



# Faithful Synthesis of Low-Dose Contrast-Enhanced Brain MRI Scans Using Noise-Preserving Conditional GANs

Thomas Pinetz<sup>1</sup>(✉) , Erich Kobler<sup>2</sup> , Robert Haase<sup>2</sup> ,  
Katerina Deike-Hofmann<sup>2,3</sup> , Alexander Radbruch<sup>2,3</sup> ,  
and Alexander Effland<sup>1</sup>

<sup>1</sup> Institute of Applied Mathematics, Rheinische Friedrich-Wilhelms-Universität Bonn,  
Bonn, Germany

{pinetz,effland}@iam.uni-bonn.de

<sup>2</sup> Department of Neuroradiology, University Medical Center Bonn, Bonn, Germany

<sup>3</sup> German Center for Neurodegenerative Diseases (DZNE),  
Helmholtz Association of German Research Centers, Bonn, Germany

**Abstract.** Today Gadolinium-based contrast agents (GBCA) are indispensable in Magnetic Resonance Imaging (MRI) for diagnosing various diseases. However, GBCAs are expensive and may accumulate in patients with potential side effects, thus dose-reduction is recommended. Still, it is unclear to which extent the GBCA dose can be reduced while preserving the diagnostic value – especially in pathological regions. To address this issue, we collected brain MRI scans at numerous non-standard GBCA dosages and developed a conditional GAN model for synthesizing corresponding images at fractional dose levels. Along with the adversarial loss, we advocate a novel content loss function based on the Wasserstein distance of locally paired patch statistics for the faithful preservation of noise. Our numerical experiments show that conditional GANs are suitable for generating images at different GBCA dose levels and can be used to augment datasets for virtual contrast models. Moreover, our model can be transferred to openly available datasets such as BraTS, where non-standard GBCA dosage images do not exist.

**Keywords:** MRI · GANs · Optimal Transport · Noise Modelling

## 1 Introduction

Magnetic Resonance Imaging (MRI) of the brain is an essential imaging modality to accurately diagnose various neurological diseases ranging from inflammatory

T. Pinetz and A. Effland—are funded the German Research Foundation under Germany’s Excellence Strategy - EXC-2047/1 - 390685813 and - EXC2151 - 390873048 and R. Haase is funded by a research grant (BONFOR; O-194.0002.1).

T. Pinetz and E. Kobler—contributed equally to this work.

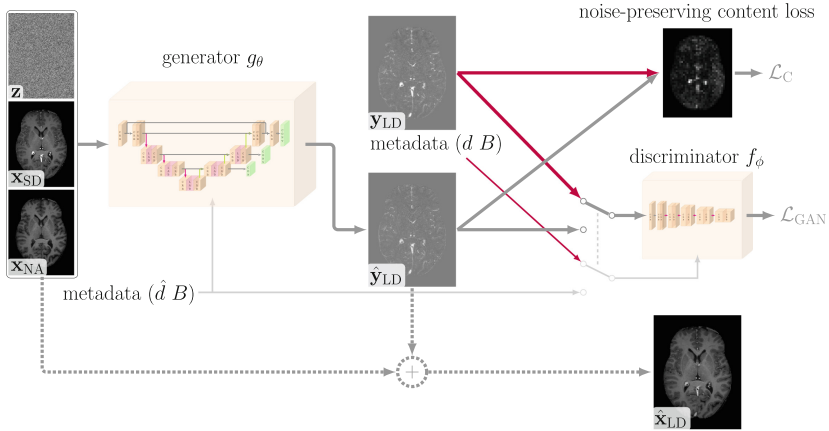
**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-43895-0\\_57](https://doi.org/10.1007/978-3-031-43895-0_57).

lesions to brain tumors and metastases. For accurate depictions of said pathologies, gadolinium-based contrast agents (GBCA) are injected intravenously to highlight brain-blood barrier dysfunctions. However, these contrast agents are expensive and may cause nephrogenic systemic fibrosis in patients with severely reduced kidney function [31]. Moreover, [17] reported that Gadolinium accumulates inside patients with unclear health consequences, especially after repeated application. The American College of Radiology recommends administering the lowest GBCA dose to obtain the needed clinical information [1].

