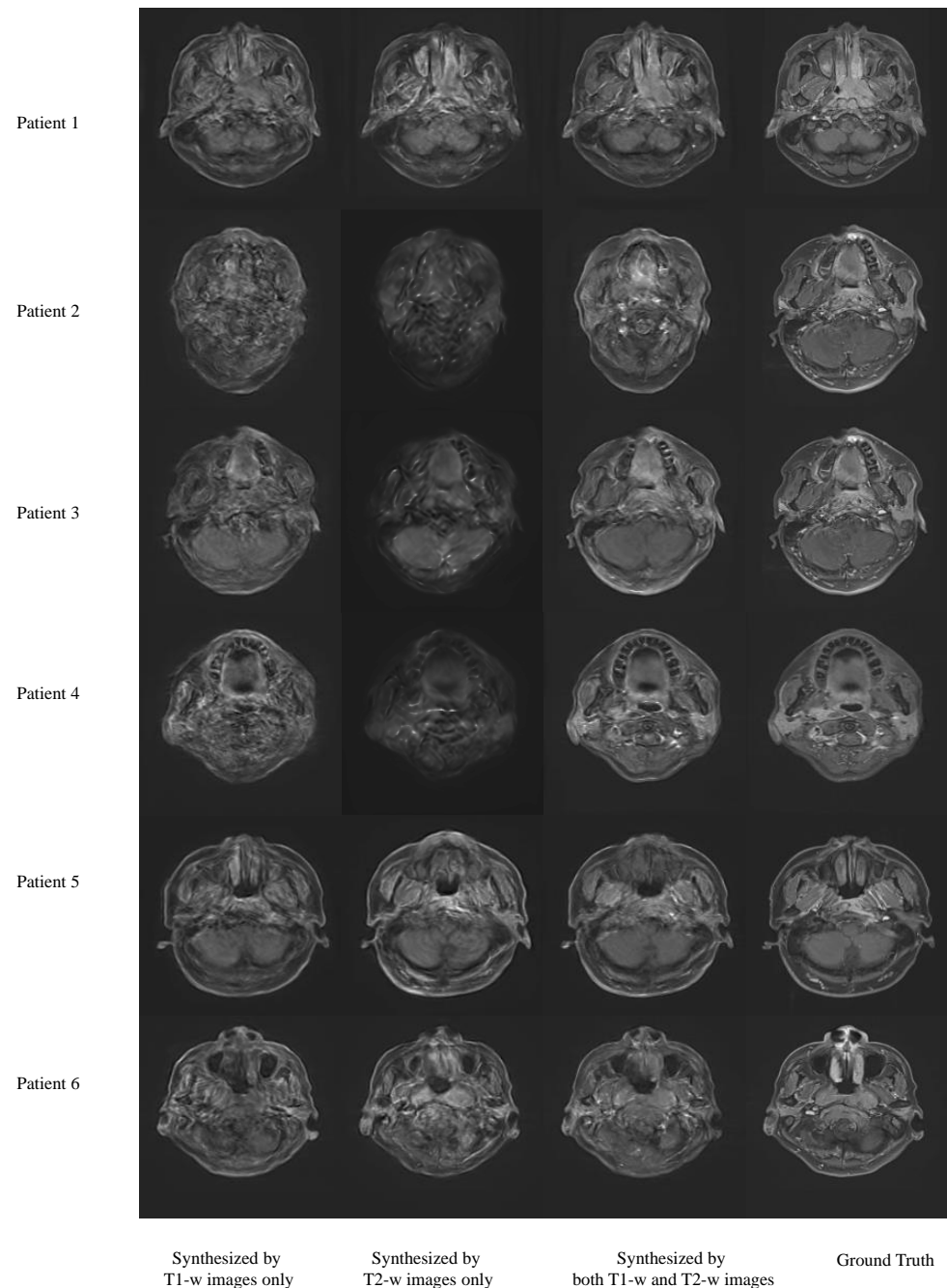


Figure 6. Results of virtual contrast enhanced (VCE) images synthesized using T1-w or T2-w images only and both T1-w and T2-w images.

4. Discussion

Our model's architecture incorporates convolutional layers for initial feature extraction, succeeded by modules that integrate and prioritize features from each imaging modality. Employing a blend of pixelwise gradient methods and GANs, our model captures intri-

cate details from the input images. The gradient method, inspired by image registration techniques [26–28], is adept at detecting subtle variations in the shape and texture characteristics of different tissues and pathologies. By calculating and normalizing the image gradient, the model discerns the geometric structure of tissues, aiding in the high-fidelity reconstruction of enhanced images [32]. GANs, recognized for their capacity to generate lifelike images, are applied here to ensure that the synthesized T1C images are not only structurally precise but visually indistinguishable from actual contrast-enhanced scans. The dynamic interaction between the discriminative and generative elements of GANs compels the model to yield results that fulfill the stringent criteria necessary for clinical diagnosis [33].

