DEEP LEARNING FOR ACCELERATED 3D MRI

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

By
Muzaffer Özbey
August 2021

Deep Learning for Accelerated 3D MRI
By Muzaffer Özbey
August 2021

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Tolga Çukur(Advisor)

Hatice Kader Karlı Oğuz

Emine Ülkü Sarıtaş Çukur

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

DEEP LEARNING FOR ACCELERATED 3D MRI

Muzaffer Özbey
M.S. in Electrical and Electronics Engineering
Advisor: Tolga Çukur
August 2021

Magnetic resonance imaging (MRI) offers the flexibility to image a given anatomic volume under a multitude of tissue contrasts. Yet, scan time considerations put stringent limits on the quality and diversity of MRI data. The gold-standard approach to alleviate this limitation is to recover high-quality images from data undersampled across various dimensions, most commonly the Fourier domain or contrast sets. A primary distinction among recovery methods is whether the anatomy is processed per volume or per cross-section. Volumetric models offer enhanced capture of global contextual information, but they can suffer from sub-optimal learning due to elevated model complexity. Cross-sectional models with lower complexity offer improved learning behavior, yet they ignore contextual information across the longitudinal dimension of the volume. Here, we introduce a novel progressive volumetrization strategy for generative models (ProvoGAN) that serially decomposes complex volumetric image recovery tasks into successive cross-sectional mappings task-optimally ordered across individual rectilinear dimensions. ProvoGAN effectively captures global context and recovers fine-structural details across all dimensions, while maintaining low model complexity and improved learning behaviour. Comprehensive demonstrations on mainstream MRI reconstruction and synthesis tasks show that ProvoGAN yields superior performance to state-of-the-art volumetric and cross-sectional models.

Keywords: MRI, Generative Adversarial Networks, Synthesis, Reconstruction, Model complexity.

ÖZET

HIZLANDIRILMIŞ 3D MRG İÇİN DERİN ÖĞRENME

Muzaffer Özbey
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Danışmanı: Tolga Çukur
Ağustos 2021

Manyetik rezonans görüntüleme (MRG), çok sayıda doku kontrastı altında belirli bir anatomik hacmi görüntüleme esnekliği sunar. Fakat tarama zamanıyla alakalı endişeler MRG verisinin kalite ve çeşitliliği üzerinde keskin kısıtlamalara sebep olmaktadır. Bu kısıtlamaları kaldırmak için uygulanan güncel standart yaklaşım değişik boyutlardan, genellikle de Fourier alanı veya kontrast kümelerinden, alınan ve çarpazlama alt-örneklenen verilerden yüksek kalite görüntüler devşirmektir. Bu devşirme yöntemleri arasındaki başlıca fark anatominin hacimsel mi yoksa kesitsel mi işlendiğidir. Hacimsel modeller gelişmiş bütünlüklü bağlamsal malumat sağlarken fazlaca karmaşık modeller yüzünden yetersiz bir öğrenmeye sebep olur. Kesit bazlı modellerse daha az karmaşık olmaları sayesinde öğrenimde artış sağlarken hacmin boylamsal boyutlarındaki bağlamsal malumatı göz ardı ederler. Burada biz karmaşık hacimsel görüntülerin devşirilme görevini münferit doğrusal boyutlarda görev uygunluğunca sıralanmış, ardıl kesitli haritalamalara bolüştüren üretim modelleri (ProvoGAN) için yeni ve ileri bir hacimleştirme stratejisi sunmaktayız. ProvoGAN düşük model karmaşıklığını koruyup gelişmis öğrenimi temin ederken etkili biçimde bütünlüklü bağlam elde edebilmekte ve tüm boyutlar boyunca ince yapısal detayları devşirebilmektedir. Anaakım MRG geriçatım ve sentez görevleri üzerinde yapılan kapsamlı izahlar ProvoGAN'ın son teknoloji hacimsel ve kesitsel modellere göre üstün performans sağladığını göstermektedir.

Anahtar sözcükler: MRG, Üretici Çekişmeli Ağlar, Sentez, Geriçatım, Model karmaşıklığı.

Acknowledgement

First of all, I would like to thank my advisor, Prof. Tolga Çukur, for his consistent support and encouragement. He accepted me in his research group although I had no MRI background and he guided me with patience whenever I had any difficulties. Thanks to him, I had the opportunity to study in an outstanding research environment in the field of MRI and Deep Learning. Consequently I have become who I am today as a growing researcher candidate. Therefore I don't think I could have had a better advisor than him for my MS life.

I would also like to thank my jury members, Prof. Emine Ülkü Sarıtaş for her time and her teaching excellence in the classes in which I acquired detailed knowledge in medical imaging, and Prof. Kader Karlı Oğuz for her time and significant collaborations.

Next, I would like to thank all my fellow labmates. I was very luck to have Salman as a second mentor who thought me how to do things in lab from scratch. I have spent most of my time in the lab together with Mahmut and he fascinated me with his perfectionist endeavor and helpfulness, his supportive friendship and discussions on the ideas made my MS life better. I would like to thank Emin, Özgür, İbrahim, Yılmaz, Onat, Emrullah and Berk for their companionship and providing an amazing working environment.

I also want to thank my closest friends during both joy and sadness in these times; Kenan, Bozkurt, Ramazan, Mehmet, Safa, Abdulsamet, Ömer, Rahmi, Mahmut, Erdem, Bilal and Zafer.

My deepest gratitude must be paid to my dear parents. Even thouht I have been far away from them, without their unconditional love, support and guidance I could not have managed to accomplish this work. I have felt very honoured to be in this family.

I would like to thank TUBİTAK BİDEB-2210, 1001 Research Grant(118E256) for the funding and support they provided throughout this work.

Contents

1 Introduction **1**

2 Fundamentals of MRI **6**

3 Methods and Materials **9**

3.1 Generative Adversarial Networks 10

3.2 MR Image Recovery via Volumetric GANs 11

3.3 MR Image Recovery via Cross-Sectional GANs 12

3.4 Progressively Volumetrized GAN 13

3.5 Datasets 17

3.6 Competing Methods 20

3.7 Experiments 29

4 Results **33**

4.1 Task-Specific Progression Order 33

4.2 Accelerated MRI Reconstruction 36

4.3 Multi-Contrast MRI Synthesis 40

4.4 Demonstrations Against Hybrid Models 43

4.5 Radiological Evaluation 48

4.6 Complexity of Cross-Sectional Mappings 50

4.7 Generalizability of Progressive Volumetrization 56

5 Conclusion **58**

List of Figures

2.1 Sample $T1-$ and T_2- weighted images of the human brain 8

3.1 ProvoGAN decomposes complex volumetric image recovery tasks into a cascade of progressive cross-sectional subtasks defined across the rectilinear orientations (axial, coronal, and sagittal). Given a specific order of progression sequence (axial → sagittal → coronal is given here for demonstration), ProvoGAN first learns a cross-sectional mapping in the first orientation, and processes cross-sections within the entire source volume to estimate the target volume. This volumetric estimate is then divided into cross-sections in the second orientation, and a separate cross-sectional model is learned in the second orientation. The volumetric estimate from the second progression is then fed onto the final progression in which a third cross-sectional model is learned for final recovery. The sequential implementation of the progressive cross-sectional models enables ProvoGAN to gradually improve capture of fine-structural details in each orientation and to ensure global contextual consistency within the volume while at the same time manifesting reduced model complexity and improved learning behaviour of cross-sectional mapping. 18

3.2 The architectural details of the generator and discriminator submodules in ProvoGAN are displayed (here a progression order of A $\rightarrow$ S $\rightarrow$ C is used for illustration). The generator architectures are the same for the progressions and consist of an encoder with 3 convolutional layers: conv2D($k = 7, f = 24, s = 1, a =$ ReLU), conv2D($k = 3, f = 48, s = 2, a =$ ReLU), conv2D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet2D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv2D($k = 3, f = 48, s = 2, a =$ ReLU), deconv2D($k = 3, f = 24, s = 2, a =$ ReLU), conv2D($k = 7, f = 1, s = 1, a =$ Tanh), where k denotes kernel size, f denotes number of filters, s denotes stride, and a denotes activation function. Similarly, the discriminator architectures are identical for the progresions and consist of a convolutional network of 5 convolutional layers in series: conv2D($k = 4, f = 24, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 48, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv2D($k = 4, f = 1, s = 1, a =$ none). 27

3.3 a) The architectural details of the generator and discriminator in sGAN are displayed (here an sGAN model trained in the axial orientation referred to as sGAN-A is used for illustration). The generator consists of an encoder with 3 convolutional layers: conv2D($k = 7, f = 24, s = 1, a =$ ReLU), conv2D($k = 3, f = 48, s = 2, a =$ ReLU), conv2D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet2D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv2D($k = 3, f = 48, s = 2, a =$ ReLU), deconv2D($k = 3, f = 24, s = 2, a =$ ReLU), conv2D($k = 7, f = 1, s = 1, a =$ Tanh), where k denotes kernel size, f denotes number of filters, s denotes stride, and a denotes activation function. The discriminator consists of a convolutional network of 5 convolutional layers in series: conv2D($k = 4, f = 24, s = 2, a =$ leakyReLU), *conv2D*($k = 4, f = 48, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv2D($k = 4, f = 1, s = 1, a =$ none). b) The architectural details of the generator and discriminator in vGAN are displayed. The generator consists of an encoder with 3 convolutional layers: conv3D($k = 7, f = 24, s = 1, a =$ ReLU), conv3D($k = 3, f = 48, s = 2, a =$ ReLU), conv3D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet3D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv3D($k = 3, f = 48, s = 2, a =$ ReLU), deconv3D($k = 3, f = 24, s = 2, a =$ ReLU), conv3D($k = 7, f = 1, s = 1, a =$ Tanh). The discriminator consists of a convolutional network of 5 convolutional layers in series: conv3D($k = 4, f = 24, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 48, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv3D($k = 4, f = 1, s = 1, a =$ none). 28

4.1 The proposed ProvoGAN method is demonstrated on the IXI dataset for single-coil reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. Representative results are displayed for all competing methods together with the zero-filled (ZF) undersampled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed ProvoGAN method offers delineation of tissues with higher acuity compared to the volumetric (vGAN) model, and alleviates undesirable discontinuities compared to cross-sectional models (sGAN, RefineGAN, SparseMRI) by improving reconstruction performance in all orientations. 37

4.2 The proposed method is demonstrated on the in vivo multi-coil knee dataset for reconstruction at an acceleration ratio of $R = 8$. Representative results are displayed for all competing methods together with the zero-filled (ZF) undersampled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, ProvoGAN achieves sharper tissue depiction compared to vGAN, and alleviates undesirable discontinuities compared to cross-sectional models (sGAN, RefineGAN, SPIRiT) by improving reconstruction performance in all orientations. 39

4.3 The proposed method is demonstrated on the IXI dataset for T_2-weighted image synthesis from T_1- and PD-weighted images. Representative results are displayed for all competing methods together with the reference target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Overall, the proposed method delineates tissues with higher spatial resolution compared to volumetric vGAN, SC-GAN, and REPLICA models, and alleviates discontinuity artifacts by improving synthesis performance in all orientations compared to cross-sectional sGAN-A, sGAN-C, and sGAN-S models. 41

4.4 The proposed method is demonstrated on the in vivo brain dataset for T_1-weighted image synthesis from T_2-, T_{1c}-weighted and FLAIR images. Representative results are displayed for all competing methods together with the reference target images (first column). The first row displays results for the axial, the second row for the coronal, and the third row for the sagittal orientation. Overall, the proposed method delineates tissues with higher spatial resolution compared to volumetric vGAN, SC-GAN, and REPLICA models, and alleviates discontinuity artifacts by improving synthesis performance in all orientations compared to cross-sectional sGAN-A, sGAN-C, and sGAN-S models. Meanwhile, the proposed method achieves more accurate depictions for tumor regions, which are suboptimally recovered by the competing methods. 42

4.5 The proposed ProvoGAN method is demonstrated on the IXI dataset against hybrid models (M^{3}NET and TransferGAN) for T_1-weighted image synthesis from T_2- and PD-weighted images. Representative results are displayed for all methods under comparison together with the ground truth target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed method offers sharper and more accurate delineation of tissues than the competing methods. Furthermore, ProvoGAN better alleviates residual discontinuity artifacts compared to M^{3}NET. 45

4.6 The proposed ProvoGAN method is demonstrated against hybrid models (M^{3}NET and TransferGAN) on the IXI dataset for reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. Representative results are displayed for the methods under comparison together with the undersampled zero-filled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed method enables sharper tissue delineation against competing methods, and it improves mitigation of discontinuity artifacts compared to M^{3}NET. 46

4.7 Methods were compared in terms of radiological opinion scores for three reconstruction tasks: a) single-coil reconstruction of T_1-weighted images undersampled by $R = 8$ in the IXI dataset, b) single-coil reconstruction of T_2-weighted images undersampled by $R = 8$ in the IXI dataset, c) multi-coil reconstruction of PD-weighted images undersampled by $R = 8$ in the in vivo knee dataset, and for two synthesis tasks: d) many-to-one synthesis task of T_2, PD $\rightarrow$ T_1 in the IXI dataset, e) many-to-one synthesis task of T_1, T_2, T_{1c} $\rightarrow$ FLAIR in the in vivo brain dataset. The quality of the recovered axial, coronal, and sagittal cross-sections were rated by an expert radiologist by assessing their similarity to the reference cross-sections via a five-point scale (0: unacceptable, 1: very poor, 2: limited, 3: moderate, 4: good, 5: perfect match). Figure legends denote the colors used for marking the methods under comparison. 49

4.8 Volumetrization approaches were compared in terms of radiological opinion scores for three representative tasks: a) reconstruction of T_1-weighted images undersampled by $R = 8$ in the IXI dataset, b) reconstruction of T_2-weighted images undersampled by $R = 8$ in the IXI dataset, c) T_2, PD $\rightarrow$ T_1 synthesis in the IXI dataset. The quality of the recovered axial, coronal, and sagittal cross-sections were rated by an expert radiologist by assessing their similarity to the reference cross-sections via a five-point scale (0: unacceptable, 1: very poor, 2: limited, 3: moderate, 4: good, 5: perfect match). Figure legend denotes the colors used for the methods under comparison. 50

4.9 Progressive volumetrization was performed on recently proposed SC-GAN architecture. Representative results for T_1-weighted image synthesis from T_2- and PD-weighted images in the IXI dataset are displayed. Results are shown for progressively volumetrized (ProvoSC-GAN), cross-sectional (sSC-GAN), and volumetric (vSC-GAN) models, along with the ground truth target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Overall, ProvoSC-GAN improves delineation of structural details compared to vSC-GAN, and enhances contextual consistency in the longitudinal dimensions compared to sSC-GAN models. 57

List of Tables

4.1 Task-Optimal Progression Order for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) measurements between the reconstructed and ground truth images in the validation set in the IXI dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and acceleration factors: $R = 4, 8, 12, 16$. Boldface indicates the highest performing progression sequence. 34

4.2 Task-Optimal Progression Order for Reconstruction in the In vivo Knee Dataset: Volumetric PSNR (dB) measurements between the reconstructed and ground truth images in the validation set in the in vivo knee dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and acceleration factors: $R = 4, 8, 12, 16$. Boldface indicates the highest performing progression sequence. 35

4.3 Task-Optimal Progression Order for Synthesis in the IXI Dataset: Volumetric PSNR (dB) measurements between the synthesized and ground truth images in the validation set in the IXI dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and all many-to-one synthesis tasks: 1) T_2, PD → T_1, 2) T_1, PD → T_2, 3) T_1, T_2 → PD. Boldface indicates the highest performing progression sequence. 35

4.4 Task-Optimal Progression Order for Synthesis in the In vivo Brain Dataset: Volumetric PSNR (dB) measurements between the synthesized and ground truth images in the validation set in the in vivo brain dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and all many-to-one synthesis tasks: 1) T_2, FLAIR, T_{1c} → T_1, 2) T_1, FLAIR, T_{1c} → T_2, 3) T_1, T_2, T_{1c} → FLAIR, 4) T_1, T_2, FLAIR → T_{1c}. Boldface indicates the highest performing progression sequence. 35

4.5 Quality of Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the IXI dataset are given as mean ± std for the test set. The measurements are reported for zero-filled images (ZF), the proposed ProvoGAN and competing sGAN, vGAN, RefineGAN, and SparseMRI reconstruction methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method. 38

4.6 Quality of Reconstruction in the In vivo Knee Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the in vivo knee dataset are given as mean ± std for the test set. The measurements are reported for zero-filled images (ZF), the proposed ProvoGAN and competing sGAN, vGAN, RefineGAN, and SPIRiT methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method. 40

4.7 Quality of Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are provided for the proposed and competing methods for all synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method. 42

4.8 Quality of Synthesis in the In vivo Brain Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set of the in vivo brain dataset are given as mean ± std. The measurements are provided for proposed and competing methods for all many-to-one synthesis tasks: 1) T_2, FLAIR, T_{1c} $\rightarrow$ T_1, 2) T_1, FLAIR, T_{1c} $\rightarrow$ T_2, 3) T_1, T_2, T_{1c} $\rightarrow$ FLAIR, 4) T_1, T_2, FLAIR $\rightarrow$ T_{1c}. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method. 43

4.9 Comparison of Volumetrization Approaches for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN and competing M^{3}NET and TransferGAN methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method. 45

4.10 Comparison of Volumetrization Approaches for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN and competing M^{3}NET and TransferGAN methods for all synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. Boldface indicates the highest performing method. 47

4.11 Model Complexity Analysis for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set of the IXI dataset are given as mean ± std for reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. The measurements are reported for ProvoGAN and sGAN while varying the complexity of the convolutional layers in both models by $n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$, yielding seven distinct ProvoGAN-sGAN pairs: ProvoGAN(n_f)-sGAN(n_f) with n_f times folded number of learnable network weights. Boldface indicates the highest performing method. 52

