MAUDGAN: Motion Artifact Unsupervised Disentanglement Generative Adversarial Network of Multicenter MRI Data with Different Brain tumors

⁴ Mojtaba Safari, MSc^{a,b,*}, Ali Fatemi, PhD^{c,d}, Louis Archambault, PhD^{a,b}

⁵ ^aDépartement de physique, de génie physique et d'optique, et Centre de recherche sur le cancer, Université Laval,

- 6 Québec, Québec, Canada
- ⁷ ^bService de physique médicale et radioprotection, Centre Intégré de Cancérologie, CHU de Québec Université
- 8 Laval et Centre de recherche du CHU de Québec, Québec, Québec, Canada
- ⁹ ^cDepartment of Physics, Jackson State University, Mississipi, USA
- ¹⁰ ^dMerit Health Central, Department of Radiation Oncology, Gamma Knife Center, Mississipi, USA

Abstract. **Purpose:** This study proposed a novel retrospective motion reduction method named motion artifact 11 unsupervised disentanglement generative adversarial network (MAUDGAN) that reduces the motion artifacts from 12 brain images with tumors and metastases. The MAUDGAN was trained using a mutlimodal multicenter 3D T1-13 Gd and T2-fluid attenuated inversion recovery MRI images. Approach: The motion artifact with different artifact 14 levels were simulated in k-space for the 3D T1-Gd MRI images. The MAUDGAN consisted of two generators, 15 two discriminators and two feature extractor networks constructed using the residual blocks. The generators map 16 the images from content space to artifact space and vice-versa. On the other hand, the discriminators attempted to 17 discriminate the content codes to learn the motion-free and motion-corrupted content spaces. Results: We compared 18 the MAUDGAN with the CycleGAN and Pix2pix-GAN. Qualitatively, the MAUDGAN could remove the motion 19 with the highest level of soft-tissue contrasts without adding spatial and frequency distortions. Quantitatively, we 20 reported six metrics including normalized mean squared error (NMSE), structural similarity index (SSIM), multi-21 scale structural similarity index (MS-SSIM), peak signal-to-noise ratio (PSNR), visual information fidelity (VIF), and 22 multi-scale gradient magnitude similarity deviation (MS-GMSD). The MAUDGAN got the lowest NMSE and MS-23 GMSD. On average, the proposed MAUDGAN reconstructed motion-free images with the highest SSIM, PSNR, and 24 VIF values and comparable MS-SSIM values. Conclusions: The MAUDGAN can disentangle motion artifacts from 25 the 3D T1-Gd dataset under a multimodal framework. The motion reduction will improve automatic and manual 26 post-processing algorithms including auto-segmentations, registrations, and contouring for guided therapies such as 27 radiotherapy and surgery. 28

29 Keywords: MRI artifact reduction, motion artifact simulation, deep learning, multi-modal MRI, k-space.

30 *Mojtaba Safari, mojtaba.safari.1@ulaval.ca

31 1 Introduction

- ³² Magnetic resonance imaging (MRI) with different sequences provides excellent soft tissue contrast
- ³³ for diagnosis and treatment planning. However, high MRI acquisition time limits the quality of
- ³⁴ high-resolution images¹ because of the increased probability of patient motion. Involuntary and
- voluntary subject motions during data acquisition cause image blurring and ghosting along the

³⁶ phase-encoding direction. The prevalence of motion artifacts is high for infants and patients with ³⁷ acute distress.²

To tackle motion artifacts, retrospective motion correction (RMC) and prospective motion cor-38 rection (PMC) methods were developed. PMC approaches modify the gradient magnetic fields 39 using the imaged object positions that are tracked during imaging to maintain a constant relation-40 ship between imaged object and imaged volume.^{3,4} PMC can maintain a uniform k-space sampling 41 density, which avoids Nyquist violation, and compensate for spin-history effects.⁵ However, PMC 42 methods require additional hardware and complicated pulse sequences that increase the imaging 43 time. On the other hand, RMC methods are post-processing approaches, and do not require ad-44 ditional hardware and pulse sequence modifications during imaging. Traditional RMC methods, 45 such as auto-focusing, attempt to optimize image quality metrics like entropy and gradient,⁶ iter-46 ative methods to estimate motion trajectory,⁷ compressed-sensing theory,⁸ and modified imaging 47 sequences.⁹ They are either limited to 2D imaging methods or require raw k-space data that are 48 not widely available. In addition, these methods are computationally expensive. 49

Recently, deep learning techniques, in particular, convolutional neural networks (CNNs) have been used to quantify^{10,11} and reduce^{12,13} MRI motion artifact retrospectively. These models learn the task through a supervised framework using the simulating motion artifacts. Unpaired deep learning models attempted to use data without the motion artifact as a ground truth to reduce the artifacts from MRI with the same imaging sequence.¹⁴