Comparative analysis with the model in [22] reveals that, while basic accuracy metrics (MAE, MSE, SSIM) are comparable, the PGMGVCE demonstrates superior texture representation. The metrics of TMSVPML, TAVPML, TFPML and VFPML for the model in [22] are significantly lower than the ground-truth T1C, indicating that its results are overly smooth and may lack critical detail. In contrast, these metrics for the PGMGVCE closely match those of the ground-truth T1C, suggesting a more realistic texture replication. This could be attributed to the PGMGVCE incorporating pixelwise gradient in its loss term, enhancing its ability to capture the authentic texture of T1C images.

Various iterations of the PGMGVCE were examined, particularly the impact of different hyperparameter ratios between pixelwise gradient and GAN components. After extensive trial and error, a 1:1 ratio was identified as optimal. Regarding image normalization methods, Sigmoid normalization was found to be superior, followed by Tanh and z-score normalization. When considering the use of single modalities for VCE image synthesis, it is evident that using both T1-w and T2-w images yields better results than using either modality alone, as the latter only captures partial anatomical information. This conclusion is supported by higher MAER, MSER and SSIMR values when using single modalities.

There are some limitations of our study. The model's performance heavily relies on the quality and diversity of the training data. Additionally, incorporating other MRI modalities or sequences might further amplify the model's diagnostic capabilities. Future investigations should focus on enhancing the model's generalizability by training it with a more varied dataset and could also explore the real-time application of this model in clinical settings to assess its practicality and effectiveness in routine clinical workflows.

5. Conclusions

This study introduces a novel method for VCE in MRI imaging through a deep learning model that effectively combines pixelwise gradient methods with GANs. Our model excels in utilizing the complementary strengths of T1-w and T2-w MRI images, thereby synthesizing T1C images that are visually akin to actual contrast-enhanced scans. The fusion of these imaging modalities is key, as it captures a more exhaustive representation of the anatomy and pathology, thus increasing the diagnostic utility of the images.

In summary, this study presents an innovative approach to virtual contrast enhancement in MRI imaging, leveraging deep learning to reduce the risks associated with contrast agents in VCE images. The ability of our model, the PGMGVCE, to generate images with authentic textures and its potential to offer safer and more precise diagnostics, represents a significant advancement in medical imaging. While the PGMGVCE demonstrates comparable accuracy to the existing model (16), its enhanced texture replication sets it apart, underlining its advantage in realistically capturing VCE images.

The clinical implications of our study are noteworthy. By offering a safer alternative to gadolinium-based contrast agents, the PGMGVCE may diminish the risks linked with contrast-enhanced MRI scans. The improved texture accuracy of the synthesized images could potentially lead to enhanced diagnosis and patient management, particularly in the detection and characterization of NPC.

A limiting aspect of our synthesis network's efficacy is its training solely on T1- and T2-weighted MRI images. It appears these types of images may not encapsulate

all the necessary details for effective contrast synthesis. This issue could potentially be mitigated by incorporating additional MRI techniques (like diffusion-weighted MRI) into our network's input. Furthermore, the model only underwent evaluation using a single dataset. Therefore, its performance and ability to generalize across different scenarios require further examination in subsequent research.

One possible future direction would be performing segmentation on the tumor region to evaluate the performance of the tumor enhancement. Different VCE methods and the ground-truth real T1C can be segmented to compare with each other. This would be a method to test if the tumor contrast of the VCE images is enhanced.

Author Contributions: Conceptualization, K.-H.C.; Methodology, K.-H.C., W.L., T.L. and J.C.; Resources, F.K.-H.L.; Writing—original draft, K.-H.C.; Writing—review and editing, K.-H.C., W.L., T.L. and J.C.; Supervision, T.L. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly supported by the research grants of the Shenzhen Basic Research Program (JCYJ20210324130209023) of the Shenzhen Science and Technology Innovation Committee, Project of Strategic Importance Fund (P0035421) and Projects of RISA (P0043001) from Hong Kong Polytechnic University, Mainland-Hong Kong Joint Funding Scheme (MHKJFS) (MHP/005/20) and Health and Medical Research Fund (HMRF 09200576), the Health Bureau, The Government of the Hong Kong Special Administrative Region.