4.12 Model Complexity Analysis for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set of the IXI dataset are given as mean ± std for T_1-weighted image synthesis from T_2- and PD-weighted images. The measurements are reported for ProvoGAN and sGAN while varying the complexity of the convolutional layers in both models by $n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$, yielding seven distinct ProvoGAN-sGAN pairs: ProvoGAN(n_f)-sGAN(n_f) with n_f times folded number of learnable network weights. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method. 53

4.13 Comparison of Single versus Multi-Cross-Section ProvoGAN and sGAN Models for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set of the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN model, its multi-cross-section variant, the cross-sectional sGAN model and its multi-cross-section variant for all single-coil reconstruction tasks: T_1-weighted and T_2-weighted image reconstruction at $R = 4, 8, 12, 16$. sGAN and sGAN(multi) are trained in the axial orientation given axial readout direction. Boldface indicates the highest performing method. 53

4.14 Comparison of Single and Multi-Cross-Section ProvoGAN and sGAN Models for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN model, its multi-cross-section variant, the cross-sectional sGAN model and its multi-cross-section variant for all many-to-one synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sGAN-A and sGAN(multi)-A denote the models trained in the axial orientation, sGAN-C and sGAN(multi)-C denote the models trained in the coronal orientation, and sGAN-S and sGAN(multi)-S denote the models trained in the sagittal orientation. Boldface indicates the highest performing method. . . 54

4.15 The proposed ProvoGAN and competing sGAN and vGAN methods are evaluated in terms instantaneous model complexity (millions of parameters, M), GPU VRAM use (gigabytes, GB), FLOPs (billions of floating point operation, G), and train duration (hours). The model complexity of the methods is given as (p_g, p_d), where p_g denotes the number of free parameters in the generator and p_d denotes the number of free parameters in the discriminator. The number of operations (FLOPs) of the methods is given as (F_g, F_d), where F_g denotes the number of operations in the generator and F_d denotes the number of operations in the discriminator. The input-output volume sizes and the number training subjects are also reported. 55

4.16 Progressive Volumetrization of the SC-GAN Architecture: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. Measurements are provided for proposed and competing methods for all many-to-one synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sSC-GAN-A denotes the sSC-GAN model trained in the axial orientation, sSC-GAN-C in the coronal orientation, and sSC-GAN-S in the sagittal orientation. Boldface indicates the highest performing method. . . 57

Chapter 1

Introduction

Magnetic resonance imaging (MRI) is a clinically preferred modality that produces volumetric images of a given anatomy under diverse tissue contrasts [1]. As MR acquisitions are intrinsically slow, there has been persistent interest in recovery methods to improve quality and diversity of images derived from accelerated imaging protocols [2, 3]. Two mainstream MRI recovery problems with pervasive applications are reconstruction and synthesis [4–27]. While reconstruction aims to recover high-quality images from undersampled k-space acquisitions [28], synthesis aims to recover high-quality images of unacquired tissue contrasts from images of collected contrasts [29]. Learning-based models have offered performance leaps in both recovery tasks, given their ability to solve inverse problems [30–32]. However, the trade-off between sensitivity to spatial context and model complexity introduces a dilemma regarding the use of volumetric versus cross-sectional recovery models [33]. The primary aim of this study is to introduce a novel volumetrization approach to achieve the contextual sensitivity of volumetric models while maintaining on par complexity with cross-sectional models.

Among learning-based models, a native recovery approach is to perform a single-shot global mapping between source and target volumes [34–43]. Volumetric models leverage spatial correlations across all dimensions to better capture

contextual information [34–37]. Introduction of these contextual priors can theoretically lead to more consistent and accurate recovery across the volume. However, three-dimensional (3D) models involve substantially more parameters than their two-dimensional (2D) counterparts [8,33]. Furthermore, each volume constitutes a single training sample for a 3D model, whereas it would yield several tens of samples for a 2D model. Taken together, these factors render heavier demand for training data and impair the learning process for volumetric models [33].

A less demanding approach in terms of training data for learning-based MRI recovery is to perform a spatially-localized mapping between individual cross-sections [44–47, 47–58]. Volumes are split along a specific rectilinear orientation, and cross-sectional models are then trained to learn the 2D mapping [59–64]. Since a lower-dimensional mapping is to be learned, cross-sectional models are of lower complexity and have reduced demand for training data [8]. This facilitates the learning process, and often results in more detailed mappings along the transverse dimensions within cross-sections compared to 3D models. Yet, 2D models do not fully utilize context across the longitudinal dimension, even when simultaneously processing multiple neighboring cross-sections [36, 37, 59]. This results in inconsistency and errors across separately recovered cross-sectional images [34, 35, 38].

An effective alternative to either approach is to build hybrid architectures that bridge 2D and 3D models. A group of studies in this domain have proposed aggregated models that fuse the outputs of parallel streams, where the streams are cross-sectional models in three orthogonal orientations [65, 66]. Pseudo target volumes are first recovered separately by the 2D streams, and a 3D fusion network then produces the final target volume [65, 66]. Other studies have instead proposed transfer of learned model weights from 2D to 3D models [67, 68]. A 2D model is first pretrained for a cross-sectional recovery task at a selected orientation, the learned weights are then used to initialize the convolutional kernels in 3D models [67, 68]. While both approaches can improve learning behavior, they involve a volumetric processing component that elevates memory requirements and places practical constraints on model complexity, potentially limiting sensitivity to detailed image features.

Contributions

In this thesis, we propose a novel progressive volumetrization strategy for deep generative models (ProvoGAN) for contextual learning of MR image recovery. To improve learning efficiency by lowering model complexity, ProvoGAN serially decomposes volumetric recovery tasks into a sequence of cross-sectional subtasks (e.g., axial, coronal, sagittal) for the first time in literature. For a given subtask in a selected orientation, the source volume is split across the respective longitudinal dimension, and a 2D model is trained to map between cross-sectional source and target images. The predicted pseudo cross-sections are reformatted into a volume and then input to the next subtask as spatial priors (Fig. 3.1). This progressive nature empowers ProvoGAN to recover fine-structural details in each orientation while ensuring contextual consistency across the volume. Furthermore, the progression order of the subtasks is adaptively optimized to enhance task-specific performance. To ensure a high degree of realism, we primarily employ ProvoGAN to volumetrize a recent conditional generative adversarial network based on the ResNet architecture [59]. Note that ProvoGAN can be viewed as a model-agnostic strategy, so it can be extended to volumetrize other 2D network models as also demonstrated here. Comprehensive demonstrations are provided for mainstream reconstruction and synthesis tasks in multi-contrast MRI protocols. Our results indicate that ProvoGAN yields enhanced recovery performance compared to cross-sectional, volumetric, and hybrid approaches in terms of image quality. Importantly, ProvoGAN maintains these performance benefits while at the same time offering reduced model complexity and improved learning behaviour.

- To our knowledge, ProvoGAN is the first volumetrized model for MRI recovery that serially decomposes a global 3D mapping into a sequence of progressive 2D mappings.
- ProvoGAN maximizes task performance via adaptive ordering of the progression sequence of 2D mappings across rectilinear orientations.
- ProvoGAN embodies a model-agnostic learning strategy, so it can be implemented to volumetrize various 2D network architectures.

- Demonstrations on mainstream reconstruction and synthesis tasks indicate that ProvoGAN yields superior performance to several prior 2D, 3D and hybrid models.

Outline

For the rest of this thesis, the organization is as follows. Chapter 2 discusses the fundamentals of MRI. Chapter 3 describes the methods, used date-set and experiments in details. Chapter 4 demonstrates the outcome of the experiments and comparison results. Lastly, Chapter 5 presents the conclusive remarks.

Chapter 2

Fundamentals of MRI

Magnetic resonance imaging (MRI) depands on atomic nuclei (^1H) which has both magnetic and angular momentum. The hydrogen nuclei are aligned based on the magnetic field. If there is not any applied external magnetic field, the hydrogen nuclei are randomly aligned and thereby net magnetization approximating to zero. If an external magnetic field (B_0) is applied, it causes that the hydrogen nuclei start to precessing around the longitudinal direction of the applied B_0 with the frequency of $\omega_0 = B_0\gamma$ which known as the Larmor frequency. For MRI scans, an external radio-frequency (RF) field pulse, which designated as B_1, can tips the net magnetization of the nuclei to the transverse plane up to a desired degree which known as the lip angle. After turning off of the external RF field, the net magnetization tend to return its equilibrium state. For this relaxation period, the transverse part of magnetization decays with a time constant T_2 and the longitudinal part of the magnetization recovers with a time constant T_1. In MRI, the receiver coils detect the flux changes along transverse direction and with these flux changes, the signal in the transverse plane can be detected. Meanwhile, for spatial encoding of the signal, additional magnetic field gradients (G) are performed. Overall, the received signal by the coil is given as:

$$s(t) = \int_x \int_y \int_z M^0(x,y,z) e^{-t/T_2(r)} exp(-i\gamma \int_0^t G(\tau).r d\tau)\, dx\, dy\, dz \tag{2.1}$$

where $M(x, y, z)$ is transverse magnetization at point (x,y,z), γ is gyromagnetic ratio, $r = [x, y, z]$ and $G = [G_x, G_y, G_z]$. To obtain a frequency domain(k-space) formulation, equation can be rearranged by expressing the gradient terms with spatial frequency variables $k_x(t) = \frac{\gamma}{2\pi} \int G_x(\tau)\, d\tau$, $k_y(t) = \frac{\gamma}{2\pi} \int G_y(\tau)\, d\tau$ and $k_z(t) = \frac{\gamma}{2\pi} \int G_z(\tau)\, d\tau$:

$$s(t) = \int_x \int_y \int_z M^0(x, y, z) e^{-t/T_2(r)} exp(-i2\pi(k_x(t) + k_y(t) + k_z(t))\, dx\, dy\, dz \quad (2.2)$$

Note that transverse magnetization can be manipulated via the applied RF pulse and gradient fields based on the tissue-specific relaxation parameters. With this manipulation, desired properties of the tissues can be imaged and this enables the multi-contrast MRI acquisition. The obtained MRI signal for one of the most regularly used spin-echo sequences can be represented as:

$$S_0 = K_0 \rho(r)[1 - e^{-TR/T_1(r)}]e^{TE/T_2(r)} \quad (2.3)$$

where K_0 is a scaling factor, $\rho(r)$ is proton density, TR is the repetition time, and TE are the echo time of RF pulse and $T_1(r)$ and $T_2(r)$ are relaxation parameters of longitudinal and transverse magnetization. This equation shows how to alter the contrast in acquired images by simply modifying TR and TE. T_1-weighted contrast images, for example, can be produced by a low TE and a moderate TR. Similarly, T_2-weighted contrast images can be acquired by a moderate TE and long TR duration. A sample $T1-$ and T_2- weighted images are shown in Figure 2.1

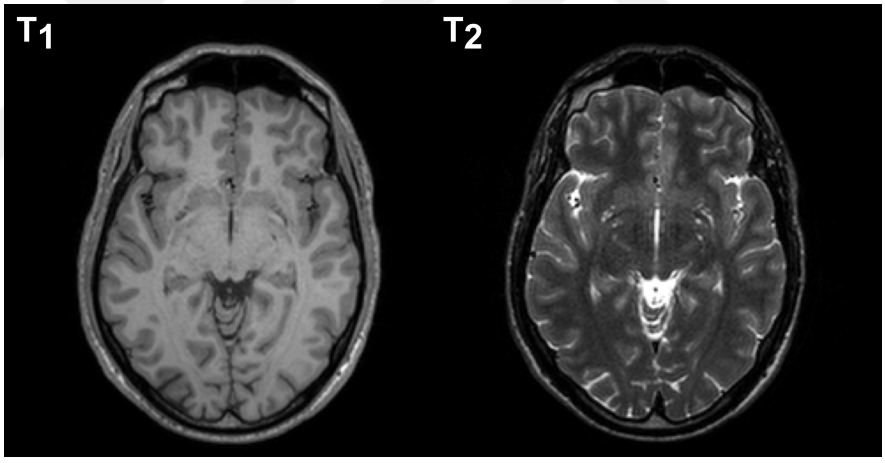


Figure 2.1: Sample $T1-$ and T_2- weighted images of the human brain

Chapter 3

Methods and Materials

The content of this chapter reflects the study described in the following publications:

- M. Özbey, M. Yurt, S. U. H. Dar, and T. Çukur, "Three Dimensional MR Image Synthesis with Progressive Generative Adversarial Networks," in *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, (Virtual Conference),4 2020
- M. Özbey, M. Yurt, S. U. H. Dar, and T. Çukur, "Three dimensional mr image synthesis with progressive generative adversarial networks," *arXiv preprint arXiv:2101.05218*, 2020
- M. Yurt, M. Özbey, S. U. H. Dar, B. Tınaz, K. K. Oğuz, and T. Çukur, "Progressively volumetrized deep generative models for data-efficient contextual learning of mr image recovery," *arXiv preprint arXiv:2011.13913*, 2020
- M. Yurt, M. Özbey, S. U. H. Dar, B. Tınaz, K. K. Oğuz, and T. Çukur, "Progressively Volumetrized Deep Generative Models for Data-Efficient Contextual Learning of MR Image Recovery," in *29th Annual Meeting of International Society for Magnetic Resonance Imaging (ISMRM)*, (Virtual Conference),5 2021

3.1 Generative Adversarial Networks

Generative adversarial networks (GAN) are generative models composed of two subnetworks. The first subnetwork is a generator (G) that aims to synthesize fake samples closely mimicking a target data distribution, while the second subnetwork is a discriminator (D) that aims to detect whether a given data sample has been drawn from the target distribution or not [73]. These subnetworks are trained alternately in a two player zero-sum min-max game in an adversarial setup:

$$L_{\text{GAN}} = \text{E}_y[\log(D(y))] + \text{E}_z[\log(1 - D(G(z)))] \tag{3.1}$$

where L_{GAN} is the adversarial loss function, E denotes expectation, z denotes a random noise vector sampled from a prior distribution, and y denotes an arbitrary real sample drawn from the target domain. In practice, the log-likelihood terms are replaced with squared-loss terms to improve stability [74]:

$$L_{\text{GAN}} = -\text{E}_y[(D(y) - 1)^2] - \text{E}_z[D(G(z))^2] \tag{3.2}$$

where D is trained to maximize L_{GAN}, whereas G is trained to minimize it.

While the basic GAN model synthesizes target data samples given a random noise input, recent studies on computer vision [75, 76] and medical imaging [37, 38, 59, 63, 77–80] have demonstrated that conditional GAN (cGAN) models [81] are highly effective in image-to-image translation tasks. The central aim in these tasks is to synthesize data samples from the target image domain, given data samples from a separate source image domain. The cGAN model is therefore modified to condition both G and D on the source domain image:

$$L_{\text{cGAN}} = -\text{E}_{x,y}[(D(x, y) - 1)^2] - \text{E}_x[D(x, G(x))^2] \tag{3.3}$$

where x denotes the source domain image, and y denotes the target domain image. When paired images from the source and target domains are available, a pixel-wise loss between the ground truth and synthesized images can also be included:

$$\begin{aligned} L_{\text{cGAN}} = &- \mathrm{E}_{x,y}[(D(x,y)-1)^2] - \mathrm{E}_x[D(x,G(x))^2] \\ &+ \mathrm{E}_{x,y}[||y - G(x)||_1] \end{aligned} \tag{3.4}$$

The pixel-wise loss is typically based on the mean-absolute error to reduce sensitivity to outliers and alleviate undesirable smoothing. The mapping learned by the cGAN model grows more accurate as the statistical dependence between source and target domains gets stronger [8].

3.2 MR Image Recovery via Volumetric GANs

As MR images are intrinsically volumetric, a comprehensive approach for 3D MR image recovery is to use volumetric GAN (vGAN) models that perform a global mapping between source and target volumes [34–37]. To learn this mapping, vGAN models commonly employ complex generator G_V and discriminator D_V modules containing 3D convolutional kernels. The loss function is defined over the entire volume in an adversarial setup:

$$\begin{aligned} L_{\text{vGAN}} = &- \mathrm{E}_{X,Y}[(D_V(X,Y)-1)^2] - \mathrm{E}_X[D_V(X,G_V(X))^2] \\ &+ \mathrm{E}_{X,Y}[||Y - G_V(X)||_1] \end{aligned} \tag{3.5}$$

where X denotes the source and Y denotes the target volumetric images. For MRI reconstruction, X is typically the Fourier reconstruction of undersampled acquisitions, and Y is the fully-sampled reference volume. For MRI synthesis, X is the source contrast volume, and Y is the target contrast volume. Note that, in MRI reconstructions, an additional constraint is introduced to enforce consistency of acquired and recovered k-space data:

$$F_u(G_V(X)) := F_u(X) \tag{3.6}$$

where F_u denotes the partial Fourier operator that is defined at the acquired k-space points.

Due to their 3D nature, vGAN models can better incorporate contextual information across MRI volumes by leveraging spatial correlations across separate cross-sections [36–38]. This contextual prior can lead to elevated consistency across the volume and increased accuracy in recovery performance. That said, learning in 3D network models is inherently more difficult since they involve substantially more parameters [33]. The learning process might be further impaired by data scarcity as the entire volume of each subject is taken as a single training sample [33]. These limitations often cause vGAN models to settle on suboptimal parameter sets, compromising recovery performance.