This study aimed to address the problem in a more practical setting where one motion-free MRI modality removes artifacts from the motion-corrupted images acquired with different MRI imaging sequences. This study reformulated MRI motion artifacts as an unsupervised disentanglement problem. Thus, we introduced a novel motion artifact unsupervised disentanglement generative



Fig 1: Content and artifact components of 3D T1-Gd MRI (x_a) in the motion-corrupted space \mathcal{T}_a and T2-FLAIR in artifact-free space \mathcal{T} are mapped to the content space \mathcal{C} and artifact space \mathcal{A} , respectively. MAUDGAN maps the data in \mathcal{T}_a space to \mathcal{T} space $(x_a \to \hat{x})$ shown by blue arrows. Conversely, MADuGAN learns to map from \mathcal{T} space to \mathcal{T}_a space $(y \to \hat{y}_a)$ shown by green arrows.

⁵⁹ adversarial network (MAUDGAN). The novel MAUDGAN was applied to reduce the motion of ⁶⁰ 3D T1-Gd MRI sequences using motion-free T2-fluid attenuated inversion recovery (FLAIR) se-⁶¹ quences for the patients with different brain cancers metastasis. This study used a multicenter ⁶² dataset to improve the MUADGAN's generalization.

This study leverages an inductive bias¹⁵ that the MAUDGAN learn to disentangle motion artifacts from motion-free contents by comparing 3D T1-Gd MRI sequences (typically with motion artifacts) with motion-free T2-FLAIR (Figure 1) in the latent space.

The MAUDGAN consists of U-net¹⁶ generators to perform different forms of image translations including motion artifact reduction and synthesis. Discriminators were used to distinguish between the motion-free and the motion-corrupted MRI sequences in the latent spaces. To our knowledge, MAUDGAN is the first study in multi-modal anatomical MRI motion artifact reduction.

The rest of this paper is as follows: Section 2 explains the dataset and motion simulation steps. Section 3 gives a detail about the MAUDGAN architecture and loss functions. Results and comparisons with two generative models are illustrated in Section 4. Finally, Sections 6 and 5 ⁷⁴ discuss the significance of the MAUDGAN and its possible use in the context of diagnosis and
 ⁷⁵ therapy.

76 2 Material

77 2.1 Dataset

We used a publicly available multicenter medical GLIS-RT dataset from the Cancer Imaging Archive¹⁷ consisting of 230 patients (100 males and 130 females). All patients with different brain tumor types underwent 3D T1-Gd, 2D T2-FLAIR MRI sequences, and a CT scan under different imaging protocols. The brain tumor types were glioblastoma (GBM - 198 cases), anaplastic astrocytoma (AAC - 23 cases), astrocytoma (AC - 5 cases), anaplastic oligodendroglioma (AODG - 2 cases), and oligodendroglioma (ODG - 2 case). We used 80% (11246 image slices) and 20% (2276 image slices) of data for training and testing our method, respectively.

The median of the T2-FLAIR and 3D T1-Gd images' resolution was $1.1 \times 1.1 \times 5$ mm³ (standard 85 deviation $0.53 \times 0.53 \times 0.87$ mm³) and $0.94 \times 0.94 \times 1$. mm³ (standard deviation $0.24 \times 0.24 \times 1.21$ 86 mm³), respectively. The T2-FLAIR imaging parameters were (median \pm std); TE = 119 ± 64.06 87 ms, TR = 9000 \pm 936.20 ms, TI = 2500 \pm 174.02 ms, and flip angle = $150^{\circ} \pm 13.56^{\circ}$. Those 88 parameters for T1-Gd were (median \pm std); TE = 2.98 ± 3.86 ms, TR = 2200 ± 1031.76 ms, TI = 89 900 ± 235.50 ms, and flip angle = $9^{\circ} \pm 5.45^{\circ}$ About 30% of data were acquired using MRI scanners 90 with B₀ of 1.5 T and the others were acquired using 3T scanners. Out of 230 cases, 55 cases were 91 obtained using GE MRI scanners and the rest were obtained using Siemens MRI scanners. 92

Finally, we evaluated the MAUDGAN performance on anonymized clinical data with real mo tion artifacts. This retrospective single-centre study was approved by the institutional review board,
 and the requirement for written informed consent was waived.

96 2.2 Motion simulation

⁹⁷ The head motion was simulated in the Fourier domain (*k*-space), and the motion-corrupted data was ⁹⁸ generated after the inverse discrete Fourier transform. We adapted the piecewise constant motion ⁹⁹ simulation approach with a low computation burden because it provides a similar generalization ¹⁰⁰ than to the complex motion simulation techniques.¹³ Moreover, the generated motion artifacts ¹⁰¹ were similar to the real motion artifacts.¹³

We assumed the phase encoding interval was much faster than the head motion. Thus, the same motion parameters could be used at each phase encoding direction (Figure 2). The *k*-space lines within the randomly selected slabs were translated in the phase encoding direction. However, the middle of the *k*-space that corresponds to the low-frequency content of the MRI images was excluded in the motion artifact simulation process, shown as a forbidden region in Figure 2. Our motion simulation method could successfully model the ghosting of the bright fat tissue, due to the motion artifact, to the background around the skull, which is common in structural MRI images.¹⁸