Driven by this recommendation, several research groups have recently published dose-reduction techniques focusing on maintaining image quality. Complementary to the development of higher relaxivity contrast agents [28], virtual contrast [3, 8] – replacing a large fraction of the GBCA dose by deep learning – has been proposed. These approaches typically acquire a contrast-enhanced (CE) scan with a lower GBCA dose along with non-CE scans, e.g., T1w, T2w, FLAIR, or ADC. These input images are then processed by a deep neural network (DNN) to replicate the corresponding standard-dose scan. While promising, virtual contrast techniques have not been integrated into clinical practice yet due to false-positive signals or missed small lesions [3, 23]. As with all deep learning-based approaches, the availability of large datasets is essential, which is problematic in the considered case since the additional CE low-dose scan is not acquired in clinical routine exams. Hence, there are no public datasets to easily benchmark and compare different algorithms or evaluate their performance. In general, the enhancement behavior of pathological tissues at various GBCA dosages has barely been researched due to a lack of data [12].

In recent years, generative models have been used to overcome data scarcity in the computer vision and medical imaging community. Frequently, generative adversarial networks (GANs) [9] are applied as state-of-the-art in image generation [30] or semantic translation/interpolation [5, 18, 21]. In a nutshell, the GAN framework trains two competing DNNs – the generator and the discriminator. The generator learns a non-linear transformation of a predefined noise distribution to fit the distribution of a target dataset, while the discriminator provides feedback by simultaneously approximating a distance or divergence between the generated and the target distribution. The choice of this distance leads to the well-known different GAN algorithms, e.g., Wasserstein GANs [4, 10], Least Squares GANs [24], or Non-saturating GANs [9]. However, Lucic et al. [22] showed that this choice has only a minor impact on the performance.

Learning conditional distributions between images can be accomplished by additionally feeding a condition (additional scans, dose level, etc.) into both the generator and discriminator. In particular, for image-to-image translation tasks, these conditional GANs have been successfully applied using paired [14, 25, 27] and unpaired training data [35]. Within these methods, an additional content (cycle) loss typically penalizes pixel-wise deviations (e.g.,  $\ell_1$ ) from a corresponding reference to enforce structural similarity, whereas a local adversarial loss (discriminator with local receptive field) controls textural similarity. In addition, embeddings have been used to inject metadata [7, 18].



**Fig. 1.** Low-dose synthesis using a conditional GAN. The generator predicts a residual low-dose image  $\hat{\mathbf{y}}_{LD}$  from a noise sample  $\mathbf{z}$  conditioned on the native  $\mathbf{x}_{NA}$  and standard-dose  $\mathbf{x}_{SD}$  images as well as the field strength  $B$  and the artificial dose  $\hat{d}$ . Along with the discriminator, a novel noise-preserving loss – penalizing the Wasserstein distance of paired patches – is used for training. At inference, the generated residual  $\hat{\mathbf{y}}_{LD}$  is added to the native image  $\mathbf{x}_{NA}$  to yield the corresponding synthetic low-dose  $\hat{\mathbf{x}}_{LD}$ .

To study the GBCA accumulation behavior, we collected 453 CE scans with non-standard GBCA doses in the set of  $\{10\%, 20\%, 33\%\}$  along with the corresponding standard-dose (0.1 mmol/kg) scan after applying the remaining contrast agent. Using this dataset, we aim at the semantic interpolation of the GBCA signal at various fractional dose levels. To this end, we use GANs to learn the contrast enhancement behavior from the dataset collective and thereby enable the synthesis of contrast signals at various dose levels for individual cases. Further, to minimize the smoothing effect [19] of typical content losses (e.g.,  $\ell_1$  or perceptual [16]), we develop a noise-preserving content loss function based on the Wasserstein distance between paired image patches calculated using a Sinkhorn-style algorithm. This novel loss enables a faithful generation of noise, which is important for the identification of enhancing pathologies and their usability as additional training data.

With this in mind, the contributions of this work are as follows:

- synthesis of GBCA behavior at various doses using conditional GANs,
- loss enabling interpolation of dose levels present in training data,
- noise-preserving content loss function to generate realistic synthetic images.

## 2 Methodology

Given a *native* image  $\mathbf{x}_{NA}$  (i.e. without any contrast agent injection) and a CE *standard-dose* image  $\mathbf{x}_{SD}$ , our conditional GAN approach synthesizes CE *low-dose* images  $\hat{\mathbf{x}}_{LD}$  for selected dose levels  $\hat{d} \in \mathcal{D} \subset [0, 1]$  from a uniform noise

image  $\mathbf{z} \sim \mathcal{N}(0, \text{Id})$ , see Fig. 1. To focus the generation on the contrast agent signal, our model predicts residual images  $\hat{\mathbf{y}}_{\text{LD}}$ ; the corresponding low-dose can be obtained by  $\hat{\mathbf{x}}_{\text{LD}} = \mathbf{x}_{\text{NA}} + \hat{\mathbf{y}}_{\text{LD}}$ .