3.3 MR Image Recovery via Cross-Sectional GANs

A more focused approach for 3D MRI recovery is based on cross-sectional GAN (sGAN) models that perform localized mappings between 2D cross-sectional images within source and target volumes [8,59,79,82]. These 2D images are typically taken to be individual cross-sections within the volume in a specific rectilinear orientation, i.e., axial, sagittal or coronal. To learn this 2D mapping, sGAN models employ relatively simpler generator G_S and discriminator D_S modules containing 2D convolutional kernels. The loss function is defined for individual cross-sections in an adversarial setup with a pixel-wise loss:

$$\begin{aligned} L_{\text{sGAN}} = &- \mathrm{E}_{x_o^i, y_o^i}[(D_S(x_o^i, y_o^i) - 1)^2] - \mathrm{E}_{x_o^i}[D_S(x_o^i, G_S(x_o^i))^2] \\ &+ \mathrm{E}_{x_o^i, y_o^i}[||y_o^i - G_S(x_o^i)||_1] \end{aligned} \tag{3.7}$$

where x_o^i and y_o^i denote the ith cross-sections within the source and target volumes in orientation o. As with sGAN models, x_o^i-y_o^i are taken as cross-sectional images for undersampled and fully-sampled acquisitions in MRI reconstruction, and x_o^i-y_o^i are taken as cross-sectional images of source and target contrasts in MRI synthesis. Consistency between acquired and recovered data can again be enforced during reconstruction via the following procedure:

$$F_u(G_S(x_o^i)) := F_u(x_o^i) \tag{3.8}$$

where F_u denotes the partial Fourier operator that is defined at the acquired k-space points. Once the mapping between the source and target cross-sections is learned, cross-sections of the target volumes are independently generated, and then the target volumes are recovered by concatenating the generated cross-sections.

Due to their 2D nature, sGAN models are less complex and so they naturally have lower demand for data [8]. Individual cross-sections within a subject's volume are taken as separate training samples, expanding the effective size of the dataset. As a result, more detailed cross-sectional mapping can be learned. However, this advantage comes at the expense of neglecting global contextual information across the volume [36–38]. Therefore, sGAN models might suffer from inconsistency or inaccuracy of recovered images across cross-sections.

3.4 Progressively Volumetrized GAN

Here, a novel architecture is proposed to address the limitations of volumetric and cross-sectional GAN models. The proposed model, named progressively volumetrized GAN (ProvoGAN), decomposes complex volumetric image recovery tasks into a series of simpler cross-sectional tasks (Fig. 3.1). The cross-sectional

recovery tasks are defined in separate orientations, and are implemented sequentially via cascaded 2D GAN models. We consider rectilinear cross-sections of volumetric MRI datasets in this study, so the selected orientations are axial, coronal and sagittal. Given a specific order of the three orientations (o_1, o_2, o_3), ProvoGAN first learns a 2D recovery model in orientation o_1. The entire source volume is processed by this model to estimate the target volume. Afterwards, this volumetric estimate is separated into cross-sections in orientation o_2, and a separate 2D recovery model is trained. The estimated target volume for o_2 is then fed onto the final stage, where a third 2D recovery model is trained in orientation o_3.

The cascaded 2D models in ProvoGAN are trained sequentially in the three rectilinear orientations, where the 2D model weights at earlier orientations are frozen upon training. This learning strategy empowers ProvoGAN to progressively recover fine-structural details at each orientation, while bypassing the need for computationally expensive calculation of error gradients across the entire volume and across all orientations. Therefore, ProvoGAN offers the ability to efficiently capture global contextual information without drastically elevating computational demand. At the same time, this step-wise training can increase sensitivity to progression order. Therefore, progression order across orientations is adaptively tuned to maximize performance in specific tasks. Detailed formulation of the ProvoGAN model is provided below.

First Progression: ProvoGAN first learns a cross-sectional mapping between the source-target volumes in o_1 via a generator (G_{o_1}) and a discriminator (D_{o_1}). The source and target cross-sections in o_1 are extracted with a division block (d_{o_1}).

$$\begin{aligned} x^i_{o_1} \in \{x^1_{o_1}, x^2_{o_1} \ldots, x^I_{o_1}\} = d_{o_1}(X) \\ y^i_{o_1} \in \{y^1_{o_1}, y^2_{o_1} \ldots, y^I_{o_1}\} = d_{o_1}(Y) \end{aligned} \tag{3.9}$$

where X denotes the source volume, Y denotes the target volume, $x^i_{o_1}$ denotes the ith cross-section of the source volume in o_1, $y^i_{o_1}$ denotes the ith cross-section of the target volume in o_1, and I denotes the total number of cross-sections within the volumes in o_1. G_{o_1} then learns to recover the cross-sections of the target

volume from the corresponding cross-sections of the source volume.

$$\hat{y}^i_{p_1,o_1} = G_{o_1}(x^i_{o_1}) \tag{3.10}$$

where $\hat{y}^i_{p_1,o_1}$ denotes the ith cross-section of the target volume in o_1 recovered via the first progression. Meanwhile, D_{o_1} learns to distinguish between the real and fake cross-sections.

$$D_{o_1}(x^i_{o_1}, m) \in [0, 1] \tag{3.11}$$

where m is either generated ($\hat{y}^i_{p_1,o_1}$) or ground truth ($y^i_{o_1}$) target cross-section. To simultaneously train G_{o_1} and D_{o_1}, a loss function (L_{o_1}) consisting of adversarial and pixel-wise losses is used.

$$\begin{aligned} L_{o_1} = &- \mathrm{E}_{x^i_{o_1}, y^i_{o_1}}[(D_{o_1}(x^i_{o_1}, y^i_{o_1}) - 1)^2] \\ &- \mathrm{E}_{x^i_{o_1}}[D_{o_1}(x^i_{o_1}, G_{o_1}(x^i_{o_1}))^2] \\ &+ \mathrm{E}_{x^i_{o_1}, y^i_{o_1}}[||y^i_{o_1} - G_{o_1}(x^i_{o_1})||_1] \end{aligned} \tag{3.12}$$

Once G_{o_1} and D_{o_1} are properly trained, cross-sections in o_1 for the target volume are independently generated, and then combined with a concatenation block (c_{o_1}) to recover the entire target volume.

$$\hat{Y}_{p_1} = c_{o_1}(\hat{y}^1_{p_1,o_1}, \ldots, \hat{y}^I_{p_1,o_1}) \tag{3.13}$$

where $\hat{Y}_{p_1}$ denotes the target volume recovered after the first progression.

Second Progression: Having learned the cross-sectional mapping in o_1, Provo-GAN then learns a separate recovery model in the second orientation o_2 to gradually enhance capture of fine-structural details and spatial correlations. The prediction for the target volume generated in the first progression is also incorporated as an input to the generator G_{o_2} to leverage global contextual priors.

$$\hat{y}^{j}_{p_2,o_2} = G_{o_2}(x^{j}_{o_2}, \hat{y}^{j}_{p_1,o_2}) \tag{3.14}$$

where $x^{j}_{o_2}$ denotes the jth cross-section of the source volume in o_2, $\hat{y}^{j}_{p_1,o_2}$ denotes the jth cross-section in o_2 of the target volume recovered in the first progression, and $\hat{y}^{j}_{p_2,o_2}$ denotes the jth cross-section in o_2 of the target volume recovered in the second progression. Meanwhile, discriminator D_{o_2} learns to distinguish between the generated and real cross-sections.

Third Progression: Lastly, ProvoGAN learns a cross-sectional mapping in the third orientation o_3. As in the second progression, the prediction from the previous progression is incorporated into the mapping as prior information. Therefore, the third generator G_{o_3} receives as input the cross-sections in o_3 of the source volume and the previously recovered volume:

$$\hat{y}^{k}_{p_3,o_3} = G_{o_3}(x^{k}_{o_3}, \hat{y}^{k}_{p_2,o_3}) \tag{3.15}$$

where $x^{k}_{o_3}$ denotes the kth cross-section of the source volume in o_3, $\hat{y}^{k}_{p_2,o_3}$ denotes the kth cross-section in o_3 of the target volume recovered in the second progression, and $\hat{y}^{k}_{p_3,o_3}$ denotes the kth cross-section in o_3 of the target volume recovered in the third progression. Meanwhile, discriminator D_{o_3} learns to distinguish between the generated and real cross-sections. The final output volume $\hat{Y}_{p_3}$ of the proposed method is recovered by combining the generated cross-sections in o_3 via a concatenation block c_{o_3}:

$$\hat{Y}_{p_3} = c_{o_3}(\hat{y}^{1}_{p_3,o_3}, \ldots, \hat{y}^{K}_{p_3,o_3}) \tag{3.16}$$

where K denotes the total number of cross-sections in o_3. Note that, in MRI reconstruction, an additional constraint is introduced after each progression to enforce consistency of the acquired and recovered k-space data via the following procedure.

$$F_u(\hat{Y}_{p_n}) := F_u(X) \tag{3.17}$$

where F_u denotes the partial Fourier operator defined on the sampling mask utilized to acquire X, and n denotes the ongoing progression index. Meanwhile, an additional consistency between the progressions is enforced in the form of residual learning for MRI synthesis, where the generator models in the second and third progressions learn to predict the cross-sectional residuals between the target volume and the previously synthesized target volume.

$$\hat{y}_{p_n,o_n} = \hat{y}_{p_{n-1},o_n} + G_{o_n}(x_{o_n}, \hat{y}_{p_{n-1},o_n}) \tag{3.18}$$

3.5 Datasets

We demonstrated the proposed ProvoGAN approach on a public brain dataset, an in vivo knee dataset, and an in vivo brain dataset. The public dataset, IXI (`https://brain-development.org/ixi-dataset/`), consisted of coil-combined magnitude multi-contrast brain MR images of healthy subjects. The in vivo knee dataset [83] consisted of multi-coil complex knee MR images of healthy subjects. The in vivo brain dataset contained multi-contrast brain MR images of both healthy subjects and glioma patients. Further details about each dataset are provided below.

IXI Dataset: T_1-, T_2-, and proton-density (PD-)weighted brain MR images of 52 subjects were used, where 37 subjects were reserved for training, 5 for validation, and 10 for testing. T_1-weighted images were acquired sagittally with repetition time = 9.813 ms, echo time = 4.603 ms, flip angle = 8°, spatial resolution = $0.94 \times 0.94 \times 1.2$ mm^3, and matrix size = $256 \times 256 \times 150$. T_2-weighted images were acquired axially with repetition time = 8178 ms, echo time = 100

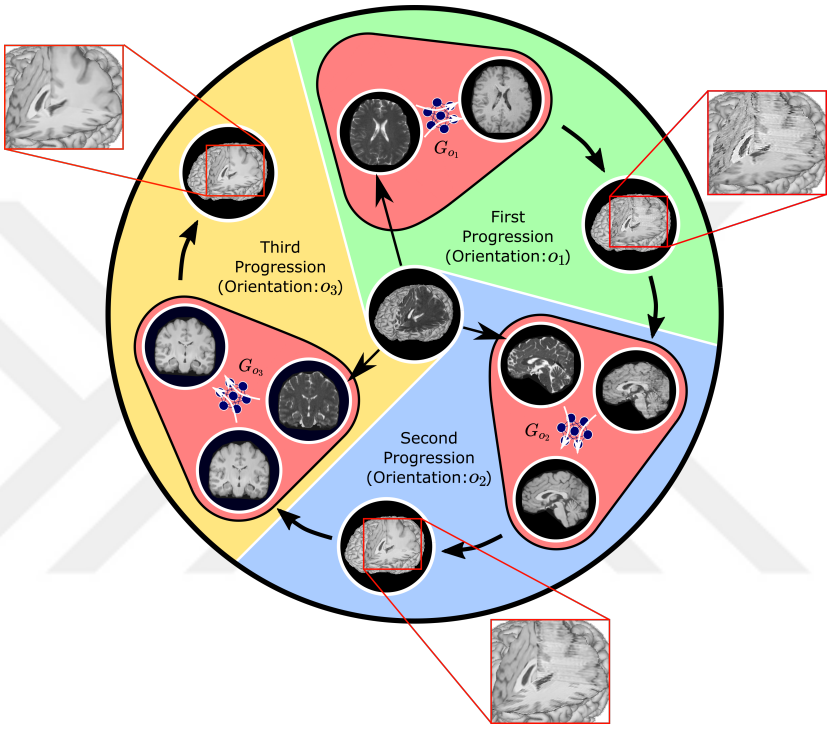


Figure 3.1: ProvoGAN decomposes complex volumetric image recovery tasks into a cascade of progressive cross-sectional subtasks defined across the rectilinear orientations (axial, coronal, and sagittal). Given a specific order of progression sequence (axial $\rightarrow$ sagittal $\rightarrow$ coronal is given here for demonstration), ProvoGAN first learns a cross-sectional mapping in the first orientation, and processes cross-sections within the entire source volume to estimate the target volume. This volumetric estimate is then divided into cross-sections in the second orientation, and a separate cross-sectional model is learned in the second orientation. The volumetric estimate from the second progression is then fed onto the final progression in which a third cross-sectional model is learned for final recovery. The sequential implementation of the progressive cross-sectional models enables ProvoGAN to gradually improve capture of fine-structural details in each orientation and to ensure global contextual consistency within the volume while at the same time manifesting reduced model complexity and improved learning behaviour of cross-sectional mapping.

ms, flip angle = 90°, spatial resolution = 0.94 × 0.94 × 1.20 mm^3, and matrix size = 256 × 256 × 150. PD-weighted images were acquired axially with repetition time = 8178.34 ms, echo time = 8 ms, flip angle = 90°, spatial resolution = 0.94 × 0.94 × 1.2 mm^3, and matrix size = 256 × 256 × 150. Since the images of separate contrasts were spatially unregistered in this dataset, T_2- and PD-weighted images were registered onto T_1-weighted images using FSL [84, 85] via an affine transformation. For synthesis images were further registered onto the MNI template of T_1-weighted images with an isotropic resolution of 1 mm^3. Registration was performed based on mutual information loss.

In vivo Knee Dataset: PD-weighted multi-coil knee MR images of 20 subjects were used, where 12 subjects were reserved for training, 3 for validation, and 5 for testing. Images were sagittally acquired with 8 receive coils, repetition time = 1550 ms, echo time = 25.661 ms, spatial resolution = $0.5 \times 0.5 \times 0.6$ mm^3, and matrix size = $320 \times 320 \times 256$. MRI scans were performed in the Richard M. Lucas Center at Stanford University, California, United States on 3T GE scanners.

In vivo Brain Dataset: T_1-weighted, contrast enhanced T_1-weighted (T_{1c}), T_2-weighted, and FLAIR coil-combined brain MR images of 11 healthy subjects, 12 glioma patients with homogenous tumor, and 62 glioma patients with heterogenous tumor were used. 55 subjects were reserved for training (healthy: 8, homogenous: 7, heterogenous: 40), 15 for validation (healthy: 2, homogenous: 2, heterogenous: 11), and 15 for testing (healthy: 2, homogenous: 2, heterogenous: 11). Data augmentation was performed to prevent class imbalance among the three subject groups. Augmentation was achieved by rotating the volumes around their longitudinal axis by a random angle in the range $[-10°, 10°]$, and repeated 10 times for healthy subjects, 9 times for glioma patients with homogenous tumor, and performed once for glioma patients with heterogenous tumor. MRI exams were performed in the Department of Radiology at Hacettepe University, Ankara, Turkey, on Siemens and Philips scanners under a diverse set of protocols with varying spatial resolution across both contrast sets and subjects. Specifically, the prescribed resolutions included $1 \times 1 \times 1$ mm^3, $0.9 \times 0.9 \times 1.5$

mm^3, $0.9 \times 0.9 \times 1.8$ mm^3, $0.9 \times 0.9 \times 2.2$ mm^3 for T_1- and T_{1c}-weighted images, and $0.3 \times 0.3 \times 5$ mm^3, $0.4 \times 0.4 \times 5$ mm^3, $0.5 \times 0.5 \times 5$ mm^3, $0.6 \times 0.6 \times 5$ mm^3, $0.7 \times 0.7 \times 5$ mm^3 for T_2-weighted and FLAIR images. For demonstrations, all images were registered onto the MNI template of T_1-weighted images with an isotropic resolution of $1\,mm^3$. Registration was performed via FSL [84, 85] using affine transformation based on mutual information loss. Imaging protocols were approved by the local ethics committee at Hacettepe University. All participants provided written informed consent.

For MRI reconstruction, volumes in the IXI and in vivo knee datasets were retrospectively undersampled with variable-density sampling patterns for acceleration factors ($R = 4, 8, 12, 16$). A sampling density function across k-space was taken a bi-variate normal distribution with mean at the center of k-space. The variance of the distribution was adjusted to achieve the expected sampling rate given R. The in-plane orientation was designated as axial. For MRI synthesis, all brain images were further skull stripped using FSL [84, 85] with functional intensity threshold of 0.5, and vertical gradient intensity threshold of 0.

3.6 Competing Methods

To demonstrate the performance of ProvoGAN in MR image recovery, we compared it against several state-of-the-art 3D models (vGAN, SC-GAN, REPLICA), 2D models (sGAN, RefineGAN, SPIRiT, SparseMRI), and hybrid models (M^3NET, TransferGAN). Baselines implemented for both reconstruction and synthesis included sGAN, vGAN, M^3NET, and TransferGAN. Meanwhile, task-specific baselines were RefineGAN, SPIRiT, and SparseMRI in MRI reconstruction, and SC-GAN and REPLICA in MRI synthesis.

vGAN: A learning-based volumetric GAN model that performs a global one-shot mapping between source and target volumes (see Section 3.2). vGAN was implemented with a ResNet-based generator and a PatchGAN discriminator.

sGAN: A learning-based cross-sectional GAN model that performs a localized mapping between cross-sections of the source and target volumes (see Section 3.3). sGAN contained a ResNet-based generator and a PatchGAN discriminator.

RefineGAN: A learning-based cross-sectional GAN model proposed for MRI reconstruction [57]. RefineGAN uses a cycle-consistency loss for acquired k-space samples in addition to adversarial and pixel-wise image loss to improve reconstruction quality. The overall architecture and loss terms were taken from [57], but a ResNet-based generator was implemented to enable fair comparisons against ProvoGAN as it was observed here to yield higher reconstruction quality.

SC-GAN: A learning-based volumetric GAN model proposed for MRI synthesis [34]. SC-GAN leverages self-attention modules to improve capture of long-range spatial dependencies. SC-GAN was implemented with a U-Net based generator and a PatchGAN discriminator as described in [34], where the encoder and decoder components in the generator and the intermediate layer in the discriminator contained a self-attention module.