109 3 Method

We denote \mathcal{T}_a and \mathcal{T} as the motion-corrupted image and the motion-free image spaces, respectively. 110 The paired and unpaired motion reduction process is formalized as a $\mathcal{M} = \{(x_a, x) \mid x_c \in \mathcal{T}_a, x \in \mathcal{T}_a,$ 11 $\mathcal{T}, f(x_a) = x$ where x_a and x were the motion-corrupted and motion-free single MRI image 112 sequence and $f: \mathcal{T}_a \to \mathcal{T}.^{14,19}$ However, we assumed there is no paired or unpaired dataset of a 113 single modality available to disentangle motion artifacts. Instead, another MRI image sequence, 114 T2-FLAIR, was employed to disentangle the motion artifact of the T1-Gd MRI sequence, which 115 is more practical in clinical settings. Thus, the MAUDGAN is formalized as $\mathcal{M} = \{(x_a, y) \mid x_a \in$ 116 $\mathcal{T}_a, y \in \mathcal{T}, f(x_a) = x, g(x_a, y) = y_a$ where $f : \mathcal{T}_a \to \mathcal{T}$ and $g : \mathcal{T} \to \mathcal{T}_a$ are the encoding into a 117



Fig 2: The motion simulation process. After choosing the phase encoding direction, several random k-space regions were selected. The randomly selected k-space lines were randomly translated within the random regions.

content space C and artifact space A. Also, x_a and y are motion-corrupted T1-Gd and the motion-118 free T2-FLAIR MRI images. After training the MAUDGAN, the image data in the content space 119 will be free of motion artifacts. In contrast, the motion-corrupted T2-FLAIR could be generated 120 using the learned motion artifact model. 121

3.1 MAUDGAN 122

127

The MAUDGAN consists of two generators $\mathcal{F} : \mathcal{T}_a \to \mathcal{T}$ and $\mathcal{G} : \mathcal{T} \to \mathcal{T}_a$ to map from motion-123 corrupted space to motion-free space and vice-versa (Figure 3). In addition, two networks \mathcal{H}_{ν_1} and 124 \mathcal{H}_{ν_2} were also employed to extract features of the images before feeding them to the generators. 125 Given multimodal MRI images T1-Gd $x_a \in T_a$ and T2-FLAIR $y \in T$, the training steps were 126 as follows:



Fig 3: The proposed MAUDGAN is illustrated. The Generators \mathcal{F} learns disentanglement while the \mathcal{G} learns to generate motion-corrupted images from motion-free images.

128 1. \mathcal{F} maps the motion-corrupted T1-Gd x_a to motion-free space \hat{x} ,

$$\hat{x} = \mathcal{F}(x_a; \theta) \tag{1}$$

129 2. \mathcal{G} maps the motion-free space T2-FLAIR y to the motion-corrupted space \hat{y}_a ,

$$\hat{y}_a = \mathcal{G}(\mathcal{H}_{\nu_1}(x_a), \mathcal{H}_{\nu_2}(y); \varrho) \tag{2}$$

3. trained \mathcal{F} in step 1 was used to recover motion-free T2-FLAIR \tilde{y} from motion-corrupted \hat{y}_a simulated in step 2,

$$\tilde{y} = \mathcal{F}(\hat{y}_a, \theta) = \mathcal{F}\left(\mathcal{G}(\mathcal{H}_{\nu_1}(x_a), \mathcal{H}_{\nu_2}(y); \varrho); \theta\right)$$
(3)

4. trained \mathcal{G} in step 2 was used to recover motion corrupted T1-Gd \tilde{x}_a from motion-free \hat{x} and motion-corrupted \hat{y}_a simulated in step 1 and 2,

$$\tilde{x}_a = \mathcal{G}(\mathcal{H}_{\nu_1}(\hat{y}_a), \mathcal{H}_{\nu_2}(\hat{x}); \varrho) \tag{4}$$

134 3.2 Learning

The MAUDGAN attempts to train generators in an adversarial scenario to achieve motion artifact disentanglement. Thus, the MAUDGAN employed loss functions to remove motion artifacts from T1-Gd using content information of T2-FLAIR as given in (1)-(4). The MAUDGAN employs four loss functions including two adversarial losses $\mathcal{L}_{adv}^{\mathcal{T}_a}$ and $\mathcal{L}_{adv}^{\mathcal{T}_a}$, reconstruction loss \mathcal{L}_{rec} , and artifact consistency loss \mathcal{L}_{arti} . The cost function is formalized as the weighted sum of the losses,

$$\mathcal{L} = \lambda_{adv} \times (\mathcal{L}_{adv}^{\mathcal{T}} + \mathcal{L}_{adv}^{\mathcal{T}_a}) + \lambda_{rec} \times \mathcal{L}_{rec} + \lambda_{arti} \times \mathcal{L}_{arti}$$
(5)

where λ_{adv} , λ_{rec} , and λ_{arti} are the hyper-parameters controlling the importance of each term.