Institutional Review Board Statement: The data collected in this manuscript are approved by the Research Ethics Committee in Hong Kong on 25/07/021 (reference number: UW21-412).

Informed Consent Statement: Patient consent was waived due to the nature of the retrospective study.

Data Availability Statement: For original data and computer programs, please contact the first author (K.-H.C.) (khcheng9209@gmail.com).

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

References

1. Prince, M.R.; Zhang, H.L.; Prowda, J.C.; Grossman, M.E.; Silvers, D.N. Nephrogenic systemic fibrosis and its impact on abdominal imaging. *Radiographics* **2009**, *29*, 1565–1574. [\[CrossRef\]](#)
2. ACR Committee on MR Safety; Greenberg, T.D.; Hoff, M.N.; Gilk, T.B.; Jackson, E.F.; Kanal, E.; McKinney, A.M.; Och, J.G.; Pedrosa, I.; Rampulla, T.L.; et al. ACR guidance document on MR safe practices: Updates and critical information 2019. *J. Magn. Reson. Imaging* **2020**, *51*, 331–338.
3. Thomsen, H.S. Nephrogenic systemic fibrosis: A serious late adverse reaction to gadodiamide. *Eur. Radiol.* **2006**, *16*, 2619–2621. [\[CrossRef\]](#)
4. Broome, D.R.; Girguis, M.S.; Baron, P.W.; Cottrell, A.C.; Kjellin, I.; Kirk, G.A.; Paulson, E.K.; Kanne, J.P.; Mankoff, D.A.; Baird, G.S.; et al. Gadodiamide-associated nephrogenic systemic fibrosis: Why radiologists should be concerned. *Am. J. Roentgenol.* **2007**, *188*, 586–592. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Kanda, T.; Ishii, K.; Kawaguchi, H.; Kitajima, K.; Takenaka, D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: Relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* **2013**, *270*, 834–841. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Nguyen, N.C.; Molnar, T.T.; Cummin, L.G.; Kanal, E. Dentate nucleus signal intensity increases following repeated gadobenate dimeglumine administrations: A retrospective analysis. *Radiology* **2020**, *296*, 122–130. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Kleesiek, J.; Morshuis, J.N.; Isensee, F.; Deike-Hofmann, K.; Paech, D.; Kickingeder, P.; Köthe, U.; Rother, C.; Forsting, M.; Wick, W.; et al. Can virtual contrast enhancement in brain MRI replace gadolinium? A feasibility study. *Invest. Radiol.* **2019**, *54*, 653–660. [\[CrossRef\]](#)
8. Wong, L.M.; Ai, Q.H.; Mo, F.K.F.; Poon, D.M.; King, A.D. Non contrast-enhanced imaging as a replacement for contrast-enhanced imaging for MRI automatic delineation of nasopharyngeal carcinoma. *medRxiv* **2020**. [\[CrossRef\]](#)
9. Olchow, C.; Cebulski, K.; Łasecki, M.; Chaber, R.; Olchow, A.; Kałwak, K.; Zaleska-Dorobisz, U. The presence of the gadolinium-based contrast agent depositions in the brain and symptoms of gadolinium neurotoxicity—A systematic review. *PLoS ONE* **2017**, *12*, e0171704. [\[CrossRef\]](#)
10. Kanda, T.; Fukusato, T.; Matsuda, M.; Toyoda, K.; Oba, H.; Kotoku, J.; Haruyama, T.; Kitajima, K.; Furui, S. Gadolinium-based contrast agent accumulates in the brain even in subjects without severe renal dysfunction: Evaluation of autopsy brain specimens with inductively coupled plasma mass spectroscopy. *Radiology* **2015**, *276*, 228–232. [\[CrossRef\]](#)