M^{3}NET: A learning-based hybrid model proposed for MRI segmentation [65]. First, M^{3}NET separately learns orthogonal cross-sectional mappings in three rectilinear orientations (i.e., axial, coronal, sagittal). Using these 2D mappings as parallel streams, it fuses their outputs with a 3D fusion module to recover the target volume. The overall architecture, 3D fusion module, and loss functions were adopted from [65], where 2D models were implemented with ResNet-based generators and PatchGAN discriminators as they were observed to yield enhanced performance in this study.

TransferGAN: A learning-based hybrid GAN model proposed for low-dose CT denoising [67]. TransferGAN pretrains a 2D model for image recovery in a specific orientation, and then performs domain transfer from 2D onto 3D by transferring model weights. The transfer learning procedure was implemented as described in [67], with 2D-3D models implemented as conditional GANs using ResNet-based generators and PatchGAN discriminators for fair comparison against ProvoGAN.

SparseMRI: A compressed sensing-based cross-sectional method for single-coil MRI reconstruction [28]. SparseMRI enforces transform domain sparsity as prior information during reconstruction from undersampled acquisitions. Here, SparseMRI was implemented as described in [28].

SPIRiT: A compressed sensing-based cross-sectional method for multi-coil MRI reconstruction [86]. SPIRiT employs k-space interpolation kernels to estimate missing k-space samples. Here, SPIRiT was implemented as described in [86].

REPLICA: A compressed sensing-based volumetric method for multi-contrast MRI synthesis [87]. REPLICA performs a nonlinear intensity transformation in multi-resolution feature space via a regression ensemble based on random forests. Here, REPLICA was implemented as described in [87].

The main effect that we seeked in comparing ProvoGAN against sGAN and vGAN was the benefit of progressive volumetrization over purely 2D or 3D processing. To improve reliability of these comparisons, we wanted to control for potential confounds from secondary factors such as network architecture or loss function. Therefore, the sGAN and vGAN models embodied consistent generator-discriminator architectures and loss functions with ProvoGAN. Detailed information of network architecture and implementation details of ProvoGAN and competing methodas are as follows:

ProvoGAN: Network Architecture

The proposed ProvoGAN model is based on conditional generative adversarial networks (GANs) [81]. Each GAN model within the progressions consisted of a generator that contains an encoder of 3 convolutional layers, a residual network of 9 ResNet blocks, and a decoder of 3 convolutional layers, and a discriminator that contains a convolutional network of 5 convolutional layers, all with 2D kernels. The kernel size, number of filters, stride, activation function, and connections of the network layers in the generators and discriminators are provided

in Figure 3.2. The generator in the first progression received as input the cross-sections of the source volume in the first orientation, whereas the generators in the second and third progressions received as input the cross-sections of both source and previously recovered target volumes in the second and third orientations, respectively. Conditional PatchGAN discriminator architectures were used to effectively incorporate priors. The discriminator received as input the concatenation of source and target cross-sectional images. The source cross-sections were identical to the inputs of the corresponding generator. Ground truth target images were designated as real target images, whereas generator-recovered target images were designated as the fake target images.

ProvoGAN: Implementation Details

The training procedure of ProvoGAN comprised three progressive phases, and within each phase the respective pair of generator and discriminator architectures were trained to learn cross-sectional recovery in the given orientation. Hyperparameter-selection procedures for all phases were adopted from [59]. The number of epochs and relative weighting of the loss terms were optimized via PSNR measurements in the validation set. The generator and discriminator were trained with the ADAM optimizer [88] ($\beta_1 = 0.5$, $\beta_2 = 0.999$) for 100 epochs. The learning rate was set to 2×10^{-4} in the first 50 epochs and was linearly decayed to 0 in the last 50 epochs. All cross-sectional training samples were processed by the networks in every training epoch with a batch size of 1, and instance normalization was performed. The optimal relative weighing of the pixel-wise loss to the adversarial loss was selected as 100.

Competing Methods: Network Architectures

The cross-sectional sGAN and RefineGAN models as well as cross-sectional components of the hybrid M^{3}NET and TransferGAN models were based on 2D conditional GANs [81] with a ResNet backbone [59]. The resultant GAN model consisted of a generator that contains an encoder of 3 convolutional layers, a residual network of 9 ResNet blocks, and a decoder of 3 convolutional layers, and

a discriminator that contains a convolutional network of 5 convolutional layers, all formed with 2D kernels. The kernel size, number of filters, stride, activation function, and connections of the network layers in the generators and discriminators are provided in Figure 3.3a. The generators in these models received as input cross-sections of the source contrast in the selected orientation. The discriminators received as input either recovered or ground truth cross-sections of the target contrast concatenated with cross-sections of the source contrast in the same orientation.

On the other hand, the volumetric vGAN model and the volumetric component of the hybrid TransferGAN model were based on 3D conditional GANs [81] again with a ResNet backbone. GAN models consisted of a generator that contains an encoder of 3 convolutional layers, a residual network of 9 ResNet blocks, and a decoder of 3 convolutional layers, and a discriminator that contains a convolutional network of 5 convolutional layers, all with 3D kernels. The kernel size, number of filters, stride, activation function, and connections of the network layers in the generators and discriminators are provided in Figure 3.3b. The generators in these volumetric models received as input the entire volume of the source contrast. The discriminators received as input the entire volume of either recovered or ground truth target contrast concatenated with the entire volume of the source contrast.

The architecture of the hybrid M^3NET method was adopted from [65]. Meanwhile, SC-GAN embodied a generator based on a U-Net architecture and a discriminator based on a PatchGAN architecture. Both subnetworks contained 3D self-attention modules as described in [34]. The generators in these volumetric models again received as input the entire volume of the source contrast. The discriminators received as input the entire volume of either recovered or ground truth target contrast concatenated with the entire volume of the source contrast. The compressed-sensing-based REPLICA, SparseMRI, and SPIRiT methods were implemented as described in [87], [28], and [86], respectively. REPLICA received as input the entire volume of the source contrast to recover the target volume. Meanwhile, SparseMRI and SPIRiT received as input cross-sections of the undersampled source volume to recover cross-sections of the target volume.

Competing Methods: Implementation Details

For the cross-sectional sGAN and RefineGAN models and cross-sectional components of hybrid M^{3}NET and TransferGAN models, generator-discriminator pairs were trained to learn a recovery task in the selected orientation. Hyperparameter selection procedures were again adopted from [59]. The number of epochs and relative weighting of the loss terms were optimized via PSNR measurements in the validation set. The generator and discriminator were trained with the ADAM optimizer [88] ($\beta_1 = 0.5$, $\beta_2 = 0.999$) for 100 epochs. The learning rate was set to 2×10^{-4} in the first 50 epochs and was linearly decayed to 0 in the last 50 epochs. All cross-sectional training samples were processed by the models in every training epoch with a batch size of 1, and instance normalization was performed. The relative weighing of the pixel-wise loss to the adversarial loss was selected as 100 for sGAN and cross-sectional components of M^{3}NET and TransferGAN. For RefineGAN, the relative weighing of the pixel-wise loss and cycle consistency loss of acquired k-space samples to the adversarial loss [57] were taken as 10 and 1. For the volumetric vGAN and volumetric component of TransferGAN, generator-discriminator pairs were trained to learn a volumetric recovery task. The number of epochs and relative weighting of the loss terms were optimized via PSNR measurements in the validation set. The generator and discriminator were trained with the ADAM optimizer [88] ($\beta_1 = 0.5$, $\beta_2 = 0.999$) for 200 epochs. The learning rate was set to 2×10^{-4} for synthesis tasks and 5×10^{-4} for reconstruction tasks in the first 100 epochs, and was linearly decayed to 0 in the last 100 epochs. All training samples were processed by the models in every training epoch with a batch size of 1, and instance normalization was performed. The relative weighing of the pixel-wise loss to the adversarial loss was selected as 50 for synthesis tasks and 200 for reconstruction tasks.

Training procedures of compressed-sensing-based REPLICA, SparseMRI and SPIRIT methods were implemented following procedures outlined in [87], [28], and [86], respectively. For REPLICA, hyperparameter selections reported in [87] were adopted as they were observed to yield high performance. For SparseMRI

and SPIRIT, number of iterations and relative weighting of the loss terms were optimized via PSNR measurements in the validation set. For SparseMRI, number of iterations was selected as 40. L1 regularization weight of wavelet coefficients was set to 0.002 for T_1reconstruction with R=4x, 0.004 for T_2reconstruction with R=16x, and 0.01 for remaining reconstruction tasks. For SPIRIT, number of iterations was selected as 30 and interpolation kernel size was set to 5. L1 and L2 regularization weights were respectively selected as $0.01 - 0.1$ for R=4,8, $0.01 - 0.001$ for R=12, and $0.01 - 0.0001$ for R=16.

In single-coil reconstruction, learning-based models were trained to recover a magnitude image given real and imaginary parts of the undersampled image. In multi-coil reconstruction, learning-based models were first trained to recover a coil-combined magnitude image given real and imaginary parts of coil-combined Fourier reconstructions of undersampled acquisitions. A complex image was then formed by mapping the phase of the coil-combined undersampled image onto the predicted magnitude image. Coil combination was performed using sensitivity maps estimated via ESPIRiT [89]. A multi-coil complex image was obtained by projecting the coil-combined network prediction onto individual coils with the estimated sensitivity maps. Data-consistency was enforced in Fourier domain using the multi-coil complex images. In synthesis, learning-based models were trained to recover the magnitude image of the target contrast given magnitude images of the source contrasts.

The volumetric vGAN, SC-GAN, and REPLICA methods received as input volumetric source images. The cross-sectional sGAN-A, sGAN-C, sGAN-S, RefineGAN, SPIRiT, and SparseMRI methods received as input individual cross-sections of source volumes. M^3NET received cross-sectional inputs, aggregated them across the volume and finally processed the entire volume. TransferGAN received cross-sectional inputs during pretraining of the 2D model, and instead received volumetric inputs during training of the 3D model.

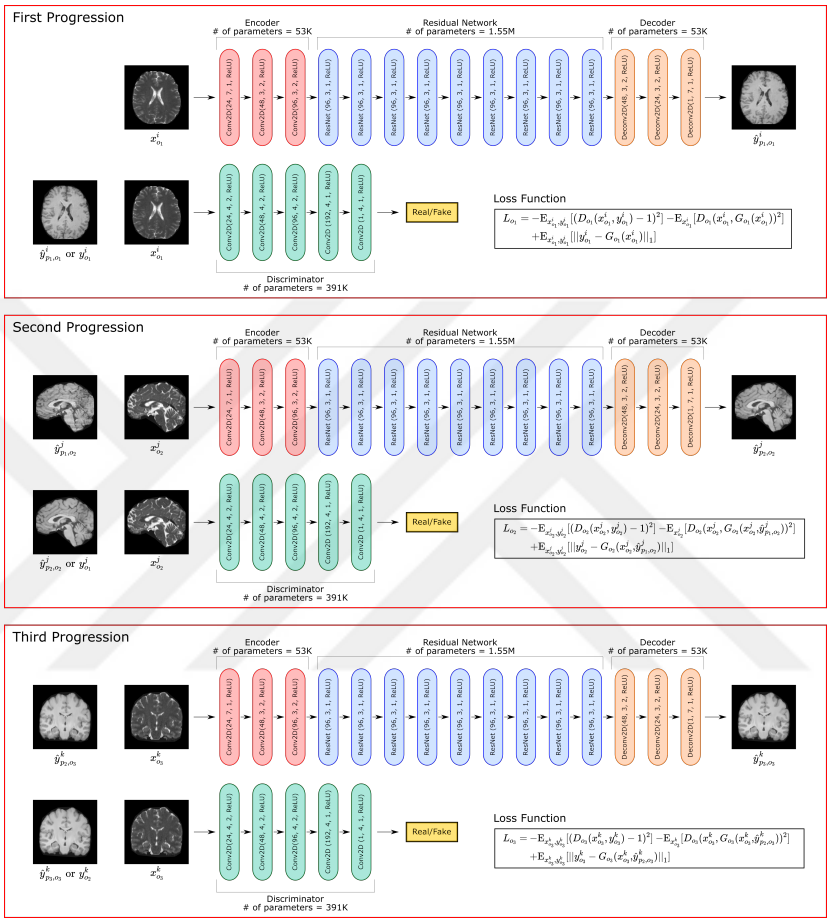


Figure 3.2: The architectural details of the generator and discriminator submodules in ProvoGAN are displayed (here a progression order of A $\rightarrow$ S $\rightarrow$ C is used for illustration). The generator architectures are the same for the progressions and consist of an encoder with 3 convolutional layers: conv2D($k = 7, f = 24, s = 1, a =$ ReLU), conv2D($k = 3, f = 48, s = 2, a =$ ReLU), conv2D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet2D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv2D($k = 3, f = 48, s = 2, a =$ ReLU), deconv2D($k = 3, f = 24, s = 2, a =$ ReLU), conv2D($k = 7, f = 1, s = 1, a =$ Tanh), where k denotes kernel size, f denotes number of filters, s denotes stride, and a denotes activation function. Similarly, the discriminator architectures are identical for the progresions and consist of a convolutional network of 5 convolutional layers in series: conv2D($k = 4, f = 24, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 48, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv2D($k = 4, f = 1, s = 1, a =$ none).

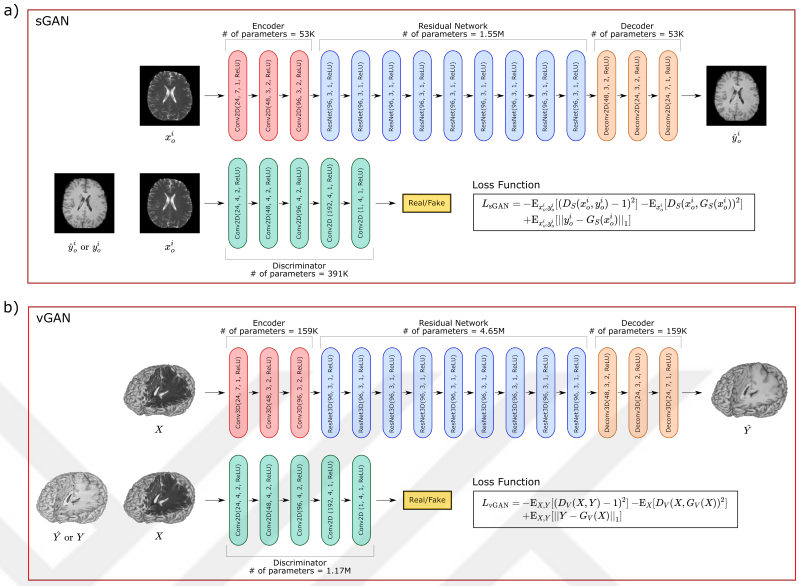


Figure 3.3: a) The architectural details of the generator and discriminator in sGAN are displayed (here an sGAN model trained in the axial orientation referred to as sGAN-A is used for illustration). The generator consists of an encoder with 3 convolutional layers: conv2D($k = 7, f = 24, s = 1, a =$ ReLU), conv2D($k = 3, f = 48, s = 2, a =$ ReLU), conv2D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet2D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv2D($k = 3, f = 48, s = 2, a =$ ReLU), deconv2D($k = 3, f = 24, s = 2, a =$ ReLU), conv2D($k = 7, f = 1, s = 1, a =$ Tanh), where k denotes kernel size, f denotes number of filters, s denotes stride, and a denotes activation function. The discriminator consists of a convolutional network of 5 convolutional layers in series: conv2D($k = 4, f = 24, s = 2, a =$ leakyReLU), $conv2D(k = 4, f = 48, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv2D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv2D($k = 4, f = 1, s = 1, a =$ none).

b) The architectural details of the generator and discriminator in vGAN are displayed. The generator consists of an encoder with 3 convolutional layers: conv3D($k = 7, f = 24, s = 1, a =$ ReLU), conv3D($k = 3, f = 48, s = 2, a =$ ReLU), conv3D($k = 3, f = 96, s = 2, a =$ ReLU), a residual network of 9 ResNet blocks: $9 \times$ ResNet3D($k = 3, f = 96, s = 1, a =$ ReLU), and a decoder of 3 convolutional layers: deconv3D($k = 3, f = 48, s = 2, a =$ ReLU), deconv3D($k = 3, f = 24, s = 2, a =$ ReLU), conv3D($k = 7, f = 1, s = 1, a =$ Tanh). The discriminator consists of a convolutional network of 5 convolutional layers in series: conv3D($k = 4, f = 24, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 48, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 96, s = 2, a =$ leakyReLU), conv3D($k = 4, f = 192, s = 1, a =$ leakyReLU), conv3D($k = 4, f = 1, s = 1, a =$ none).

3.7 Experiments

Task-Specific Progression Order in ProvoGAN: Experiments were performed on ProvoGAN to optimize its progression order across the rectilinear orientations for specific tasks. To do this, multiple independent ProvoGAN models were trained while varying the progression order: 1) A $\rightarrow$ C $\rightarrow$ S, 2) A $\rightarrow$ S $\rightarrow$ C, 3) C $\rightarrow$ A $\rightarrow$ S, 4) C $\rightarrow$ S $\rightarrow$ A, 5) S $\rightarrow$ A $\rightarrow$ C, 6) S $\rightarrow$ C $\rightarrow$ A, where A denotes the axial, C denotes the coronal, and S denotes the sagittal orientation. Performance of these models were evaluated on the validation set via PSNR measurements. The experiments were performed separately for all synthesis and reconstruction tasks, and the progression orders optimized for specific tasks were used in all evaluations thereafter.