141 3.2.1 Adversarial loss

The MAUDGAN was trained to map from motion-corrupted space to motion-free space as given in (1) and (3) and vice versa as given in (2) and (4). Learning those two tasks are important to disentangle motion artifact from the image content. As the MAUDGAN is trained on multimodal MRI sequences, regression losses like \mathcal{L}_1 and \mathcal{L}_2 could not be employed due to the domain difference between T2-FLAIR and T1-Gd MRI images. Therefore, the adversarial learning technique,²⁰ introduced $\mathcal{D}_{\mathcal{T}}$ and $\mathcal{D}_{\mathcal{T}_a}$ discriminators, was employed to regularize the plausibility between motion-corrected and motion-free images using $\mathcal{L}_{adv}^{\mathcal{T}}$ loss and between motion-corrupted

and motion-simulated images using $\mathcal{L}_{adv}^{\tau_a}$ loss. Thus, the MAUDGAN is trained to fool the discriminators, so they could not determine whether the motion was generated or real. The adversarial losses are as follows;

$$\mathcal{L}_{adv}^{\mathcal{T}} = \mathbb{E}_{z \sim q_{\theta}(z|\hat{y}_{a})} \left[\| \mathcal{D}_{\mathcal{T}}(z) - \mathbb{I} \|_{1} \right] + \mathbb{E}_{z \sim q_{\theta}(z|x_{a})} \left[\| \mathcal{D}_{\mathcal{T}}(z) \|_{1} \right]$$

$$\mathcal{L}_{adv}^{\mathcal{T}_{a}} = \mathbb{E}_{z \sim q_{\varrho}(z|\hat{y}_{a},\hat{x})} \left[\| \mathcal{D}_{\mathcal{T}a}(z) - \mathbb{I} \|_{1} \right] + \mathbb{E}_{z \sim q_{\varrho}(z|y,x_{a})} \left[\| \mathcal{D}_{\mathcal{T}a}(z) \|_{1} \right]$$
(6)

where z is the latent variable generators, $\mathcal{D}_{\mathcal{T}}$ and $\mathcal{D}_{\mathcal{T}_a}$ are the discriminators to distinguish between motion-corrupted and motion-free content data sampled from \mathcal{T} and \mathcal{T}_a domains, respectively. I is an unit matrix with a size $M \times M$, where M is substantially smaller than the image dimension size, that matched the discriminators' output.

156 3.2.2 Reconstruction loss

¹⁵⁷ Despite motion artifact disentanglement, the whole process needed to be lossless. In other words, ¹⁵⁸ the MAUDGAN was required to recover the original motion-corrupted T1-Gd \tilde{x}_a from motion-¹⁵⁹ corrected \hat{x} and to recover motion-free T2-FLAIR \tilde{y} from motion-simulated \hat{y}_a . Therefore, two ¹⁶⁰ reconstruction losses given in (7) were used to encourage the MAUDGAN to preserve the infor-¹⁶¹ mation.

$$\mathcal{L}_{rec} = \mathbb{E}_{x_a \sim \mathcal{T}_a, y \sim \mathcal{T}} \left[\|\tilde{x}_a - x_a\|_1 + \|\tilde{y} - y\|_1 \right] \tag{7}$$

where \tilde{y} and \tilde{x}_a shown in Figure 3 are given in (3) and (4). We adapted the \mathcal{L}_1 loss rather that than the \mathcal{L}_2 to generate sharper images.²¹



Fig 4: The Generator with the blocks used to construct discriminator and \mathcal{H}_{ν_s} are illustrated.

164 3.2.3 Artifact consistency loss

Adversarial losses encouraged the content of generated motion-corrupted \hat{y}_a and motion-free \hat{x} images to be indistinguishable from T1-Gd x_a and T2-FLAIR y images, respectively. However, the discriminators lose the spatial resolution. To preserve the spatial resolution, \mathcal{L}_1 and \mathcal{L}_2 could be used. But, due to the domain difference between T1-Gd and T2-FLAIR, direct use of losses would transfer the images' domain. We proposed artifact loss \mathcal{L}_{artif} given in (8) to induce motion artifacts to the motion-corrected \hat{x} images. Thus, \mathcal{L}_{artif} conflicts with adversarial losses and comprises the overall learning process.

$$\mathcal{L}_{artif} = \mathbb{E}_{x_a \sim \mathcal{T}_a, y \sim \mathcal{T}} \left[\| (\hat{x} - x_a) - (\hat{y}_a - y) \|_1 \right]$$
(8)

Equation (8) encourages the difference between x_a and \hat{x} to be similar to y and \hat{y}_a . Unlike a direct minimization by \mathcal{L}_1 that would cause an image domain translation, \mathcal{L}_{artif} requires the \hat{x} and x_a to be anatomically close rather be exactly close to preserve structural information.