11. Marckmann, P.; Skov, L.; Rossen, K.; Dupont, A.; Damholt, M.B.; Heaf, J.G.; Thomsen, H.S. Nephrogenic systemic fibrosis: Suspected causative role of gadodiamide used for contrast-enhanced magnetic resonance imaging. *J. Am. Soc. Nephrol.* **2006**, *17*, 2359–2362. [\[CrossRef\]](#)
12. Grobner, T.; Prischl, F.C. Gadolinium and nephrogenic systemic fibrosis. *Kidney Int.* **2007**, *72*, 260–264. [\[CrossRef\]](#)
13. Kanal, E.; Tweedle, M.F. Residual or retained gadolinium: Practical implications for radiologists and our patients. *Radiology* **2015**, *275*, 630–634. [\[CrossRef\]](#)
14. Saba, T.; Mohamed, A.S.; El-Affendi, M.; Amin, J.; Sharif, M. Brain tumor detection using fusion of hand crafted and deep learning features. *Cogn. Syst. Res.* **2020**, *59*, 221–230. [\[CrossRef\]](#)
15. Amin, J.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Big data analysis for brain tumor detection: Deep convolutional neural networks. *Future Gener. Comput. Syst.* **2018**, *87*, 290–297. [\[CrossRef\]](#)
16. Li, Q.; Xu, Y.; Chen, Z.; Liu, D.; Feng, S.T.; Law, M.; Ye, Y.; Huang, B. Tumor segmentation in contrast-enhanced magnetic resonance imaging for nasopharyngeal carcinoma: Deep learning with convolutional neural network. *BioMed Res. Int.* **2018**, *2018*, 1–7. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **2017**, *10*, 257–273. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Shen, D.; Wu, G.; Suk, H.I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [\[CrossRef\]](#) [\[PubMed\]](#)
19. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
20. Liang, X.; Chen, L.; Nguyen, D.; Zhou, Z.; Gu, X.; Yang, M.; Wang, J.; Jiang, S. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys. Med. Biol.* **2019**, *64*, 125002. [\[CrossRef\]](#)
21. Ren, G.; Zhang, J.; Li, T.; Xiao, H.; Cheung, L.Y.; Ho, W.Y.; Qin, J.; Cai, J. Deep learning-based computed tomography perfusion mapping (DL-CTPM) for pulmonary CT-to-perfusion translation. *Int. J. Radiat. Oncol. Biol. Phys.* **2021**, *110*, 1508–1518. [\[CrossRef\]](#)
22. Li, W.; Xiao, H.; Li, T.; Ren, G.; Lam, S.; Teng, X.; Liu, C.; Zhang, J.; Lee, F.K.; Au, K.H.; et al. Virtual contrast-enhanced magnetic resonance images synthesis for patients with nasopharyngeal carcinoma using multimodality-guided synergistic neural network. *Int. J. Radiat. Oncol. Biol. Phys.* **2022**, *112*, 1033–1044. [\[CrossRef\]](#)
23. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
24. Chilamkurthy, S.; Ghosh, R.; Tanamala, S.; Biviji, M.; Campeau, N.G.; Venugopal, V.K.; Mahajan, V.; Rao, P.; Warier, P. Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study. *Lancet* **2018**, *392*, 2388–2396. [\[CrossRef\]](#)
25. Li, C.; Sun, H.; Liu, Z.; Wang, M.; Zheng, H.; Wang, S. Learning cross-modal deep representations for multi-modal MR image segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part II 22; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 57–65.
26. Haber, E.; Modersitzki, J. Intensity gradient based registration and fusion of multi-modal images. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006: 9th International Conference, Copenhagen, Denmark, 1–6 October 2006; Proceedings, Part II 9; Springer: Berlin/Heidelberg, Germany, 2006; pp. 726–733.
27. Rühaak, J.; König, L.; Hallmann, M.; Papenberg, N.; Heldmann, S.; Schumacher, H.; Fischer, B. A fully parallel algorithm for multimodal image registration using normalized gradient fields. In Proceedings of the 2013 IEEE 10th International Symposium on Biomedical Imaging, San Francisco, CA, USA, 7 April 2013; IEEE: New York, NY, USA, 2013; pp. 572–575.
28. König, L.; Rühaak, J. A fast and accurate parallel algorithm for non-linear image registration using normalized gradient fields. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April 2014; pp. 580–583.
29. Yeo, T.T.; Ong, S.H.; Sinniah, R. Autofocusing for tissue microscopy. *Image Vis. Comput.* **1993**, *11*, 629–639. [\[CrossRef\]](#)
30. Pieper, S.; Halle, M.; Kikinis, R. 3D Slicer. In Proceedings of the 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), Arlington, VA, USA, 18 April 2004; pp. 632–635.
31. Li, W.; Lam, S.; Wang, Y.; Liu, C.; Li, T.; Kleesiek, J.; Cheung, A.L.; Sun, Y.; Lee, F.K.; Au, K.H.; et al. Model Generalizability Investigation for GFCE-MRI Synthesis in NPC Radiotherapy Using Multi-institutional Patient-based Data Normalization. *IEEE J. Biomed. Health Inform.* **2023**, *28*, 100–109. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Maintz, J.A.; Viergever, M.A. A survey of medical image registration. *Med. Image Anal.* **1998**, *2*, 1–36. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 1–24.

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.