MRI Reconstruction: Reconstruction experiments were performed on the IXI and in vivo knee datasets to compare ProvoGAN against sGAN, vGAN, RefineGAN, SparseMRI, and SPIRiT. In the IXI dataset, the proposed and competing methods were demonstrated separately for single-coil reconstruction of T_1- and T_2-weighted images with four distinct acceleration factors ($R = 4, 8, 12, 16$). Meanwhile, in the in vivo knee dataset, the proposed and competing methods were demonstrated for multi-coil reconstruction of PD-weighted images again with ($R = 4, 8, 12, 16$). Note that a single sGAN model was trained in the axial orientation (sGAN-A) given the axial readout direction.

MRI Synthesis: Synthesis experiments were performed on the IXI and in vivo brain datasets to demonstrate ProvoGAN against sGAN, vGAN, SC-GAN, and REPLICA. All synthesis experiments were conducted on coil-combined magnitude images. In the IXI dataset, three synthesis tasks were considered: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. In the in vivo brain dataset, four synthesis tasks were considered: 1) T_2, FLAIR, T_{1c} $\rightarrow$ T_1, 2) T_1, FLAIR, T_{1c} $\rightarrow$ T_2, 3) T_1, T_2, T_{1c} $\rightarrow$ FLAIR, 4) T_1, T_2, FLAIR $\rightarrow$ T_{1c}. For each task, three independent sGAN models were implemented to recover target cross-sections in separate orientations: sGAN-A for the axial, sGAN-C for the

coronal, sGAN-S for the sagittal orientation.

Progressive Volumetrization versus Hybrid Models: Experiments were conducted on the IXI dataset to demonstrate ProvoGAN against M^{3}NET and TransferGAN. Reconstruction experiments were conducted for T_1- and T_2-weighted image recovery tasks at four distinct acceleration factors ($R = 4, 8, 12, 16$). Meanwhile, synthesis experiments were conducted for the many-to-one recovery tasks of T_2, PD $\rightarrow$ T_1, T_1, PD $\rightarrow$ T_2, and T_1, T_2 $\rightarrow$ PD.

Radiological Evaluation: To assess the clinical value of the recovered images, an expert radiologist (25+ years of experience) gave opinion scores to the images while blinded to the method name and order of presentation. Reconstructed images were evaluated for single-coil reconstructions of T_1- and T_2-weighted acquisitions at $R = 8$ in the IXI dataset, and multi-coil reconstructions of PD-weighted acquisitions at $R = 8$ in the in vivo knee dataset. Synthesized images were evaluated for T_2, PD $\rightarrow$ T_1 in IXI and T_1, T_2, T_{1c} $\rightarrow$ FLAIR in the in vivo brain datasets. From each recovered volume, intermediate axial, coronal, and sagittal cross-sections were randomly selected, and the image quality was rated as the similarity to ground truth images on a five-point scale (5: perfect match, 4: good, 3: moderate, 2: limited, 1: very poor, 0: unacceptable).

Multi-Cross-Section Models: To demonstrate the benefit of leveraging contextual priors by incorporating multiple neighboring cross-sections at the input level, variants of ProvoGAN and sGAN, referred to as ProvoGAN(multi) and sGAN(multi), were implemented, which receive as input n_c consecutive cross-sections to recover the corresponding central cross-section in the target volume. Here $n_c = 3$ was selected as higher number of cross-sections did not yield a notable benefit in recovery performance [59]. Experiments were performed on the IXI dataset for reconstruction of T_1- and T_2-weighted images with distinct acceleration factors ($R = 4, 8, 12, 16$), and for many-to-one synthesis tasks (T_2, PD $\rightarrow$ T_1, T_1, PD $\rightarrow$ T_2, T_1, T_2 $\rightarrow$ PD). The ordering of the progressions across the orientations in ProvoGAN was optimized via PSNR measurements in the validation set.

Three separate sGAN(multi) models were implemented in each individual rectilinear orientation: sGAN(multi)-A for the axial, sGAN(multi)-C for the coronal, and sGAN(multi)-S for the sagittal orientation.

Cross-Sectional Models of Varying Complexity: An additional analysis was performed on ProvoGAN and sGAN to examine recovery performance as a function of the complexity of convolutional layers. Several variants of ProvoGAN and sGAN were implemented while the number of network weights in individual convolutional layers were scaled by $n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$, where the kernel size, number of layers, number of hidden units were kept fixed but the number of filters were modified. This resulted in seven distinct ProvoGAN and sGAN pairs: ProvoGAN(n_f)-sGAN(n_f). Experiments were performed on the IXI dataset for single-coil reconstruction of T_1-weighted acquisitions undersampled at $R = 8$ and a many-to-one synthesis task of T_2, PD $\rightarrow$ T_1. The ordering of the progressions across the orientations in ProvoGAN(n_f) was optimized via PSNR measurements in the validation set. Separate sGAN models were trained in the individual rectilinear orientations (axial, coronal, and sagittal) for each model complexity level: sGAN(n_f)-A, sGAN(n_f)-C, sGAN(n_f)-S.

Generalizability of Progressive Volumetrization: Experiments were conducted to demonstrate the generalizability of the proposed progressive volumetrization to another network architecture. Demonstrations were performed on the IXI dataset for T_1, T_2 $\rightarrow$ PD, T_1, PD $\rightarrow$ T_2, and T_2, PD $\rightarrow$ T_1 synthesis tasks. A recent state-of-the-architecture, SC-GAN, with a U-Net backbone using intermittent self-attention layers [34] was considered. Variants of sGAN, vGAN, and ProvoGAN were implemented based on this architecture: sSC-GAN, vSC-GAN, ProvoSC-GAN. Again, three separate sSC-GAN models were trained in each orientation: sSC-GAN-A in the axial, sSC-GAN-C in the coronal, and sSC-GAN-S in the sagittal orientation.

Statistical Assessments: PSNR, SSIM, and opinion scores were utilized to quantitatively evaluate the recovery quality of the methods under comparison. Since the performance measurements from these metrics followed a non-normal

distribution ($p < 0.05$ with Shapiro-Willks test), significance of differences in quantitative metrics were evaluated using non-parametric statistical tests. Assessments of progression order in ProvoGAN were performed via Kruskal-Wallis tests, whereas performance comparisons among competing methods were performed via Wilcoxon signed-rank tests.

Chapter 4

Results

4.1 Task-Specific Progression Order

ProvoGAN serially decomposes a given volumetric recovery task into cross-sectional mappings in three rectilinear orientations. Subsequent 2D mappings are residually learned based on outputs from earlier progressions. Please note that spatial distribution of the tissues and the correlations between the source-target images may vary uniquely across orientations for each recovery task. In this setup, if an earlier 2D model yields relatively higher artifacts, the task difficulty for the remaining progressions would be elevated. Contrarily, initiating the progression at a different orientation with lower artifacts can reduce task difficulty for the remaining stages. Therefore, we predicted that the progression sequence in ProvoGAN can significantly affect task-specific recovery performance.

To test this prediction, we performed reconstruction and synthesis experiments separately on the IXI, in vivo brain, and in vivo knee datasets (see Section 3.5 for details). We comparatively evaluated performance of multiple independent ProvoGAN models for the six possible permutations of the progression sequence: 1) A $\rightarrow$ C $\rightarrow$ S, 2) A $\rightarrow$ S $\rightarrow$ C, 3) C $\rightarrow$ A $\rightarrow$ S, 4) C $\rightarrow$ S $\rightarrow$ A, 5) S $\rightarrow$ A $\rightarrow$ C, 6) S $\rightarrow$ C $\rightarrow$ A, where A denotes the axial, C denotes the coronal, and S denotes

the sagittal orientation. Here, we considered volumetric PSNR measurements between the recovered and reference target volumes within the validation set. The highest and lowest performing ProvoGAN models yield an average PSNR difference of 3.44 dB for single-coil reconstruction tasks in IXI and 3.42 dB for multi-coil reconstruction tasks in the in vivo knee dataset (see Tables 4.1, 4.2 for details). Meanwhile, the average PSNR difference between the highest and lowest performing ProvoGAN models is 1.46 dB for synthesis in IXI, and 1.01 dB for synthesis in the in vivo brain dataset (see Tables 4.3, 4.4). Optimization of the progression order enables a significant performance increase in both reconstruction ($p < 0.05$, Kruskal-Wallis test) and synthesis ($p < 0.05$). Therefore, the optimal orders were utilized for each recovery task in all evaluations thereafter unless otherwise stated.

Table 4.1: Task-Optimal Progression Order for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) measurements between the reconstructed and ground truth images in the validation set in the IXI dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and acceleration factors: $R = 4, 8, 12, 16$. Boldface indicates the highest performing progression sequence.

		A → C → S	**A → S → C**	**S → A → C**	**S → C → A**	**C → S → A**	**C → A → S**
R=4	T_1	35.12 ± 1.73	**36.12 ± 1.22**	35.80 ± 1.81	35.50 ± 1.63	35.84 ± 1.05	35.21 ± 1.87
	T_2	36.42 ± 1.84	36.62 ± 1.24	**36.69** ± 1.72	26.92 ± 1.02	29.50 ± 0.64	36.51 ± 2.68
R=8	T_1	32.48 ± 0.42	33.51 ± 0.75	**33.52 ± 0.70**	33.41 ± 0.49	32.85 ± 0.83	32.66 ± 0.33
	T_2	32.66 ± 1.66	**33.73 ± 2.67**	32.71 ± 2.13	28.52 ± 1.14	30.77 ± 0.92	33.10 ± 1.78
R=12	T_1	30.43 ± 0.44	30.20 ± 0.72	30.94 ± 0.94	30.18 ± 0.73	30.04 ± 0.70	**31.26 ± 1.08**
	T_2	30.76 ± 0.76	30.57 ± 1.09	30.86 ± 1.13	30.75 ± 2.27	27.79 ± 0.59	**31.09 ± 0.78**
R=16	T_1	29.44 ± 0.69	29.11 ± 0.97	29.35 ± 0.73	29.82 ± 0.61	29.39 ± 0.90	**30.38 ± 1.37**
	T_2	30.74 ± 0.77	29.73 ± 0.38	**31.84 ± 0.62**	27.15 ± 2.19	29.26 ± 1.37	30.27 ± 0.61

Table 4.2: Task-Optimal Progression Order for Reconstruction in the In vivo Knee Dataset: Volumetric PSNR (dB) measurements between the reconstructed and ground truth images in the validation set in the in vivo knee dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and acceleration factors: $R = 4, 8, 12, 16$. Boldface indicates the highest performing progression sequence.

	A → C → S	**A → S → C**	**S → A → C**	**S → C → A**	**C → S → A**	**C → A → S**
R=4	37.56 ± 1.09	37.91 ± 0.49	**39.72 ± 1.63**	38.73 ± 1.60	38.18 ± 1.99	38.27 ± 1.52
R=8	38.80 ± 1.39	38.25 ± 1.13	36.38 ± 1.09	**39.95 ± 0.64**	37.44 ± 1.18	35.59 ± 1.19
R=12	35.20 ± 0.04	38.36 ± 0.42	37.22 ± 1.29	**39.14 ± 0.44**	37.49 ± 0.81	37.61 ± 1.07
R=16	37.26 ± 1.19	35.05 ± 3.30	37.38 ± 1.10	38.01 ± 0.27	**38.28 ± 0.25**	37.49 ± 0.94

Table 4.3: Task-Optimal Progression Order for Synthesis in the IXI Dataset: Volumetric PSNR (dB) measurements between the synthesized and ground truth images in the validation set in the IXI dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and all many-to-one synthesis tasks: 1) T_2, PD → T_1, 2) T_1, PD → T_2, 3) T_1, T_2 → PD. Boldface indicates the highest performing progression sequence.

	A → C → S	**A → S → C**	**S → A → C**	**S → C → A**	**C → S → A**	**C → A → S**
T_2, PD → T_1	23.38 ± 2.38	23.75 ± 2.59	24.24 ± 3.81	24.01 ± 3.65	**25.35 ± 2.23**	24.90 ± 2.06
T_1, PD → T_2	28.35 ± 1.65	28.34 ± 1.67	28.50 ± 1.16	28.37 ± 1.04	**29.51 ± 1.41**	28.62 ± 1.08
T_1, T_2 → PD	30.42 ± 1.29	**31.50 ± 1.61**	30.46 ± 1.08	30.50 ± 1.14	31.44 ± 1.87	30.25 ± 1.51

Table 4.4: Task-Optimal Progression Order for Synthesis in the In vivo Brain Dataset: Volumetric PSNR (dB) measurements between the synthesized and ground truth images in the validation set in the in vivo brain dataset are given as mean ± std. The measurements are provided for all possible progression orders: 1) A → C → S, 2) A → S → C, 3) S → A → C, 4) S → C → A, 5) C → S → A, 6) C → A → S and all many-to-one synthesis tasks: 1) T_2, FLAIR, T_{1c} → T_1, 2) T_1, FLAIR, T_{1c} → T_2, 3) T_1, T_2, T_{1c} → FLAIR, 4) T_1, T_2, FLAIR → T_{1c}. Boldface indicates the highest performing progression sequence.

	A → C → S	**A → S → C**	**S → A → C**	**S → C → A**	**C → S → A**	**C → A → S**
T_2, FLAIR, T_{1c} → T_1	24.48 ± 2.91	24.49 ± 2.85	24.48 ± 3.32	24.45 ± 3.27	**25.63 ± 3.53**	24.81 ± 3.14
T_1, FLAIR, T_{1c} → T_2	27.37 ± 2.96	27.14 ± 2.86	27.66 ± 2.91	27.19 ± 2.76	26.96 ± 2.81	**27.68 ± 2.99**
T_1, T_2, T_{1c} → FLAIR	24.93 ± 2.40	**25.58 ± 2.47**	24.90 ± 3.36	25.35 ± 3.45	25.51 ± 2.93	**25.58 ± 3.09**
T_1, T_2, FLAIR → T_{1c}	28.83 ± 2.08	**29.88 ± 2.44**	28.75 ± 2.54	28.83 ± 2.34	28.94 ± 2.26	28.43 ± 2.14

4.2 Accelerated MRI Reconstruction

Next, we performed comprehensive experiments on the IXI and in vivo knee datasets for accelerated MRI reconstruction. We comparatively demonstrated the recovery quality of ProvoGAN against state-of-the-art cross-sectional (sGAN, RefineGAN, SparseMRI, and SPIRiT), and volumetric (vGAN) models (see Section 3.6 for details). We first assessed the performance of the competing methods quantitatively based on volumetric PSNR and SSIM measurements between the reconstructed and high-quality reference images in the test set. We considered single-coil reconstruction tasks in the IXI dataset for T_1- and T_2-weighted images with distinct acceleration factors ($R = 4, 8, 12, 16$). The proposed ProvoGAN model offers enhanced recovery performance compared to competing methods ($p < 0.05$), where it achieves in the range of $[1.85, 6.55]$ dB higher PSNR and $[3.26, 23.17]$ % higher SSIM (see Table 4.5). We then considered multi-coil reconstruction of PD-weighted images in the in vivo brain dataset with $R = 4, 8, 12, 16$. ProvoGAN again maintains superior performance to the competing methods ($p < 0.05$), where it achieves in the range of $[1.03, 7.35]$ dB higher PSNR and $[1.48, 5.59]$ % higher SSIM (see Table 4.6).

To corroborate quantitative assessments, we visually examined the reconstructed volumes from individual methods to identify the nature of reconstruction errors ProvoGAN alleviates. Representative results from the competing methods are shown in Fig. 4.1 for IXI and in Fig. 4.2 for the in vivo knee dataset. Overall, cross-sectional models (sGAN, RefineGAN, SparseMRI, SPIRiT) that perform 2D mapping via compressed sensing or deep learning suffer from discontinuity artifacts across individually recovered cross-sections and retrograded capture of fine-structural details. Meanwhile, the volumetric vGAN model performing 3D mapping suffers from loss of spatial resolution within the reconstructed volumes due to noticeable over-smoothing. In contrast, ProvoGAN reconstructs the target volumes with higher consistency across the cross-sections in all orientations and offers sharper delineation of brain and knee tissues. Taken together, these findings clearly outline ProvoGAN's potential to mitigate the limitations of volumetric and cross-sectional models for accelerated MRI reconstruction.

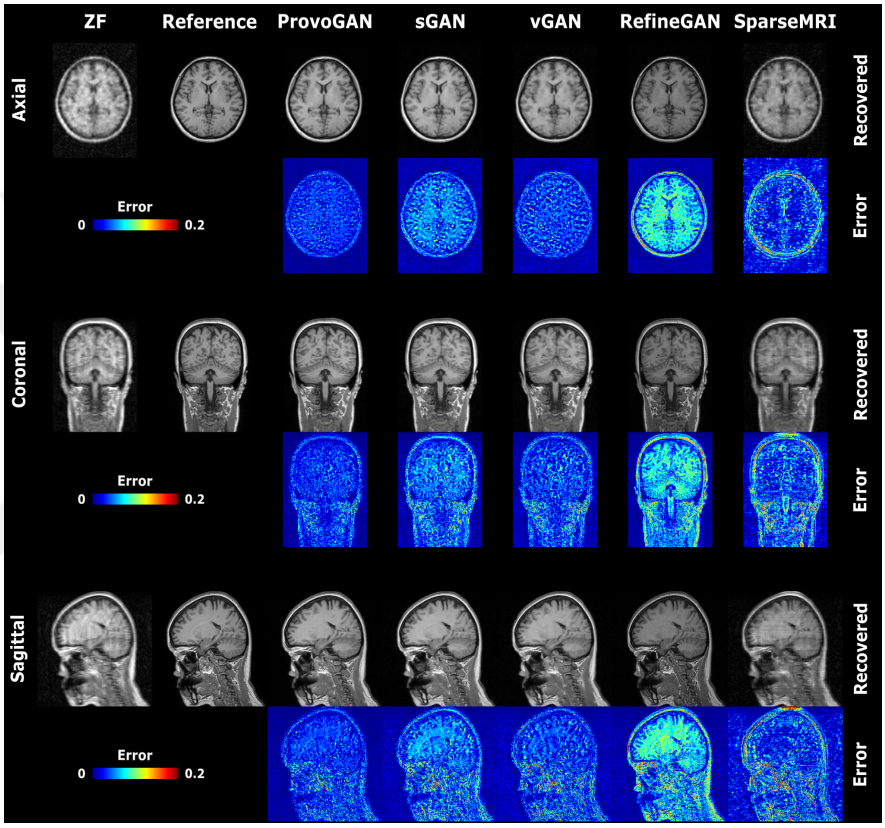


Figure 4.1: The proposed ProvoGAN method is demonstrated on the IXI dataset for single-coil reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. Representative results are displayed for all competing methods together with the zero-filled (ZF) undersampled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed ProvoGAN method offers delineation of tissues with higher acuity compared to the volumetric (vGAN) model, and alleviates undesirable discontinuities compared to cross-sectional models (sGAN, RefineGAN, SparseMRI) by improving reconstruction performance in all orientations.