175 3.3 Network architecture

The MAUDGAN network generator is illustrated in Figure 4-(a). The generator employed residual blocks²² (Figure 4-(b)) for a better generalization than convolution blocks without skip connection. To improve the generators' performance,²³ the convolution layers were used to down-sample the data in the encoder part of the generator. However, the decoder part of the generator employed the up-sampling layers rather than the transpose convolution layers to preserve the image edge information and avoid the checkerboard effect.²⁴

The discriminator consists of four residual blocks (Figure 4-(b)) and down-sampling blocks. Finally, the discriminators were constructed by four convolution blocks shown in Figure 4-(c) and the final layer with one convolution layer. The feature extractors (\mathcal{H}_{ν_i} for $i \in \{1, 2\}$) were constructed using five residual blocks (Figure 4-(b)).

We implemented the MAUDGAN under the PyTorch 1.12.0¹ deep learning framework using two NVIDIA GPUs RTX 3090. The batch size, optimizer, and the learning rate were 6, RAdam,²⁵ and 2×10^{-4} . We trained the network using hyper-parameters $\lambda_{rec} = 10$, $\lambda_{adv} = 5$, and $\lambda_{artif} = 50$.

190 4 Results

¹⁹¹ To our knowledge the MAUDGAN is the first network that employs the multi-modal MRI images ¹⁹² to reduce MRI motion artifacts. Thus, we could only compare the MAUDGAN with two well-¹⁹³ known unsupervised image-to-image translation approaches including CycleGAN²⁶ and Pix2pix.²¹

¹⁹⁴ The original implementations of the CycleGAN and Pix2pix were used² to compare the results.

¹https://pytorch.org/ ²https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix

The supervised methods like U-Net²⁷ were excluded since the ground truth targets were unavailable, and the domain shifts between the multi-modal images transfer the domain of the input motion-corrupted 3D T1-Gd images to the motion-free T2-FLAIR dataset. We compared the MAUDGAN with those networks for different motion artifact levels. Finally, we evaluated the performance of the MAUDAN to remove real motion artifacts from the patients with head & neck cancer.

Motion simulated dataset allowed us to perform qualitative and quantitative comparisons. We report six quantitative metrics including normalized mean squared error (NMSE), structural similarity index (SSIM),²⁸ multi-scale structural similarity index (MS-SSIM),²⁹ peak signal-to-noise ratio, visual information fidelity (VIF),³⁰ and multi-scale gradient magnitude similarity deviation (MS-GMSD).³¹ The higher metric values are better regarding motion artifact reduction and distortion levels except with the NMSE and MS-GMSD metrics.

Qualitative comparisons are illustrated in Figure 5 for different motion levels. Qualitatively, the Pix2Pix method had the lowest performance in preserving the MRI soft tissue contrast. CycleGAN reduced soft-tissue contrasts, smeared out the signal intensity, and unrealistically elevated the skull signals. MAUDGAN remove motion artifact with better soft tissue contrast and realistic skull signal intensity.

In addition, CycleGAN generated images with high signal intensity voxels mimicking the false tumors (see Figure 6). The false tumors were generated might be attributed to the wrong sampling from data manifolds. Those false tumors differ from water droplet-like artifacts³² cause by the normalization layers. Especially, the false tumor shown in Figure 6b is similar to the post-surgery cases.

²¹⁷ The quantitative metrics evaluating the motion-corrected image contrast, image distortion level,



Fig 5: Visual comparisons of the motion-reduction methods on the motion-simulated data. The simulated motion artifact was added along the row in (a) and column in (b). The heavy, moderate, and minor motion simulation data and the motion-corrected results are from top to bottom rows.



Fig 6: The white arrows illustrate the false tumors generated by the CycleGAN dataset.

and structure and texture similarity to the ground truth data are illustrated in Figure 7. The 218 MAUDGAN with the lowest NMSE and the highest PSNR values indicates the removing the 219 motion artifact with small spatial distortion. However, NMSE and PSNR tend to favor smooth-220 ness. The MS-SSIM and SSIM were reported to evaluate the structural similarity of the motion-221 corrected images and the ground truth. Higher MS-SSIM and SSIM indicate better similarity. Our 222 method got better SSIM values and comparable MS-SSIM values for different distortion levels. 223 The MAUDGAN with the highest value of VIF could preserve more information than the other 224 utilized methods. Finally, to evaluate the image gradient, which is related to image contrast, the 225 MS-GMSD was reported for different distortion levels. Lower MS-GMSD indicates a smaller de-226 viation between the gradients of motion-corrected and ground truth data. The MAUDGAN with 227 smaller MS-GMSD could preserve more, say soft-tissue, the contrast of the ground truth data. 228

We tested the MAUDGAN model on the data with real data with motion artifacts. The data were extracted anonymized from the PACS system. The real artifact was reduced using the MAUDGAN as shown in Figure 8.