For training and evaluation, we consider samples  $(\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, \mathbf{y}_{\text{LD}}, d, B)$  of a dataset  $\mathcal{DS}$ , where  $\mathbf{y}_{\text{LD}} = \mathbf{x}_{\text{LD}} - \mathbf{x}_{\text{NA}}$  is the residual image of a real CE low-dose scan  $\mathbf{x}_{\text{LD}}$  with dose level  $d \in \mathcal{D}$  and  $B \in \{1.5, 3\}$  is the field-strength in Tesla of the used scanner. To simplify learning of the contrast accumulation behavior, we adapt the preprocessing pipeline of BraTS [6]. Further details of the dataset and the preprocessing are in the supplementary material.

## 2.1 Conditional GANs for Contrast Signal Synthesis

Our approach is built on the insight that contrast enhancement is an inherently local phenomenon and the necessary information for the synthesis task can be extracted from a local neighborhood within an image. Therefore, we use as generator  $g_\theta$  a convolutional neural network (CNN) that is based on the U-Net [29] along with a local attention mechanism. The architecture design and the implementation details can be found in the supplementary material. As illustrated in Fig. 1, the generator uses a 3D noise sample  $\mathbf{z} \sim \mathcal{N}(0, \text{Id})$  along with the native and SD images  $(\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}})$  as input. The synthesis is guided by the meta-data  $(\hat{d} B)^\top$ , containing the artificial dose level  $\hat{d} \in \mathcal{D}$  as well as the field strength of the corresponding scanner  $B \in \{1.5, 3\}$ . In particular, the metadata is injected into every residual block of the generator using an embedding, motivated by the recent success of diffusion-based models [13].

To learn this generator, a convolutional discriminator  $f_\phi$  is used, which is in turn trained to distinguish the generated residual images  $\hat{\mathbf{y}}_{\text{LD}}$  with random dose level  $\hat{d}$  from the real residual images  $\mathbf{y}_{\text{LD}}$  with the associated real dose level  $d$ . To make this a non-trivial task, label smoothing on the metadata is used, i.e., the real dose is augmented by zero-mean additive Gaussian noise with standard deviation 0.05. The discriminator architecture essentially implements the encoding side of the generator, however, no local attention layers are used as suggested by [20]. Like the generator, the discriminator is conditioned on the metadata using an embedding, which is not shared between both networks.

For training of the generator  $\theta$  and discriminator  $\phi$ , we consider the loss

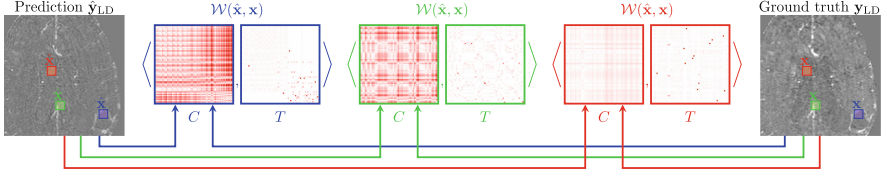
$$\min_{\theta} \max_{\phi} \{ \mathcal{L}_{\text{GAN}}(\theta, \phi) + \lambda_{\text{GP}} \mathcal{L}_{\text{GP}}(\phi) + \lambda_{\text{C}} \mathcal{L}_{\text{C}}(\theta) \}, \quad (1)$$

which consists of a Wasserstein GAN loss  $\mathcal{L}_{\text{GAN}}$ , a gradient penalty loss  $\mathcal{L}_{\text{GP}}$ , and a content loss  $\mathcal{L}_{\text{C}}$  that are relatively weighted by scalar non-negative factors  $\lambda_{\text{GP}}$  and  $\lambda_{\text{C}}$ . In detail, the Wasserstein GAN loss reads as

$$\mathcal{L}_{\text{GAN}}(\theta, \phi) :=$$

$$\mathbb{E}_{(\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, \mathbf{y}_{\text{LD}}, d, B) \sim \mathcal{U}(\mathcal{DS})} \left\{ f_\phi(\mathbf{y}_{\text{LD}}, c) - \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(0, \text{Id}), \hat{d} \sim \mathcal{U}(\mathcal{D})} \{ f_\phi(g_\theta(\mathbf{z}, \hat{c}), \hat{c}) \} \right\}$$

using condition tuples  $c = (\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, (d B)^\top)$  and  $\hat{c} = (\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, (\hat{d} B)^\top)$  to simplify notation.  $\mathcal{U}(\mathcal{S})$  denotes a uniform distribution over a set  $\mathcal{S}$ . We highlight that the artificial dose levels  $\hat{d}$  for the generated images are uniformly