Table 4.5: Quality of Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the IXI dataset are given as mean ± std for the test set. The measurements are reported for zero-filled images (ZF), the proposed ProvoGAN and competing sGAN, vGAN, RefineGAN, and SparseMRI reconstruction methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method.

		ProvoGAN		sGAN		vGAN		RefineGAN		SparseMRI		ZF	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
R=4	T_1	**35.25 ± 1.78**	**96.73 ± 0.57**	33.85 ± 1.29	93.21 ± 0.78	30.82 ± 1.24	88.61 ± 0.93	31.04 ± 1.74	92.13 ± 0.94	26.25 ± 0.80	72.16 ± 1.92	24.59 ± 1.17	64.96 ± 3.07
	T_2	**35.50 ± 2.62**	**96.08 ± 1.07**	32.95 ± 1.50	86.44 ± 1.26	33.09 ± 1.73	93.13 ± 1.20	33.96 ± 0.91	94.00 ± 0.57	28.08 ± 0.73	82.44 ± 1.42	27.54 ± 1.04	75.14 ± 2.09
R=8	T_1	**31.38 ± 1.26**	**94.93 ± 0.86**	30.08 ± 1.32	91.18 ± 1.01	29.71 ± 0.88	88.73 ± 1.23	27.63 ± 1.56	91.81 ± 0.68	25.92 ± 0.36	72.26 ± 1.52	23.04 ± 1.05	62.53 ± 3.32
	T_2	**33.49 ± 2.21**	**95.92 ± 1.01**	32.24 ± 2.14	90.47 ± 0.95	31.35 ± 0.66	92.57 ± 0.86	30.70 ± 1.51	93.82 ± 0.40	26.55 ± 0.51	79.35 ± 1.39	26.67 ± 0.87	74.16 ± 2.10
R=12	T_1	**29.67 ± 0.91**	**92.48 ± 0.90**	27.34 ± 1.06	86.23 ± 1.36	27.56 ± 1.20	82.70 ± 1.70	27.48 ± 1.08	88.56 ± 1.07	23.84 ± 0.28	60.77 ± 1.74	20.76 ± 1.04	52.72 ± 3.62
	T_2	**30.41 ± 1.03**	**91.98 ± 1.40**	28.48 ± 1.06	79.50 ± 2.15	27.97 ± 0.79	85.76 ± 1.74	29.16 ± 1.54	91.49 ± 0.87	24.67 ± 0.59	70.16 ± 1.67	24.60 ± 0.70	66.90 ± 2.12
R=16	T_1	**29.15 ± 1.09**	**91.40 ± 1.09**	26.73 ± 1.53	85.23 ± 1.74	25.47 ± 1.07	79.56 ± 2.16	22.37 ± 1.07	84.31 ± 1.36	23.42 ± 0.47	61.64 ± 1.86	20.95 ± 1.07	52.85 ± 3.46
	T_2	**30.66 ± 1.60**	**93.74 ± 1.35**	29.05 ± 1.04	83.38 ± 1.22	27.66 ± 0.50	85.47 ± 1.57	28.12 ± 1.26	91.07 ± 0.84	24.37 ± 0.69	69.15 ± 1.60	24.44 ± 0.76	66.70 ± 2.19

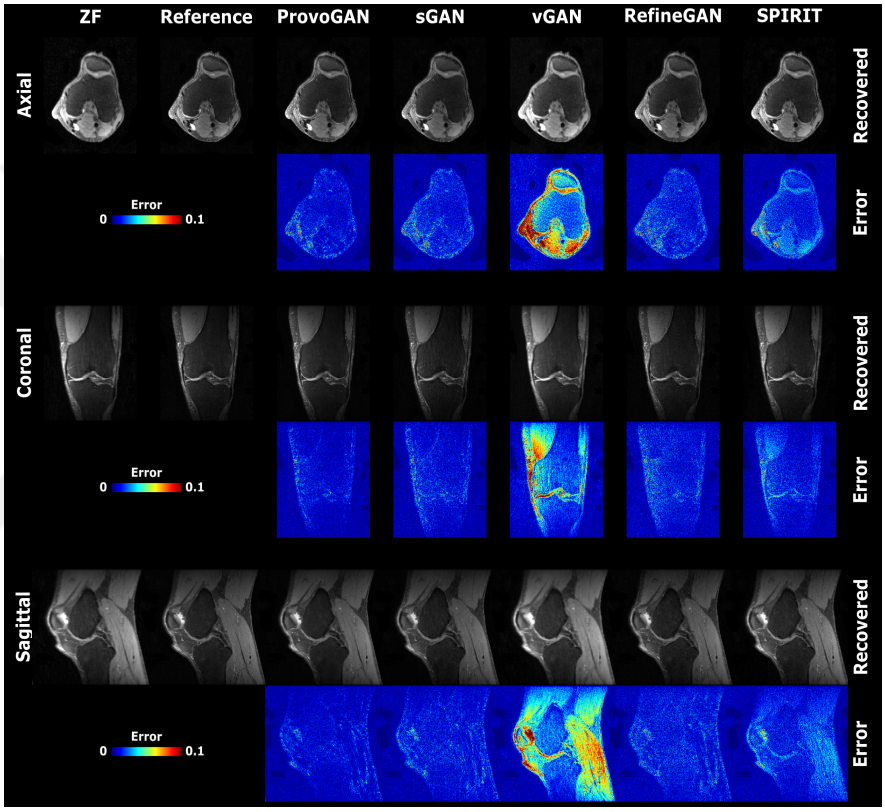


Figure 4.2: The proposed method is demonstrated on the in vivo multi-coil knee dataset for reconstruction at an acceleration ratio of $R = 8$. Representative results are displayed for all competing methods together with the zero-filled (ZF) undersampled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, ProvoGAN achieves sharper tissue depiction compared to vGAN, and alleviates undesirable discontinuities compared to cross-sectional models (sGAN, RefineGAN, SPIRiT) by improving reconstruction performance in all orientations.

Table 4.6: Quality of Reconstruction in the In vivo Knee Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the in vivo knee dataset are given as mean ± std for the test set. The measurements are reported for zero-filled images (ZF), the proposed ProvoGAN and competing sGAN, vGAN, RefineGAN, and SPIRiT methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method.

	ProvoGAN		sGAN		vGAN		RefineGAN		SPIRiT		ZF	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
R=4	**40.75**	**95.74**	40.34	95.69	36.80	92.79	40.31	95.21	39.46	95.35	32.17	93.50
	± 1.35	**± 0.94**	± 1.43	± 0.87	± 1.69	± 1.36	± 1.50	± 1.15	± 1.39	± 1.12	± 2.38	± 1.85
R=8	**39.45**	**95.13**	38.73	93.73	30.83	87.54	38.22	92.43	35.61	93.16	29.85	90.81
	± 2.15	**± 1.08**	± 1.01	± 0.98	± 1.44	± 2.80	± 1.43	± 1.58	± 2.70	± 1.60	± 1.77	± 1.88
R=12	**36.99**	**93.63**	36.76	91.99	29.03	88.10	36.21	90.82	33.58	91.66	28.65	89.17
	± 1.29	**± 0.95**	± 1.18	± 0.97	± 1.49	± 1.78	± 0.88	± 1.47	± 3.17	± 1.86	± 1.25	± 1.64
R=16	**37.86**	**92.27**	35.11	89.34	28.99	86.00	35.42	89.70	32.28	90.69	27.93	87.94
	± 0.50	**± 1.45**	± 1.06	± 1.61	± 2.21	± 2.28	± 0.55	± 1.13	± 2.63	± 1.85	± 1.25	± 1.60

4.3 Multi-Contrast MRI Synthesis

We further conducted experiments on the IXI and in vivo brain datasets for multi-contrast MRI synthesis to demonstrate ProvoGAN against state-of-the-art cross-sectional (sGAN) and volumetric (vGAN, SC-GAN, REPLICA) models (see Section 3.6 for details). We again measured volumetric PSNR and SSIM between the synthesized and reference target images for quantitative performance evaluation. In the IXI dataset, we considered synthesis tasks of T_2, PD $\rightarrow$ T_1, T_1, PD $\rightarrow$ T_2, and T_1, T_2 $\rightarrow$ PD. ProvoGAN outperforms the competing methods in all tasks ($p < 0.05$), where it achieves in the range of $[1.20, 2.90]$ dB higher PSNR and $[2.08, 4.37]$ % higher SSIM (see Table 4.7). In the in vivo brain dataset, we considered synthesis tasks of T_2, FLAIR, T_{1c} $\rightarrow$ T_1, T_1, FLAIR, T_{1c} $\rightarrow$ T_2, T_1, T_2, T_{1c} $\rightarrow$ FLAIR, and T_1, T_2, FLAIR $\rightarrow$ T_{1c}. ProvoGAN again yields enhanced recovery performance in all tasks compared to the competing methods ($p < 0.05$), where it maintains $[0.59, 5.01]$ dB higher PSNR and $[1.96, 5.36]$ % higher SSIM (see Table 4.8).

The superior synthesis quality offered by ProvoGAN is clearly visible in representative results displayed in Fig. 4.3 for the IXI dataset and Fig. 4.4 for the in vivo brain dataset. These results indicate that the cross-sectional sGAN-A,

sGAN-C, and sGAN-S models suffer from suboptimal recovery in the longitudinal dimension due to independent synthesis of cross-sections. Meanwhile volumetric vGAN, SC-GAN, and REPLICA models suffer from poor recovery of fine-structural details and loss of spatial resolution in the target images due to increased model complexity. In comparison to cross-sectional baselines, ProvoGAN alleviates discontinuity artifacts by pooling global contextual information via progressive execution of cross-sectional mappings. In comparison to volumetric baselines, ProvoGAN offers sharper and improved tissue depiction particularly near tumor regions due to its improved learning behavior. Overall, these findings demonstrate ProvoGAN's utility for diverse synthesis tasks in multi-contrast MRI exams.

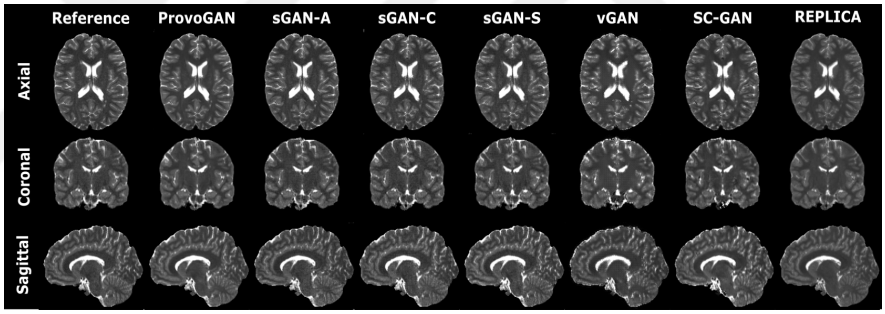


Figure 4.3: The proposed method is demonstrated on the IXI dataset for T_2-weighted image synthesis from T_1- and PD-weighted images. Representative results are displayed for all competing methods together with the reference target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Overall, the proposed method delineates tissues with higher spatial resolution compared to volumetric vGAN, SC-GAN, and REPLICA models, and alleviates discontinuity artifacts by improving synthesis performance in all orientations compared to cross-sectional sGAN-A, sGAN-C, and sGAN-S models.

Table 4.7: Quality of Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are provided for the proposed and competing methods for all synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method.

		ProvoGAN	sGAN-A	sGAN-C	sGAN-S	vGAN	SC-GAN	REPLICA
T_2, PD $\rightarrow$ T_1	PSNR	**24.15 ± 2.80**	23.20 ± 2.08	22.58 ± 2.11	23.65 ± 1.98	23.35 ± 2.89	22.58 ± 2.99	21.14 ± 4.15
	SSIM	**90.33 ± 4.47**	85.81 ± 3.95	86.60 ± 4.05	87.71 ± 4.15	85.48 ± 4.18	85.32 ± 4.00	86.30 ± 3.80
T_1, PD $\rightarrow$ T_2	PSNR	**28.97 ± 2.91**	27.64 ± 2.59	27.74 ± 2.67	27.93 ± 2.19	25.97 ± 1.81	25.29 ± 1.95	26.98 ± 2.37
	SSIM	**94.17 ± 4.16**	92.49 ± 4.20	92.671 ± 4.31	93.28 ± 2.88	90.61 ± 4.04	89.81 ± 4.31	92.51 ± 4.57
T_1, T_2 $\rightarrow$ PD	PSNR	**29.81 ± 2.96**	27.69 ± 2.20	29.00 ± 2.41	27.12 ± 1.61	26.17 ± 1.41	26.36 ± 1.41	26.96 ± 2.93
	SSIM	**95.41 ± 2.75**	93.64 ± 3.00	94.21 ± 2.99	92.67 ± 2.95	90.70 ± 2.47	92.12 ± 2.59	94.10 ± 3.03

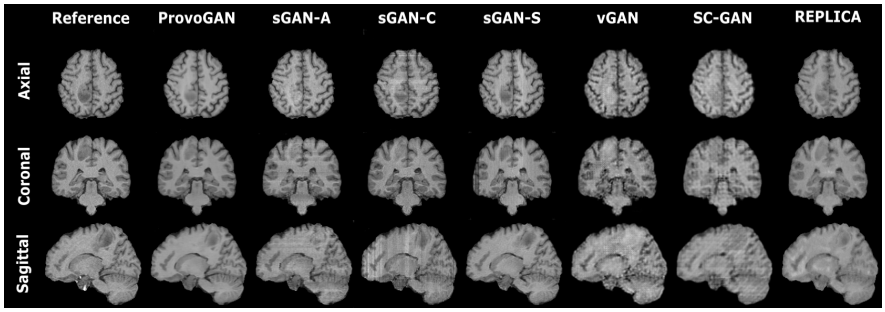


Figure 4.4: The proposed method is demonstrated on the in vivo brain dataset for T_1-weighted image synthesis from T_2-, T_{1c}-weighted and FLAIR images. Representative results are displayed for all competing methods together with the reference target images (first column). The first row displays results for the axial, the second row for the coronal, and the third row for the sagittal orientation. Overall, the proposed method delineates tissues with higher spatial resolution compared to volumetric vGAN, SC-GAN, and REPLICA models, and alleviates discontinuity artifacts by improving synthesis performance in all orientations compared to cross-sectional sGAN-A, sGAN-C, and sGAN-S models. Meanwhile, the proposed method achieves more accurate depictions for tumor regions, which are suboptimally recovered by the competing methods.

Table 4.8: Quality of Synthesis in the In vivo Brain Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set of the in vivo brain dataset are given as mean ± std. The measurements are provided for proposed and competing methods for all many-to-one synthesis tasks: 1) T_2, FLAIR, $T_{1c} \rightarrow T_1$, 2) T_1, FLAIR, $T_{1c} \rightarrow T_2$, 3) T_1, T_2, $T_{1c} \rightarrow$ FLAIR, 4) T_1, T_2, FLAIR $\rightarrow T_{1c}$. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method.

		ProvoGAN	sGAN-A	sGAN-C	sGAN-S	vGAN	SC-GAN	REPLICA
T_2, FLAIR, $T_{1c} \rightarrow T_1$	PSNR	**26.92 ± 4.55**	24.17 ± 3.83	25.31 ± 4.12	26.22 ± 3.09	22.73 ± 3.69	21.70 ± 2.96	17.14 ± 4.43
	SSIM	**94.24 ± 3.41**	88.23 ± 4.68	90.78 ± 4.09	91.17 ± 3.03	87.73 ± 3.60	85.60 ±3.45	83.34 ± 7.43
T_1, FLAIR, $T_{1c} \rightarrow T_2$	PSNR	**26.87 ± 2.40**	25.67 ± 1.75	25.98 ± 2.18	26.85 ± 2.38	25.48 ± 1.82	24.48 ± 1.72	24.68 ± 1.94
	SSIM	**92.79 ± 4.22**	89.76 ± 3.20	90.95 ± 4.10	92.11 ± 4.11	89.75 ± 3.93	88.63 ± 3.54	89.06 ± 3.16
T_1, T_2, $T_{1c} \rightarrow$ FLAIR	PSNR	**25.52 ± 2.22**	24.50 ± 1.84	24.95 ± 2.03	24.81 ± 2.21	22.94 ± 1.61	23.07 ± 2.23	22.70 ±2.79
	SSIM	**90.39 ± 3.06**	87.21 ± 2.73	88.09 ± 3.18	88.39 ± 2.67	85.67 ± 3.23	85.88 ±2.62	87.63 ± 2.97
T_1, T_2, FLAIR $\rightarrow T_{1c}$	PSNR	**29.67 ± 2.23**	28.53 ± 1.98	28.48 ± 2.11	28.75 ± 2.23	27.21 ± 1.46	27.57 ± 1.68	24.44 ± 2.49
	SSIM	**94.14 ± 2.09**	91.70 ± 2.34	89.91 ± 3.09	92.06 ± 2.44	89.11 ± 1.74	90.19 ± 2.00	90.09 ± 3.03

4.4 Demonstrations Against Hybrid Models

Having demonstrated the superior performance of ProvoGAN against several state-of-the-art cross-sectional and volumetric models, we conducted additional experiments to comparatively evaluate it against alternative volumetrization methods. In particular, ProvoGAN was compared with hybrid models based on fusion (M^3NET) and transfer learning strategies (TransferGAN) that both involve a mixture of cross-sectional and volumetric mappings (see Section 3.6 for details). Experiments were performed on the IXI dataset for accelerated MRI reconstruction and multi-contrast MRI synthesis. For reconstruction, T_1- and T_2-weighted image recovery tasks at four distinct acceleration factors ($R = 4, 8, 12, 16$) were examined. Table 4.9 reports performance measurements for the methods under

comparison. ProvoGAN yields superior performance compared to both hybrid models in all reconstruction tasks ($p < 0.05$), where on average it achieves 1.87 dB higher PSNR and 4.34 % higher SSIM compared to M^{3}NET, and 1.83 dB higher PSNR and 5.02 % higher SSIM compared to TransferGAN. Meanwhile, T_1, T_2 $\rightarrow$ PD, T_1, PD $\rightarrow$ T_2, T_2, PD $\rightarrow$ T_1 recovery tasks were considered for synthesis. The respective measurements are reported in Table 4.10. We find that ProvoGAN again maintains enhanced recovery performance in all synthesis tasks ($p < 0.05$), where it achieves an average of 2.94 dB higher PSNR and 2.81 % higher SSIM compared to M^{3}NET, and 0.87 dB higher PSNR and 1.76 % higher SSIM compared to TransferGAN.