Fig 7: Quantitative metrics to evaluate the quality of the motion-corrected data. The proposed MAUDGAN, Pix2Pix, and CycleGAN were evaluated on three motion distortion levels heavy, moderate, and minor.

232 5 Discussion

This study aimed to reduce 3D T1-Gd motion artifacts using T2-FLAIR images. 3D T1-Gd images with high acquisition times are more likely to corrupt with the motion artifact.² In addition, the high-resolution images' quality acquired with the high B0 magnetic fields is limited due to the motion artifact, which the PMC methods could partially remove the motion artifacts.¹ Motion artifacts reduce the image quality reducing the performance of manual and automatic post-processing approaches like tumor and organ at risks auto-segmentation.^{33,34} This study introduced MAUDGAN to tackle motion reduction as a disentanglement problem. The multi-center dataset with different



Fig 8: The anonymized data with real motion artifacts were exported from the PACS system to evaluate the MAUDGAN model to remove the real motion artifacts. The first row is the data with real artifact, and the second row illustrates the data after motion reduction. The arrows indicate the motion artifact.

brain tumors and metastases was used to train the MAUDGAN, which is expected to improve its 240 generalization. Our qualitative and quantitative comparisons with two well-known GAN methods 24 indicate that the MAUDGAN could disentangle the motion artifact using T2-FLAIR with a lower 242 spatial distortion and a better spatial contrast. 243 The MAUDGAN was qualitatively compared with generative models CycleGAN and Pix2pix.

244

The MAUDGAN could preserve better soft-tissue contrast (see Figure 5). The Pix2pix approach 245 did not preserve soft-tissue contrast, which might because this method was proposed to work under 246 the paired framework which is different from the theory of this study. On the other hand, the 247 CycleGAN smeared out the MRI soft-tissue contrast, which was better than the pix2pix. Finally, 248 the MAUDGAN reduced the motion artifact with better soft-tissue contrast. 249

When a network is trained on datasets with tumors, it is crucial that the network to be ro-250 bust against spatial distortions because those distortions could be misinterpreted as a tumor. The 25 MAUDGAN was free of spatial distortion, while the CycleGAN added spatial distortions (see Fig-252 ure 6). The added spatial distortions were similar to the brain tumor of the patient with edema and 253

after tumor resection as illustrated in Figure 6(a) and (b), respectively.

The quantitative comparisons shown in Figure 7 between the motion-free ground truth dataset and motion-corrected reconstructed by the CycleGAN, Pix2pix, and MAUDGAN suggest that the MAUDGAN-generated images were more distortion-free with a lower NMSE and a higher PSNR. In addition, MAUDGAN with the higher SSIM, MS-SSIM, and VIF and lower gradient deviations (MS-GMSD) generated more similar to the ground truth dataset.

To the best of our knowledge, this is the first study reporting on the feasibility of an approach enabling to disentangle motion of 3D T1-Gd using T2-FLAIR. The dataset contains different brain tumors and metastases, which are enhanced differently on the different MRI sequences. Thus, we did not use motion-free images of other patients, which need to be exported from PACS. This way, the dataset of the patients without motion artifacts remain in the clinical system. Moreover, we can use all the data to train the network, which is more than training under an unpaired scenario since we do not need to export the same number of patients' data without motion artifacts.

This study is more challenging compared with the unpaired studies^{14,35} because the data space domain of 3D T1-Gd differs from T2-FLAIR. Thus, the MAUDGAN must be robust to the domain shift between datasets. Due to the MAUDGAN's robustness, it could employ other image modalities like the T1-w dataset instead of T2-FLAIR. Thus, the MAUDGAN applies to other available MRI sequences than T2-FLAIR. However, this study is limited to the in-plane motion artifact due to the fact T2-FLAIR images were acquired in 2D that is inherently contain geometry distortion along the slice directions.³⁶

274 6 Conclusion

Our method, MAUDGAN, could disentangle motion artifacts from the 3D T1-Gd dataset under a multi-modal framework. The motion reduction will improve post-processing methods like manual and automatic brain tumors and organ at risk delineations and might increase the CT/MRI coregistration accuracy. Especially, the MAUDGAN would benefit elderly and infant patients with more involuntary motions during the 3D T1-Gd imaging with a long acquisition time. This retrospective motion correction is free from additional hardware or sequence modifications during the imaging, which makes it more practical.

282 Disclosures

²⁸³ There are no conflicts of interest declared by the authors.

284 Acknowledgments

²⁸⁵ This work was supported by NSERC CREATE RHHDS program and NSERC discovery grant.

286 Data Availability

- ²⁸⁷ The brain dataset was obtained from The Cancer Imaging Archieve
- 288 (https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=
- ²⁸⁹ 95224486).