**Fig. 2.** Illustration of our patch-wise noise-preserving content loss. For each patch pair  $((\hat{\mathbf{x}}, \mathbf{x}), (\tilde{\mathbf{x}}, \mathbf{x}), (\tilde{\mathbf{x}}, \mathbf{x}))$  extracted at the same position, the loss accounts for the Wasserstein distance  $\mathcal{W}$  of the associated empirical distributions. In the center, the corresponding cost matrices  $C$  (pixel-wise absolute difference) along with the optimal transport maps  $T$  are shown, which are obtained by solving (2). The final loss is the sum of the element-wise multiplication of all  $C$  and  $T$  for every non-overlapping patch. (Color figure online)

sampled from  $\mathcal{D} = [0.05, 0.5]$ , which enables an interpolation around the dose levels present in the dataset  $\mathcal{DS}$ . This is necessary since only a few distinct dose levels  $\{0.1, 0.2, 0.33\}$  have been acquired. For regularizing the discriminator  $f_\phi$ , we include the gradient penalty loss

$$\mathcal{L}_{\text{GP}}(\phi) := \mathbb{E}_{(\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, \mathbf{y}_{\text{LD}}, d, B) \sim \mathcal{U}(\mathcal{DS}), \mathbf{z} \sim \mathcal{N}(0, \text{Id}), \hat{d} \sim \mathcal{U}(\mathcal{D}), \alpha \sim \mathcal{U}(0, 1)} \left\{ (\|\nabla f_\phi(h(\alpha, \mathbf{y}_{\text{LD}}, \hat{\mathbf{y}}_{\text{LD}}), h(\alpha, c, \hat{c}))\|_2 - 1)^2 \right\}$$

using  $h(\alpha, \mathbf{y}, \hat{\mathbf{y}}) = \alpha \hat{\mathbf{y}} + (1 - \alpha) \mathbf{y}$ . A penalty term is introduced, if  $f_\phi$  is not Lipschitz continuous with factor 1 in its arguments as required by Wasserstein GANs [10]. Here,  $\hat{\mathbf{y}}_{\text{LD}} = g_\theta(\mathbf{z}, \hat{c})$  and the Lipschitz penalty is evaluated at convex combinations of real and synthetic images and condition tuples (essentially dose levels). Finally, using a distance  $\ell_C$ , the content loss

$$\mathcal{L}_C(\theta) := E_{(\mathbf{x}_{\text{NA}}, \mathbf{x}_{\text{SD}}, \mathbf{y}_{\text{LD}}, d, B) \sim \mathcal{U}(\mathcal{DS}), \mathbf{z} \sim \mathcal{N}(0, \text{Id})} \left\{ \ell_C(g_\theta(\mathbf{z}, c), \mathbf{y}_{\text{LD}}) \right\}$$

guides the generator  $g_\theta$  towards residual images in the dataset. Thus, it teaches the generator the principles of contrast enhancement. Typically, the  $\ell_1$ -norm is used as a distance function, which leads to smooth results since it also penalizes deviations from the noise in  $\mathbf{y}_{\text{LD}}$ .

## 2.2 Noise-Preserving Content Loss

To generate realistic CE images, it is also important to retain the original noise characteristics. Therefore, we introduce a novel loss that accounts for deviations in local statistics using optimal transport between empirical distributions of paired patches, as illustrated in Fig. 2.