Quantitative improvements that ProvoGAN offers are also visible in representative images displayed in Fig. 4.6 for reconstruction and in Fig. 4.5 for synthesis. The M^{3}NET model that performs 3D fusion of 2D model outputs at separate orientations moderately increases contextual sensitivity, but suffers from residual discontinuity artifacts and over-smoothing. Meanwhile, the TransferGAN model that transfers pretrained weights from a 2D model to condition the final 3D model improves learning behavior, but it suffers from elevated model complexity leading to loss of spatial resolution and structural details. In contrast, ProvoGAN yields enhanced recovery performance in all orientations with higher contextual consistency and sensitivity to structural details in the recovered images.

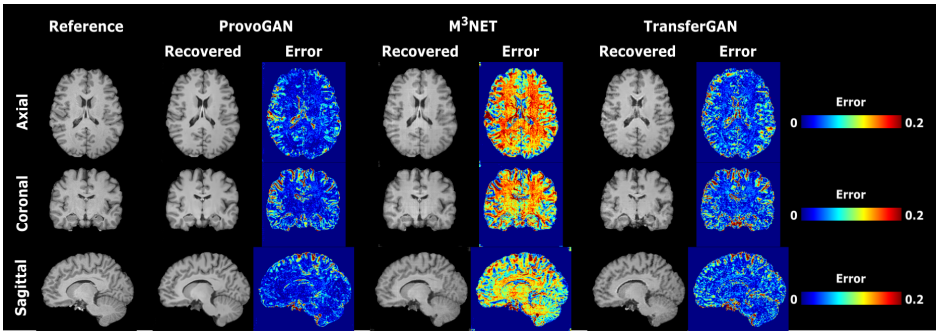


Figure 4.5: The proposed ProvoGAN method is demonstrated on the IXI dataset against hybrid models (M^{3}NET and TransferGAN) for T_1-weighted image synthesis from T_2- and PD-weighted images. Representative results are displayed for all methods under comparison together with the ground truth target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed method offers sharper and more accurate delineation of tissues than the competing methods. Furthermore, ProvoGAN better alleviates residual discontinuity artifacts compared to M^{3}NET.

Table 4.9: Comparison of Volumetrization Approaches for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN and competing M^{3}NET and TransferGAN methods for four distinct acceleration factors ($R = 4, 8, 12, 16$). Boldface indicates the best performing method.

		ProvoGAN		M^{3}NET		TransferGAN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
R=4	T_1	**35.25 ± 1.78**	**96.73 ± 0.57**	31.25 ± 1.87	92.07 ± 1.18	31.80 ± 0.88	89.11 ± 1.10
	T_2	**35.50 ± 2.62**	**96.08 ± 1.07**	33.85 ± 2.73	92.88 ± 1.12	34.18 ± 1.30	94.12 ± 0.84
R=8	T_1	**31.38 ± 1.26**	**94.93 ± 0.86**	29.61 ± 1.03	90.78 ± 1.05	30.37 ± 1.22	88.44 ± 0.99
	T_2	**33.49 ± 2.21**	**95.92 ± 1.01**	32.12 ± 2.25	93.09 ± 1.32	31.69 ± 0.83	93.49 ± 0.71
R=12	T_1	**29.67 ± 0.91**	**92.48 ± 0.90**	28.55 ± 1.05	87.61 ± 1.46	28.79 ± 0.89	85.31 ± 1.54
	T_2	**30.41 ± 1.03**	**91.98 ± 1.40**	30.29 ± 1.00	88.31 ± 1.08	29.18 ± 0.81	90.49 ± 0.94
R=16	T_1	**29.15 ± 1.09**	**91.40 ± 1.09**	25.83 ± 1.13	85.59 ± 1.90	27.61 ± 0.88	83.31 ± 1.65
	T_2	**30.66 ± 1.60**	**93.74 ± 1.35**	29.02 ± 1.24	88.03 ± 1.71	27.28 ± 1.69	88.84 ± 1.36

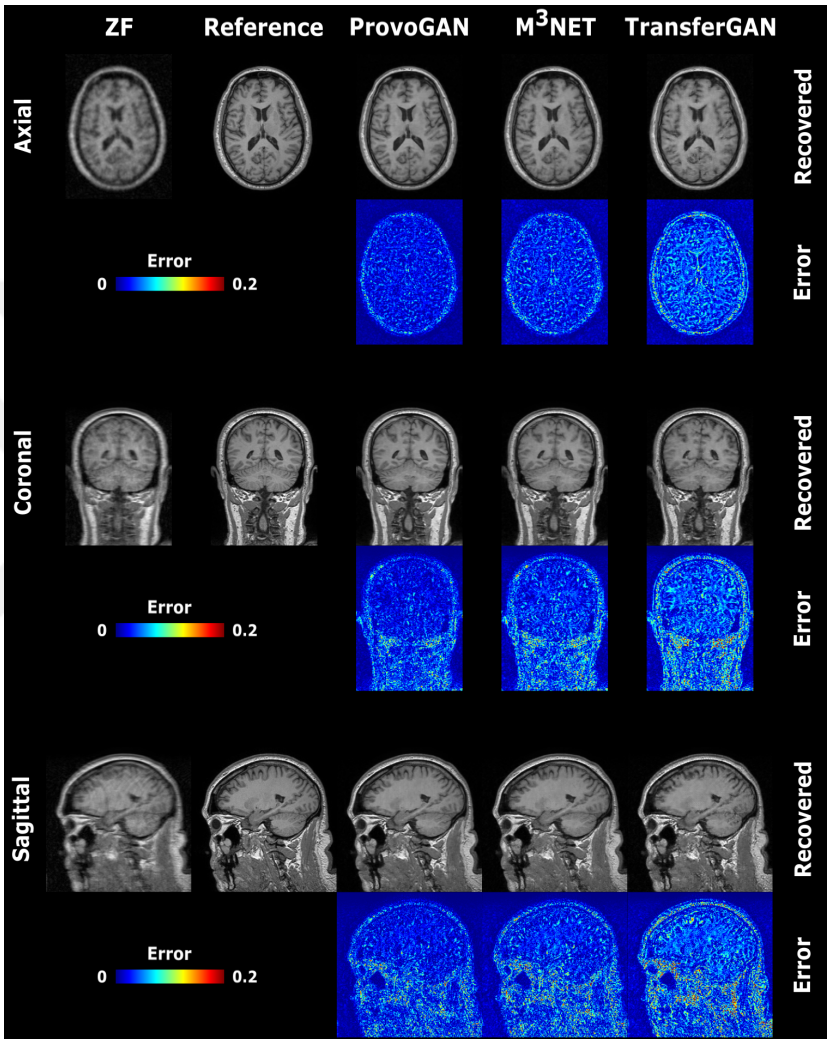


Figure 4.6: The proposed ProvoGAN method is demonstrated against hybrid models (M^{3}NET and TransferGAN) on the IXI dataset for reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. Representative results are displayed for the methods under comparison together with the undersampled zero-filled source images (first column) and the reference target images (second column). The top two rows display results for the axial, the middle two rows for the coronal, and the last two rows for the sagittal orientation. Error was taken as the absolute difference between the reconstructed and reference images (see colorbar). Overall, the proposed method enables sharper tissue delineation against competing methods, and it improves mitigation of discontinuity artifacts compared to M^{3}NET.

Table 4.10: Comparison of Volumetrization Approaches for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN and competing M^{3}NET and TransferGAN methods for all synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. Boldface indicates the highest performing method.

	ProvoGAN		M^{3}NET		TransferGAN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
T_2, PD $\rightarrow$ T_1	**24.15** ± **2.80**	**90.33** ± **4.47**	20.85 ± 4.08	86.81 ± 5.00	23.84 ± 3.37	87.09 ± 4.09
T_1, PD $\rightarrow$ T_2	**28.97** ± **2.91**	**94.17** ± **4.16**	23.79 ± 1.23	89.82 ± 3.12	27.78 ± 2.47	93.10 ± 3.98
T_1, T_2 $\rightarrow$ PD	**29.81** ± **2.96**	**95.41** ± **2.75**	29.48 ± 2.24	94.84 ± 2.41	28.71 ± 1.89	94.45 ± 2.52

4.5 Radiological Evaluation

Quantitative performance assessments in MRI recovery tasks clearly indicate that ProvoGAN outperforms competing volumetric and cross-sectional models in terms of image quality metrics (i.e., PSNR, SSIM). Yet, an important question concerns to what extent these quantitative improvements will benefit diagnostic assessments. Given its ability to effectively capture global context as well as fine structural details, we hypothesized that ProvoGAN will recover MR images of equivalent or higher diagnostic value than competing models. To test this hypothesis, radiological evaluations were performed on images recovered via ProvoGAN, sGAN and vGAN, as well as SC-GAN, REPLICA for synthesis, and RefineGAN, compressed-sensing methods (SparseMRI and SPIRIT) for reconstruction (see Section 2.7 for details). Opinion scores for all methods in axial, coronal, and sagittal orientations denoted as (OS_A, OS_C, OS_S) are reported in Fig. 4.7a-c for reconstruction and in Fig. 4.7d,e for synthesis. Across reconstruction tasks, ProvoGAN achieves average opinion scores of $(4.13, 3.97, 3.93)$ where sGAN yields $(3.33, 2.77, 2.67)$, vGAN yields $(2.00, 1.93, 1.70)$, RefineGAN yields $(3.63, 3.67, 3.13)$ and compressed-sensing reconstructions yield $(2.00, 2.27, 2.13)$. Across synthesis tasks, ProvoGAN achieves average opinion scores of $(3.73, 3.85, 4.10)$ whereas vGAN yields $(1.50, 1.63, 1.33)$, SC-GAN yields $(1.47, 1.43, 1.40)$ and REPLICA yields $(2.40, 2.17, 1.83)$. Meanwhile, transverse sGAN models [1] maintain $(2.90, 3.23, 3.38)$ and longitudinal sGAN models [2] yield $(1.64, 1.81, 1.47)$. Overall, ProvoGAN outperforms all competing methods in synthesis ($p < 0.05$, Wilcoxon signed-ranked test) and reconstruction ($p < 0.05$) tasks, except for RefineGAN and SPIRIT in the in-vivo knee dataset where the three methods perform similarly ($p > 0.05$). In synthesis, ProvoGAN surpasses not only longitudinal sGAN models in transverse dimensions but also transverse sGAN models in transverse dimensions for which they have been optimized.

[1] In radiological evaluation, an sGAN model is called transverse for those opinion scores given for the orientation where that sGAN model is trained, e.g., sGAN-A for OS_A.

[2] In radiological evaluation, an sGAN model is called longitudinal for those opinion scores given for the orientation where that sGAN model is not trained, e.g., sGAN-A for OS_C or OS_S.

Radiological evaluations were also performed on recovered images from M^{3}NET and TransferGAN, as competing volumetrization baselines. Opinion scores for all volumetrization approaches in axial, coronal, and sagittal orientations denoted as (OS_A, OS_C, OS_S) are reported for the IXI dataset in Fig. 4.8. Across reconstruction tasks, ProvoGAN achieves average opinion scores of $(4.00, 3.95, 3.90)$ where M^{3}NET yields $(2.95, 3.25, 3.50)$ and TransferGAN yields $(2.80, 2.75, 2.8)$. For synthesis, ProvoGAN achieves average opinion scores of $(4.00, 4.3, 4.2)$ where M^{3}NET yields $(2.90, 3.20, 3.00)$ and TransferGAN yields $(1.90, 2.10, 1.9)$. Overall, ProvoGAN outperforms competing volumetrization methods in both synthesis ($p < 0.05$) and reconstruction ($p < 0.05$) tasks. Taken together, these findings strongly suggest that ProvoGAN can offer improved diagnostic quality in accelerated multi-contrast MRI protocols.

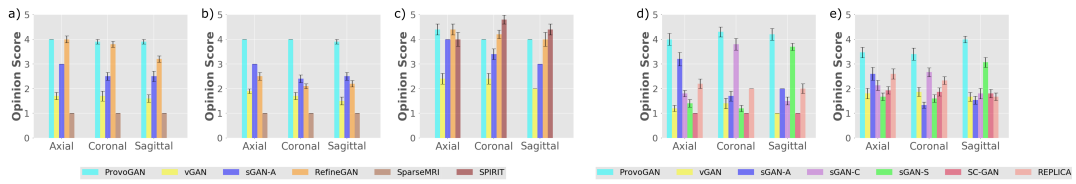


Figure 4.7: Methods were compared in terms of radiological opinion scores for three reconstruction tasks: a) single-coil reconstruction of T_1-weighted images undersampled by $R = 8$ in the IXI dataset, b) single-coil reconstruction of T_2-weighted images undersampled by $R = 8$ in the IXI dataset, c) multi-coil reconstruction of PD-weighted images undersampled by $R = 8$ in the in vivo knee dataset, and for two synthesis tasks: d) many-to-one synthesis task of T_2, PD $\rightarrow$ T_1 in the IXI dataset, e) many-to-one synthesis task of T_1, T_2, T_{1c} $\rightarrow$ FLAIR in the in vivo brain dataset. The quality of the recovered axial, coronal, and sagittal cross-sections were rated by an expert radiologist by assessing their similarity to the reference cross-sections via a five-point scale (0: unacceptable, 1: very poor, 2: limited, 3: moderate, 4: good, 5: perfect match). Figure legends denote the colors used for marking the methods under comparison.

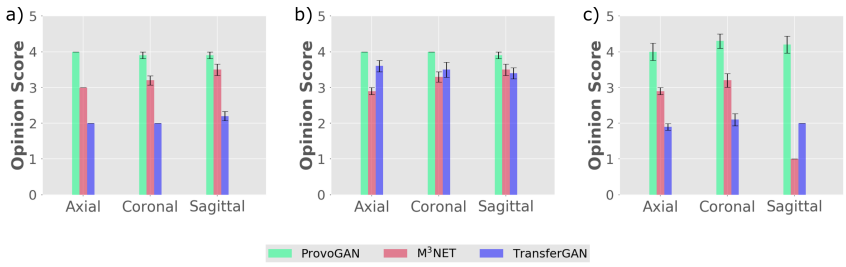


Figure 4.8: Volumetrization approaches were compared in terms of radiological opinion scores for three representative tasks: a) reconstruction of T_1-weighted images undersampled by $R = 8$ in the IXI dataset, b) reconstruction of T_2-weighted images undersampled by $R = 8$ in the IXI dataset, c) T_2, PD $\rightarrow$ T_1 synthesis in the IXI dataset. The quality of the recovered axial, coronal, and sagittal cross-sections were rated by an expert radiologist by assessing their similarity to the reference cross-sections via a five-point scale (0: unacceptable, 1: very poor, 2: limited, 3: moderate, 4: good, 5: perfect match). Figure legend denotes the colors used for the methods under comparison.

4.6 Complexity of Cross-Sectional Mappings

Model complexity in deep neural networks depends on several architectural choices, including the number of layers, number of filters in each layer, and kernel size. To minimize bias in performance comparisons, here we aligned the architectural designs as closely as possible among the competing methods. To do this, the number of layers, number of filters, and kernel size were all kept fixed across methods, except that 2D convolutional kernels were used in sGAN and ProvoGAN whereas 3D convolutional kernels were used in vGAN (see Figs. 3.2, 3.3). The precise parameter values were guided by the demanding vGAN model. We selected the parameter set that resulted in maximal model complexity while still allowing us to fit a single vGAN model into the VRAM of the GPUs used to conduct the experiments here. Thus, it is reasonable to consider that vGAN is at its performance limits. That said, a relevant question is whether and how the relative performance benefits of ProvoGAN over sGAN change with model complexity. To examine this issue, we performed separate experiments on reconstruction and synthesis tasks (see Section 3.7 for details) while systematically varying the complexity of the convolutional layers in both models by a factor of

$n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$. This resulted in seven distinct pairs of models: ProvoGAN(n_f)-sGAN(n_f) with n_f-fold change in number of learnable network weights. PSNR and SSIM measurements between the recovered and reference volumes reported in Table 4.11 demonstrate that ProvoGAN achieves superior reconstruction performance to sGAN at all complexity levels, with on average 1.42 dB higher PSNR and 3.20% higher SSIM ($p < 0.05$, Wilcoxon signed-rank test). Meanwhile, PSNR and SSIM measurements reported in Table 4.12 indicate that ProvoGAN increases synthesis performance on average by 1.22 dB in PSNR, and 2.82% in SSIM compared to sGAN across complexity levels ($p < 0.05$). Taken together, these findings suggest that the benefits of ProvoGAN over sGAN in MRI recovery tasks are reliable across variations in complexity of network layers.