290 References

- ²⁹¹ 1 D. Stucht, K. A. Danishad, P. Schulze, *et al.*, "Highest resolution in vivo human brain mri
- using prospective motion correction," *PloS one* **10**(7), e0133921 (2015).

293	2	J. M. Slipsager, S. L. Glimberg, J. Søgaard, et al., "Quantifying the financial savings of
294		motion correction in brain mri: a model-based estimate of the costs arising from patient
295		head motion and potential savings from implementation of motion correction," Journal of
296		<i>Magnetic Resonance Imaging</i> 52 (3), 731–738 (2020).
297	3	J. G. Pipe, "Motion correction with propeller mri: application to head motion and free-
298		breathing cardiac imaging," Magnetic Resonance in Medicine: An Official Journal of the
299		International Society for Magnetic Resonance in Medicine 42 (5), 963–969 (1999).
300	4	J. Maclaren, M. Herbst, O. Speck, et al., "Prospective motion correction in brain imaging: a
301		review," <i>Magnetic resonance in medicine</i> 69 (3), 621–636 (2013).
302	5	M. Zaitsev, J. Maclaren, and M. Herbst, "Motion artifacts in mri: A complex problem with
303		many partial solutions," Journal of Magnetic Resonance Imaging 42(4), 887–901 (2015).
304	6	W. Lin and H. K. Song, "Improved optimization strategies for autofocusing motion compen-
305		sation in mri via the analysis of image metric maps," Magnetic resonance imaging 24(6),
306		751–760 (2006).
307	7	M. W. Haskell, S. F. Cauley, and L. L. Wald, "Targeted motion estimation and reduction
308		(tamer): data consistency based motion mitigation for mri using a reduced model joint opti-
309		mization," IEEE transactions on medical imaging 37(5), 1253–1265 (2018).
310	8	M. Usman, D. Atkinson, F. Odille, et al., "Motion corrected compressed sensing for free-
311		breathing dynamic cardiac mri," Magnetic resonance in medicine 70(2), 504–516 (2013).
312	9	S. Kecskemeti, A. Samsonov, J. Velikina, et al., "Robust motion correction strategy for struc-
313		tural mri in unsedated children demonstrated with three-dimensional radial mpnrage," Radi-
314		ology 289 (2), 509 (2018).

315	10	A. Sciarra, S. Chatterjee, M. Dünnwald, et al., "Automated ssim regression for detection
316		and quantification of motion artefacts in brain mr images," arXiv preprint arXiv:2206.06725
317		(2022).
318	11	I. Oksuz, B. Ruijsink, E. Puyol-Antón, et al., "Deep learning using k-space based data
319		augmentation for automated cardiac mr motion artefact detection," in International Confer-
320		ence on Medical Image Computing and Computer-Assisted Intervention, 250–258, Springer
321		(2018).
322	12	T. Küstner, K. Armanious, J. Yang, et al., "Retrospective correction of motion-affected mr
323		images using deep learning frameworks," Magnetic resonance in medicine 82(4), 1527–1540
324		(2019).
325	13	B. A. Duffy, L. Zhao, F. Sepehrband, et al., "Retrospective motion artifact correction of struc-
326		tural mri images using deep learning improves the quality of cortical surface reconstructions,"
327		<i>Neuroimage</i> 230 , 117756 (2021).

- 14 G. Oh, J. E. Lee, and J. C. Ye, "Unpaired mr motion artifact deep learning using outlier rejecting bootstrap aggregation," *IEEE Transactions on Medical Imaging* 40(11), 3125–3139
 (2021).
- ³³¹ 15 F. Locatello, S. Bauer, M. Lucic, *et al.*, "Challenging common assumptions in the unsupervised learning of disentangled representations," in *international conference on machine learning*, 4114–4124, PMLR (2019).
- ³³⁴ 16 O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical im-
- age segmentation," in International Conference on Medical image computing and computer-
- assisted intervention, 234–241, Springer (2015).

20

337	17	N. Shusharina & T. Bortfeld, "Glioma image segmentation for radiotherapy: Rt targets, bar-
338		riers to cancer spread, and organs at risk [data set]," (2021). The Cancer Imaging Archive,
339		https://doi.org/10.7937/TCIA.T905-ZQ20.
340	18	B. Mortamet, M. A. Bernstein, C. R. Jack Jr, et al., "Automatic quality assessment in struc-
341		tural brain magnetic resonance imaging," Magnetic Resonance in Medicine: An Official Jour-
342		nal of the International Society for Magnetic Resonance in Medicine 62 (2), 365–372 (2009).
343	19	M. Torop, S. V. Kothapalli, Y. Sun, et al., "Deep learning using a biophysical model for robust
344		and accelerated reconstruction of quantitative, artifact-free and denoised images," Magnetic
345		resonance in medicine 84 (6), 2932–2942 (2020).
346	20	I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial networks," Com-
347		munications of the ACM 63(11), 139–144 (2020).
348	21	P. Isola, JY. Zhu, T. Zhou, et al., "Image-to-image translation with conditional adversarial
349		networks," in Proceedings of the IEEE conference on computer vision and pattern recogni-
350		tion, 1125–1134 (2017).
351	22	K. He, X. Zhang, S. Ren, et al., "Deep residual learning for image recognition," in Proceed-
352		ings of the IEEE conference on computer vision and pattern recognition, 770–778 (2016).
353	23	A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep con-
354		volutional generative adversarial networks," arXiv preprint arXiv:1511.06434 (2015).
355	24	A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard artifacts," Distill
356		1 (10), e3 (2016).
357	25	L. Liu, H. Jiang, P. He, et al., "On the variance of the adaptive learning rate and beyond,"
358		arXiv preprint arXiv:1908.03265 (2019).