Let  $\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}^{n^3}$  be patches of size  $n \times n \times n$  extracted from the same location of a real and synthetic image, respectively. The Wasserstein distance of the associated empirical distributions using a transport plan  $T \in \mathbb{R}_+^{n^3 \times n^3}$  and cost matrix  $C \in \mathbb{R}_+^{n^3 \times n^3}$  given by  $(C_{ij} = |\hat{x}_i - x_j|)$  is defined as

$$\mathcal{W}(\hat{\mathbf{x}}, \mathbf{x}) = \min_{T \in \mathbb{R}_+^{n^3 \times n^3}} \langle C, T \rangle_F \quad \text{s.t.} \quad T\mathbf{1} = \mathbf{1}_{\frac{1}{n^3}}, \quad T^\top \mathbf{1} = \mathbf{1}_{\frac{1}{n^3}}, \quad (2)$$

where  $\mathbf{1}$  is the vector of ones of size  $n^3$ . In contrast to the element-wise difference penalization of the  $\ell_1$ -distance, the Wasserstein distance accounts for mismatches between distributions. To illustrate this, let us, for instance, assume that both patches are Gaussian distributed ( $x \sim \mathcal{N}(\mu, \sigma)$ ,  $\hat{x} \sim \mathcal{N}(\hat{\mu}, \hat{\sigma})$ ), which is a coarse simplification of real MRI noise [2]. In this case, the Wasserstein distance reduces to second-order momentum matching, i.e.,  $\mathcal{W}^2(\hat{\mathbf{x}}, \mathbf{x}) = (\mu - \hat{\mu})^2 + (\sigma - \hat{\sigma})^2$ . Thus, the Wasserstein distance generalizes this distributional loss to any distribution within paired patches.

To efficiently solve problem (2), we use the inexact proximal point algorithm [34]. This algorithm is parallelized and applied to all non-overlapping patch pairs, to obtain our noise-preserving content loss

$$\ell_{\text{NP}}(\hat{\mathbf{y}}, \mathbf{y}) = \mathbb{E}_{\mathbf{o} \sim \mathcal{U}(\mathcal{O})} \left\{ \sum_{\mathbf{p} \in \mathcal{P}} \mathcal{W}(P_{\mathbf{p}+\mathbf{o}}\hat{\mathbf{y}}, P_{\mathbf{p}+\mathbf{o}}\mathbf{y}) \right\},$$

where  $P_{\mathbf{p}}$  extracts a local  $n \times n \times n$  patch at location  $\mathbf{p} \in \mathcal{P} = \{0, n, 2n, \dots\}^3$  using periodic boundary conditions. Note that we compute the expectation over offsets  $\mathbf{o} \in \mathcal{O} = \{0, 1, \dots, \lceil \frac{n}{2} \rceil\}^3$  to avoid patching artifacts. In the numerical implementation, only a single offset is sampled for time and memory constraints.

### 3 Numerical Results

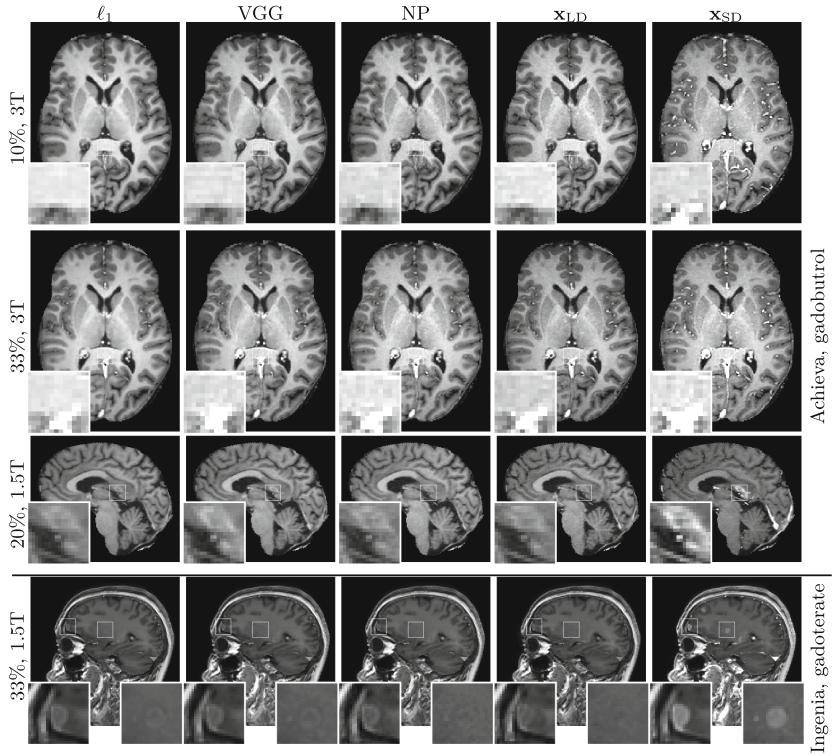
In this section, we evaluate the proposed conditional GAN approach with a particular focus on different content loss distance functions. All synthesis models were trained on 250 samples acquired on 1.5T and 3T Philips Achieva scanners and evaluated on 193 test cases, all collected at site ①. Further details of the dataset, model and training can be found in the supplementary. In our experiments, we observed that the choice of the content loss distance function  $\ell_C(\hat{\mathbf{y}}, \mathbf{y})$  strongly influences the performance. Thus, we consider the different cases:

$$\ell_1 : \quad \|\hat{\mathbf{y}} - \mathbf{y}\|_1 \quad \text{VGG:} \quad \|h(\hat{\mathbf{y}}) - h(\mathbf{y})\|_1 \quad \text{NP:} \quad \ell_{\text{NP}}(\hat{\mathbf{y}}, \mathbf{y})$$

Following Johnsen et al. [16],  $h(\mathbf{x})$  is the VGG-16 model [32] up to `relu3_3`.

A qualitative comparison of the different distance functions  $\ell_C$  is visualized in Fig. 3. The first column depicts synthesized images using the  $\ell_1$ -norm as the distance function. These images depict a plausible contrast signal, however, suffer from unrealistic smooth homogeneous regions. An improvement thereof is shown by the perceptual content loss (VGG). The NP-loss leads to a further improvement not only in the contrast signal behavior but also in the realism of the noise texture, cf. zoom regions in the lower corners.

To highlight the generalization capabilities, we depict in the bottom row of Fig. 3 a sample from site ②, which was acquired using a Philips Ingenia scanner. Moreover, the GBCA gadoterate was used, while our training data only consists of scans using the GBCA gadobutrol. Nevertheless, all models present realistically synthesized LD images. Comparing the zooms of the LD images, we observe that our NP-loss leads to a better synthesis of noise and thereby to



**Fig. 3.** Qualitative comparison of synthesized images using different loss functions to the corresponding reference  $\mathbf{x}_{LD}$ . While the  $\ell_1$  loss yields smooth low-dose images, the noise pattern is preserved to some extent using the VGG loss; our loss helps to further retain the noise characteristics.

more realistic LD images. In the  $\ell_1$  and VGG columns, the noise is not faithfully synthesized, thus it is visually easy to spot the enhancing pathological regions.

For completeness, a quantitative ablation of the considered distance functions on the test images of site ① is shown in Table 1. Although neither maximizing PSNR nor SSIM [33] is our objective, we observe on-par performances of the perceptual (VGG) and our proposed content loss (NP) with the standard  $\ell_1$  distance function. Using the SD image, we define CE pixels as those pixels at which the intensity increases by at least 10% compared to the native scan. An example of these CE regions is illustrated in the supplementary. Thus, the mean absolute error for CE pixels ( $\text{MAE}_{CE}$ ) quantifies the enhancement behavior. Further, we estimate the standard deviation of the non-CE pixels and report the MAE to the ground truth standard deviation ( $\text{MAE}_\sigma$ ). As shown in Table 1, our loss outperforms the other content losses to a large extent on both metrics, proving its effectiveness for faithful contrast enhancement and noise generation. Further statistical analyses are presented in the supplementary.



**Table 1.** Quantitative comparison of the low-dose synthesis methods. The central columns present metrics evaluated on the synthesized low-dose images, whereas the right columns evaluate the effect of purely synthesized data for training the standard-dose prediction model [26]. Note, that the PSNR/SSIM of the standard dose prediction model was always evaluated on real LD images. The definitions of the mean absolute error on the contrast enhancement ( $\text{MAE}_{\text{CE}}$ ) and on the noise standard deviation ( $\text{MAE}_{\sigma}$ ) are in Sect. 3. A \* denotes if a Wilcoxon signed rank test between VGG and  $\text{NP}_{(\text{our})}$  row is significant.

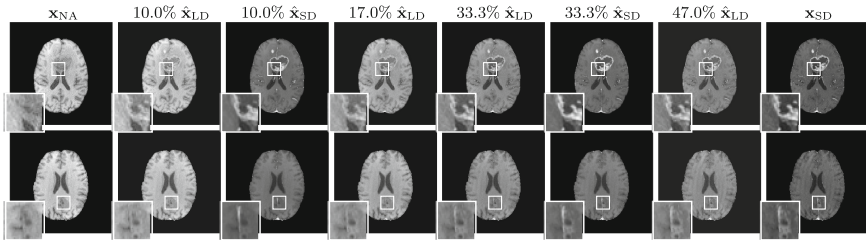
	low-dose synthesis				standard-dose prediction	
	PSNR	SSIM	$\text{MAE}_{\text{CE}}$	$\text{MAE}_{\sigma}$	PSNR	SSIM
$\ell_1$	<b>38.34</b>	<b>0.978</b>	0.022	0.012	33.83	0.922
VGG	37.84	0.976	0.019	0.009	36.33	0.958
$\text{NP}_{(\text{our})}$	38.05*	0.976	<b>0.011*</b>	<b>0.004*</b>	<b>37.15*</b>	<b>0.960*</b>
$\mathbf{x}_{\text{LD}}(\text{real})$					39.07	0.974

Next, we evaluate the effect of synthesized LD images on the performance of a virtual contrast model (VCM). In particular, we consider the state-of-the-art 2.5D U-Net model [11, 23, 26], which predicts an SD image given a corresponding native and LD image, see supplementary for further details. The columns on the right of Table 1 list the average PSNR and SSIM score on the *real* 33% LD subset of our test data from site ①. The bottom row depicts the performance if just *real* 33% LD images are used for training the VCM as an upper bound. In contrast, the other entries on the right list the performance if *only synthesized* LD images are used for training. Both metrics show that the samples synthesized using our NP-loss model are superior to both  $\ell_1$  and VGG.

To determine the effectiveness of the LD synthesis models at different settings, we acquired 160 data samples from 1.5T and 3T Philips Ingenia scanners at site ②. This site used the GBCA gadoterate, which has a lower relaxivity compared to gadobutrol used at site ① [15]. For 80 samples real LD images were acquired, which are used for testing. Using the VCM solely trained on the real 33% LD data of site ① yields an average PSNR and  $\text{MAE}_{\text{CE}}$  on the test samples of site ② of 40.04 and 0.092, respectively. Extending the training data for the VCM by synthesized LD images from our model with NP-loss, we get a significantly improved ( $p < 0.001$ ) PSNR score of 40.37 and  $\text{MAE}_{\text{CE}}$  of 0.075.

Finally, Fig. 4 visualizes synthesized LD images on the BraTS dataset [6] along with the associated VCM outputs. Comparing the predicted SD images  $\hat{\mathbf{x}}_{\text{SD}}$  using 10% and 33% synthesized LD images  $\hat{\mathbf{x}}_{\text{LD}}$ , we observe that the weakly enhancing tumor at the bottom zoom is not preserved in the case of 10%, enabling evaluation of dose reduction methods on known pathological regions.





**Fig. 4.** Comparison of synthesized LD images  $\hat{x}_{LD}$  and corresponding predicted SD images  $\hat{x}_{SD}$  for different dose levels on BraTS [6] along with the native (left) and real SD image (right). We also included non-fractional dosage levels (17% and 47%) to showcase the wide applicability of our algorithm. Top: the tumor is well contrasted in all  $\hat{x}_{SD}$  even for 10%. Bottom: the subtle enhancement of the tumor cannot be recovered from the 10% LD image.

## 4 Conclusions

In this work, we used conditional GANs to synthesize contrast-enhanced images using non-standard GBCA doses. To this end, we introduced a novel noise-preserving content loss motivated by optimal transport theory. Numerous numerical experiments showed that our content loss improves the faithful synthesis of low-dose images. Further, the performance of virtual contrast models increases if training data is extended by synthesized images from our GAN model trained by the noise-preserving content loss.

## References

1. ACR Manual on Contrast Media. American College of Radiology (2022)
2. Aja-Fernández, S., Vegas-Sánchez-Ferrero, G.: Statistical Analysis of Noise in MRI. Springer, Cham (2016). <https://doi.org/10.1007/978-3-319-39934-8>
3. Ammari, S., et al.: Can deep learning replace gadolinium in neuro-oncology?: A reader study. *Invest. Radiol.* **57**(2), 99–107 (2022)
4. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: Proceedings of the 34th International Conference on Machine Learning (ICML), pp. 214–223 (2017)
5. Armanious, K., et al.: MedGAN: medical image translation using GANs. *Comput. Med. Imaging Graph.* **79**, 101684 (2020)
6. Baid, U., et al.: The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314* (2021)
7. Choi, Y., Uh, Y., Yoo, J., Ha, J.W.: StarGAN v2: diverse image synthesis for multiple domains. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8188–8197 (2020)
8. Gong, E., Pauly, J.M., Wintermark, M., Zaharchuk, G.: Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. *J. Magn. Reson. Imaging* **48**(2), 330–340 (2018)