359	26	JY. Zhu, T. Park, P. Isola, et al., "Unpaired image-to-image translation using cycle-
360		consistent adversarial networks," in Proceedings of the IEEE international conference on
361		computer vision, 2223–2232 (2017).
362	27	K. H. Jin, M. T. McCann, E. Froustey, et al., "Deep convolutional neural network for inverse
363		problems in imaging," IEEE Transactions on Image Processing 26(9), 4509–4522 (2017).
364	28	Z. Wang, A. C. Bovik, H. R. Sheikh, et al., "Image quality assessment: from error visibility
365		to structural similarity," IEEE transactions on image processing 13(4), 600–612 (2004).
366	29	Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality
367		assessment," in The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers,
368		2003, 2 , 1398–1402, Ieee (2003).
369	30	H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Transactions on
370		<i>image processing</i> 15 (2), 430–444 (2006).
371	31	B. Zhang, P. V. Sander, and A. Bermak, "Gradient magnitude similarity deviation on mul-
372		tiple scales for color image quality assessment," in 2017 IEEE International Conference on
373		Acoustics, Speech and Signal Processing (ICASSP), 1253–1257, IEEE (2017).
374	32	T. Karras, S. Laine, M. Aittala, et al., "Analyzing and improving the image quality of style-
375		gan," in Proceedings of the IEEE/CVF conference on computer vision and pattern recogni-
376		tion, 8110–8119 (2020).
377	33	P. Kemenczky, P. Vakli, E. Somogyi, et al., "Effect of head motion-induced artefacts on the
378		reliability of deep learning-based whole-brain segmentation," Scientific reports 12(1), 1-13
379		(2022).

22

380	4 N. Aldoj, F. Biavati, F. Michallek, et al., "Automatic prostate and prostate zones se	gmenta-
381	tion of magnetic resonance images using densenet-like u-net," Scientific reports 10(1), 1–17
382	(2020).	

- 383 35 S. Liu, K.-H. Thung, L. Qu, *et al.*, "Learning mri artefact removal with unpaired data," *Nature* Machine Intelligence 3(1), 60–67 (2021).
- 36 R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, *et al.*, *Magnetic resonance imaging: physical principles and sequence design*, ch. 20. John Wiley & Sons (2014).

387 List of Figures

388	1	Content and artifact components of 3D T1-Gd MRI (x_a) in the motion-corrupted
389		space \mathcal{T}_a and T2-FLAIR in artifact-free space \mathcal{T} are mapped to the content space
390		${\cal C}$ and artifact space ${\cal A},$ respectively. MAUDGAN maps the data in ${\cal T}_a$ space to ${\cal T}$
391		space $(x_a \rightarrow \hat{x})$ shown by blue arrows. Conversely, MADuGAN learns to map
392		from \mathcal{T} space to \mathcal{T}_a space $(y \to \hat{y}_a)$ shown by green arrows.
393	2	The motion simulation process. After choosing the phase encoding direction, sev-
394		eral random k -space regions were selected. The randomly selected k -space lines
395		were randomly translated within the random regions.
396	3	The proposed MAUDGAN is illustrated. The Generators \mathcal{F} learns disentanglement
397		while the \mathcal{G} learns to generate motion-corrupted images from motion-free images.
398	4	The Generator with the blocks used to construct discriminator and \mathcal{H}_{ν_s} are illus-
399		trated.

400	5	Visual comparisons of the motion-reduction methods on the motion-simulated data.
401		The simulated motion artifact was added along the row in (a) and column in (b).
402		The heavy, moderate, and minor motion simulation data and the motion-corrected
403		results are from top to bottom rows.
404	6	The white arrows illustrate the false tumors generated by the CycleGAN dataset.
405	7	Quantitative metrics to evaluate the quality of the motion-corrected data. The pro-
406		posed MAUDGAN, Pix2Pix, and CycleGAN were evaluated on three motion dis-
407		tortion levels heavy, moderate, and minor.
408	8	The anonymized data with real motion artifacts were exported from the PACS sys-
409		tem to evaluate the MAUDGAN model to remove the real motion artifacts. The
410		first row is the data with real artifact, and the second row illustrates the data after
411		motion reduction. The arrows indicate the motion artifact.