9. Goodfellow, I., et al.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K. (eds.) *NeurIPS*. vol. 27. Curran Associates, Inc. (2014)
10. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein GANs. In: *Advances in Neural Information Processing Systems (NeurIPS)*. vol. 30 (2017)
11. Haase, R., et al.: Reduction of gadolinium-based contrast agents in MRI using convolutional neural networks and different input protocols: limited interchangeability of synthesized sequences with original full-dose images despite excellent quantitative performance. *Invest. Radiol.* **58**(6), 420–430 (2023)
12. Haase, R., et al.: Artificial contrast: Deep learning for reducing gadolinium-based contrast agents in neuroradiology. *Invest. Radiol.* **58**(8), 539–547 (2023)
13. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *NeurIPS* **33**, 6840–6851 (2020)
14. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1125–1134 (2017)
15. Jacques, V., Dumas, S., Sun, W.C., Troughton, J.S., Greenfield, M.T., Caravan, P.: High relaxivity MRI contrast agents part 2: optimization of inner-and second-sphere relaxivity. *Invest. Radiol.* **45**(10), 613 (2010)
16. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46475-6\\_43](https://doi.org/10.1007/978-3-319-46475-6_43)
17. 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* **270**(3), 834–841 (2014)
18. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4401–4410 (2019)
19. Larsen, A.B.L., Sønderby, S.K., Larochelle, H., Winther, O.: Autoencoding beyond pixels using a learned similarity metric. In: *International Conference on Machine Learning (ICML)*, pp. 1558–1566 (2016)
20. Lee, K., Chang, H., Jiang, L., Zhang, H., Tu, Z., Liu, C.: ViTGAN: training GANs with vision transformers. In: *International Conference on Learning Representations (ICLR)* (2022)
21. Liu, J., Pasumarthi, S., Duffy, B., Gong, E., Zaharchuk, G., Datta, K.: One model to synthesize them all: Multi-contrast multi-scale transformer for missing data imputation. *arXiv preprint arXiv:2204.13738* (2022)
22. Lucic, M., Kurach, K., Michalski, M., Gelly, S., Bousquet, O.: Are GANs created equal? A large-scale study. In: *Neural Information Processing Systems (NeurIPS)* (2018)
23. Luo, H., et al.: Deep learning-based methods may minimize GBCA dosage in brain MRI. *Eur. Radiol.* **31**(9), 6419–6428 (2021)
24. Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Paul Smolley, S.: Least squares generative adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2794–2802 (2017)

25. Nie, D., et al.: Medical image synthesis with context-aware generative adversarial networks. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 417–425. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-66179-7\\_48](https://doi.org/10.1007/978-3-319-66179-7_48)
26. Pasumarthi, S., Tamir, J.I., Christensen, S., Zaharchuk, G., Zhang, T., Gong, E.: A generic deep learning model for reduced gadolinium dose in contrast-enhanced brain MRI. *Magn. Reson. Med.* **86**(3), 1687–1700 (2021)
27. Preetha, C.J., et al.: Deep-learning-based synthesis of post-contrast t1-weighted MRI for tumour response assessment in neuro-oncology: a multicentre, retrospective cohort study. *Lancet Digital Health* **3**(12), e784–e794 (2021)
28. Robic, C., et al.: Physicochemical and pharmacokinetic profiles of Gadopiclenol: a new macrocyclic gadolinium chelate with high t1 relaxivity. *Invest. Radiol.* **54**(8), 475 (2019)
29. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
30. Sauer, A., Schwarz, K., Geiger, A.: StyleGAN-XL: Scaling styleGAN to large diverse datasets. In: ACM SIGGRAPH, pp. 1–10 (2022)
31. Schieda, N., et al.: Gadolinium-based contrast agents in kidney disease: a comprehensive review and clinical practice guideline issued by the Canadian association of radiologists. *Can. J. Kidney Health Dis.* **5**, 136–150 (2018)
32. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (ICLR), pp. 1–14 (2015)
33. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *Trans. Image Process.* **13**(4), 600–612 (2004)
34. Xie, Y., Wang, X., Wang, R., Zha, H.: A fast proximal point method for computing exact wasserstein distance. In: Uncertainty in Artificial Intelligence, pp. 433–453 (2020)
35. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE International Conference on Computer Vision (ICCV), pp. 2223–2232 (2017)