An alternative approach to help improve performance of cross-sectional models without substantially altering model complexity would be to admit inputs from multiple neighboring cross-sections. Given several neighboring cross-sections as input, this would enable a 2D model to incorporate local context in the vicinity of the central cross-section. To examine the utility of this approach in cross-sectional processing, we implemented multi-cross-section variants of the two methods, namely ProvoGAN(multi) and sGAN(multi). Both variants received as input three consecutive cross-sections and learned to recover the central cross-section of the target volume. We postulated that while this approach might increase sGAN performance to a limited degree, ProvoGAN that leverages broad spatial priors across all orientations should still yield superior performance. To test this prediction, we performed comprehensive experiments on the IXI dataset for reconstruction and synthesis tasks (see Section 3.7 for details). PSNR and SSIM measurements were performed between the recovered and reference target volumes (see Tables 4.13, 4.14). Overall, ProvoGAN enhances recovery performance compared to sGAN(multi) in both tasks ($p < 0.05$), where it achieves on average 1.48 dB higher PSNR and 6.87% higher SSIM in reconstruction, and 0.87 dB higher PSNR and 1.71% higher SSIM in synthesis. These findings reveal that ProvoGAN outperforms cross-sectional mappings implemented with extended spatial priors across the longitudinal dimension. Note that ProvoGAN and ProvoGAN(multi) perform similarly across tasks ($p > 0.05$), whereas sGAN(multi) generally yields

on par or higher performance than sGAN. This result suggests that sGAN processing each cross-section independently suffers from loss of spatial context across the longitudinal dimension, whereas sGAN(multi) improves performance by incorporating short-range context across this dimension. In contrast, ProvoGAN(multi) captures limited additional information from multiple neighboring cross-sections, since ProvoGAN readily captures global context across the volume.

Comprehensive experiments performed on various reconstruction and synthesis tasks in the public IXI, in vivo knee, and in vivo brain datasets clearly demonstrate that ProvoGAN advances recovery quality in accelerated multi-contrast MRI compared to cross-sectional and volumetric models. Table 4.15 lists comparisons between sGAN, ProvoGAN and vGAN in terms of model complexity, memory load, number of floating point operations per second (FLOPS), and total training time. Compared to vGAN, ProvoGAN reduces model complexity by 3-fold, memory load by 20-fold, FLOPS by 80-fold approximately. Collectively, these benefits empower ProvoGAN to offer improved learning behavior and computationally efficient inference while enhancing the capture of global context in volumetric images. Compared to sGAN, each progression in ProvoGAN naturally maintains the same model complexity and memory load. As such, ProvoGAN offers a desirable solution in the trade-off between cross-sectional and volumetric models

Table 4.11: Model Complexity Analysis for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set of the IXI dataset are given as mean ± std for reconstruction of T_1-weighted acquisitions undersampled at $R = 8$. The measurements are reported for ProvoGAN and sGAN while varying the complexity of the convolutional layers in both models by $n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$, yielding seven distinct ProvoGAN-sGAN pairs: ProvoGAN(n_f)-sGAN(n_f) with n_f times folded number of learnable network weights. Boldface indicates the highest performing method.

		1/16	1/9	1/4	1	4	9	16
ProvoGAN	PSNR	**31.53 ± 1.31**	**31.02 ± 0.96**	**30.96 ± 0.85**	**31.38 ± 1.26**	**32.16 ± 1.80**	**32.23 ± 1.81**	**32.52 ± 1.76**
	SSIM	**93.74 ± 1.10**	**91.29 ± 1.02**	**90.62 ± 1.10**	**94.93 ± 0.86**	**95.68 ± 0.97**	**95.78 ± 1.00**	**95.81 ± 1.05**
sGAN-A	PSNR	29.58 ± 0.99	30.14 ± 0.93	30.09 ± 1.45	30.08 ± 1.32	30.17 ± 1.56	30.82 ± 1.61	30.96 ± 1.53
	SSIM	89.07 ± 1.19	89.84 ± 1.03	90.50 ± 1.07	91.18 ± 1.01	91.47 ± 1.15	91.73 ± 1.20	91.66 ± 1.13

Table 4.12: Model Complexity Analysis for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set of the IXI dataset are given as mean ± std for T_1-weighted image synthesis from T_2- and PD-weighted images. The measurements are reported for ProvoGAN and sGAN while varying the complexity of the convolutional layers in both models by $n_f \in \{1/16, 1/9, 1/4, 1, 4, 9, 16\}$, yielding seven distinct ProvoGAN-sGAN pairs: ProvoGAN(n_f)-sGAN(n_f) with n_f times folded number of learnable network weights. sGAN-A denotes the sGAN model trained in the axial orientation, sGAN-C in the coronal orientation, and sGAN-S in the sagittal orientation. Boldface indicates the highest performing method.

		1/16	1/9	1/4	1	4	9	16
ProvoGAN	PSNR	**24.25 ± 2.12**	**24.32 ± 2.11**	**24.10 ± 2.14**	**24.15 ± 2.80**	**24.74 ± 2.59**	**25.27 ± 2.65**	**25.36 ± 2.14**
	SSIM	**88.61 ± 4.16**	**88.49 ± 4.30**	**88.53 ± 4.23**	**90.33 ± 4.47**	**89.51 ± 4.39**	**90.05 ± 4.36**	**89.92 ± 4.34**
sGAN-A	PSNR	23.13 ± 2.03	23.13 ± 1.95	22.90 ± 1.78	23.20 ± 2.08	23.63 ± 2.22	23.96 ± 87.34	24.22 ± 1.76
	SSIM	84.84 ± 3.52	84.66 ± 3.61	85.08 ± 3.71	85.81 ± 3.95	86.51 ± 4.03	87.34 ± 3.95	87.38 ± 3.96
sGAN-C	PSNR	22.75 ± 2.51	22.82 ± 2.19	22.84 ± 2.04	22.58 ± 2.11	23.32 ± 2.35	23.77 ± 2.27	23.32 ± 2.51
	SSIM	85.24 ± 3.66	85.75 ± 3.93	86.21 ± 3.88	86.60 ± 4.05	87.47 ± 3.94	88.19 ± 4.04	88.10 ± 4.10
sGAN-S	PSNR	23.08 ± 1.63	22.33 ± 1.81	23.74 ± 1.76	23.65 ± 1.98	24.07 ± 2.14	24.14 ± 2.18	24.30 ± 2.17
	SSIM	85.23 ± 3.65	81.14 ± 4.65	87.43 ± 4.26	87.71 ± 4.15	88.76 ± 4.74	88.70 ± 4.81	88.90 ± 4.84

Table 4.13: Comparison of Single versus Multi-Cross-Section ProvoGAN and sGAN Models for Reconstruction in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the reconstructed and ground truth images in the test set of the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN model, its multi-cross-section variant, the cross-sectional sGAN model and its multi-cross-section variant for all single-coil reconstruction tasks: T_1-weighted and T_2-weighted image reconstruction at $R = 4, 8, 12, 16$. sGAN and sGAN(multi) are trained in the axial orientation given axial readout direction. Boldface indicates the highest performing method.

		R=4		R=8		R=12		R=16	
		T_1	T_2	T_1	T_2	T_1	T_2	T_1	T_2
ProvoGAN	PSNR	**35.25 ± 1.78**	35.50 ± 2.62	**31.83 ± 1.26**	**33.49 ± 2.21**	**29.67 ± 0.91**	**30.28 ± 1.31**	**29.14 ± 1.09**	**30.66 ± 1.60**
	SSIM	**96.73 ± 0.57**	96.08 ± 1.07	94.93 ± 0.86	**95.92 ± 1.01**	**92.48 ± 0.90**	91.96 ± 1.40	**91.40 ± 1.09**	**93.74 ± 1.35**
ProvoGAN(multi)	PSNR	34.65 ± 1.61	**36.36 ± 2.05**	31.71 ± 1.31	32.63 ± 1.58	29.26 ± 0.92	30.24 ± 1.05	28.44 ± 1.05	29.29 ± 1.09
	SSIM	96.27 ± 0.63	**96.73 ± 0.70**	**95.08 ± 0.95**	94.45 ± 0.91	88.58 ± 1.26	**92.10 ± 1.26**	88.11 ± 1.50	91.74 ± 1.27
sGAN	PSNR	33.85 ± 1.29	32.95 ± 1.50	30.08 ± 1.32	32.24 ± 2.14	27.34 ± 1.06	28.48 ± 1.06	26.73 ± 1.53	29.05 ± 1.04
	SSIM	93.21 ± 0.78	86.44 ± 1.23	91.18 ± 1.01	90.47 ± 0.95	86.23 ± 1.36	79.50 ± 2.15	85.23 ± 1.75	83.38 ± 1.22
sGAN(multi)	PSNR	33.71 ± 1.35	34.11 ± 1.58	31.00 ± 1.05	29.51 ± 0.93	28.43 ± 1.04	29.28 ± 1.27	27.80 ± 0.96	29.83 ± 1.28
	SSIM	93.69 ± 0.65	88.86 ± 1.24	92.17 ± 0.77	80.54 ± 1.88	88.03 ± 1.07	82.51 ± 1.69	87.07 ± 1.32	85.43 ± 1.22

Table 4.14: Comparison of Single and Multi-Cross-Section ProvoGAN and sGAN Models for Synthesis in the IXI Dataset: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. The measurements are reported for the proposed ProvoGAN model, its multi-cross-section variant, the cross-sectional sGAN model and its multi-cross-section variant for all many-to-one synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sGAN-A and sGAN(multi)-A denote the models trained in the axial orientation, sGAN-C and sGAN(multi)-C denote the models trained in the coronal orientation, and sGAN-S and sGAN(multi)-S denote the models trained in the sagittal orientation. Boldface indicates the highest performing method.

		T_2, PD $\rightarrow$ T_1	T_1, PD $\rightarrow$ T_2	T_1, T_2 $\rightarrow$ PD
ProvoGAN	PSNR	24.15 ± 2.80	**28.97 ± 2.91**	**29.81 ± 2.96**
	SSIM	90.33 ± 4.47	**94.17 ± 4.16**	**95.41 ± 2.75**
ProvoGAN(multi)	PSNR	24.23 ± 2.56	28.32 ± 2.69	29.64 ± 2.42
	SSIM	**90.49 ± 4.50**	93.52 ± 3.93	94.96 ± 2.62
sGAN-A	PSNR	23.20 ± 2.08	27.64 ± 2.60	27.70 ± 2.20
	SSIM	85.81 ± 3.95	92.94 ± 4.20	93.64 ± 3.00
sGAN(multi)-A	PSNR	23.55 ± 2.24	27.87 ± 2.54	28.69 ± 2.38
	SSIM	86.53 ± 4.26	93.09 ± 3.89	93.89 ± 3.07
sGAN-C	PSNR	22.56 ± 2.11	27.74 ± 2.67	29.00 ± 2.41
	SSIM	86.60 ± 4.05	92.67 ± 4.31	94.21 ± 2.99
sGAN(multi)-C	PSNR	22.96 ± 2.07	27.97 ± 2.61	29.18 ± 2.42
	SSIM	87.53 ± 4.27	92.97 ± 4.06	94.46 ± 2.97
sGAN-S	PSNR	23.65 ± 1.98	27.93 ± 2.19	27.12 ± 1.61
	SSIM	87.71 ± 4.15	93.28 ± 2.88	92.67 ± 2.95
sGAN(multi)-S	PSNR	**24.33 ± 2.29**	28.08 ± 2.18	28.36 ± 1.98
	SSIM	88.67 ± 4.45	93.43 ± 2.90	93.80 ± 2.87

Table 4.15: The proposed ProvoGAN and competing sGAN and vGAN methods are evaluated in terms instantaneous model complexity (millions of parameters, M), GPU VRAM use (gigabytes, GB), FLOPs (billions of floating point operation, G), and train duration (hours). The model complexity of the methods is given as (p_g, p_d), where p_g denotes the number of free parameters in the generator and p_d denotes the number of free parameters in the discriminator. The number of operations (FLOPs) of the methods is given as (F_g, F_d), where F_g denotes the number of operations in the generator and F_d denotes the number of operations in the discriminator. The input-output volume sizes and the number training subjects are also reported.

(a) Reconstruction in the IXI Dataset (Input Volume Size: $256 \times 150 \times 256$, Output Volume Size: $256 \times 150 \times 256$, Number of Training Subjects: 37)

	sGAN	vGAN	ProvoGAN
Model Complexity (M)	(1.60, 0.39)	(4.8, 1.17)	(1.6, 0.39)
GPU VRAM Usage (GB)	0.6	19	0.6
FLOPs (G)	(4.66, 0.26)	(1428.30, 31.79)	(17.55, 0.98)
Train Duration (hours)	11	20	33

(b) Reconstruction in the In vivo Knee Dataset (Input Subvolume Size: $320 \times 320 \times 256$, Output Subvolume Size: $320 \times 320 \times 256$, Number of Training Subjects: 12)

	sGAN	vGAN	ProvoGAN
Model Complexity (M)	(1.60, 0.39)	(4.8, 1.17)	(1.6, 0.39)
GPU VRAM Usage (GB)	0.8	19	0.8
FLOPs (G)	(10.07, 0.58)	(3860.27, 88.73)	(33.12, 1.90)
Train Duration (hours)	7	28	22

(c) Synthesis in the IXI Dataset (Input Volume Size: $2 \times 192 \times 160 \times 160$, Output Volume Size: $192 \times 160 \times 160$, Number of Training Subjects: 37)

	sGAN	vGAN	ProvoGAN
Model Complexity (M)	(1.60, 0.39)	(4.8, 1.17)	(1.6, 0.39)
GPU VRAM Usage (GB)	0.6	9	0.6
FLOPs (G)	(3.57, 0.20)	(723.80, 16.27)	(10.77, 0.60)
Train Duration (hours)	4	12	12

(d) Synthesis in the In-vivo Brain Dataset (Input Volume Size: $3 \times 192 \times 160 \times 160$, Output Volume Size: $192 \times 160 \times 160$, Number of Training Subjects: 237)

	sGAN	vGAN	ProvoGAN
Model Complexity (M)	(1.60, 0.39)	(4.8, 1.17)	(1.6, 0.39)
GPU VRAM Usage (GB)	0.6	9	0.6
FLOPs (G)	(3.60, 0.20)	(764.26, 16.98)	(10.87, 0.61)
Train Duration (hours)	6	16	18

4.7 Generalizability of Progressive Volumetrization

Here we primarily implemented ProvoGAN on a recent conditional GAN architecture with a ResNet backbone within the generator [59]. That said, progressive volumetrization can be viewed as a model-agnostic approach that can be adapted to various 2D network architectures. To illustrate the generalizability of ProvoGAN, we performed progressive volumetrization on another state-of-the-art architecture SC-GAN with a U-Net backbone injected with self-attention layers [34]. Cross-sectional (sSC-GAN), volumetric (vSC-GAN) and volumetrized (ProvoSC-GAN) variants of this architecture were built. Demonstrations were performed for T_1, $T_2 \rightarrow$ PD, T_1, PD $\rightarrow T_2$, and T_2, PD $\rightarrow T_1$ synthesis tasks on the IXI dataset. Resulting PSNR and SSIM measurements are listed in Table 4.16, where ProvoSC-GAN achieves in the range $[0.35, 2.42]$ dB higher PSNR and $[0.47, 3.05]$ % higher SSIM compared to sSC-GAN and vSC-GAN ($p < 0.05$). The superior synthesis quality offered by ProvoSC-GAN is also visible in representative results displayed in Fig. 4.9. Specifically, sSC-GAN models manifest discontinuity artifacts across the respective longitudinal dimensions, and vSC-GAN is suboptimal in recovering fine-structural details. In contrast, ProvoSC-GAN alleviates the limitations of both sSC-GAN and vSC-GAN models to enable more detailed and spatially-coherent tissue depiction. Taken together, these results strongly suggest that the proposed progressive volumetrization strategy can be extended to other network architectures while preserving its advantages against cross-sectional and volumetric mappings.

Table 4.16: Progressive Volumetrization of the SC-GAN Architecture: Volumetric PSNR (dB) and SSIM (%) measurements between the synthesized and ground truth images in the test set in the IXI dataset are given as mean ± std. Measurements are provided for proposed and competing methods for all many-to-one synthesis tasks: 1) T_2, PD $\rightarrow$ T_1, 2) T_1, PD $\rightarrow$ T_2, 3) T_1, T_2 $\rightarrow$ PD. sSC-GAN-A denotes the sSC-GAN model trained in the axial orientation, sSC-GAN-C in the coronal orientation, and sSC-GAN-S in the sagittal orientation. Boldface indicates the highest performing method.

	ProvoSC-GAN		sSC-GAN-A		sSC-GAN-C		sSC-GAN-S		vSC-GAN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
T_2, PD $\rightarrow$ T_1	**23.64** ± **3.12**	**88.27** ± **4.50**	22.74 ± 2.27	85.63 ± 3.74	23.03 ± 2.42	86.94 ± 3.80	23.30 ± 2.95	87.53 ± 4.46	22.58 ± 2.99	85.32 ± 4.00
T_1, PD $\rightarrow$ T_2	**28.20** ± **2.97**	**93.26** ± **4.53**	27.79 ± 2.71	92.64 ± 4.44	27.72 ± 2.74	92.45 ± 4.53	28.09 ± 2.86	93.22 ± 4.35	25.29 ± 1.95	89.81 ± 4.31
T_1, T_2 $\rightarrow$ PD	**29.66** ± **2.37**	**94.86** ± **2.96**	28.83 ± 2.12	94.02 ± 2.86	28.93 ± 2.11	93.93 ± 2.95	29.05 ± 1.92	94.21 ± 2.46	26.36 ± 1.41	92.12 ± 2.59

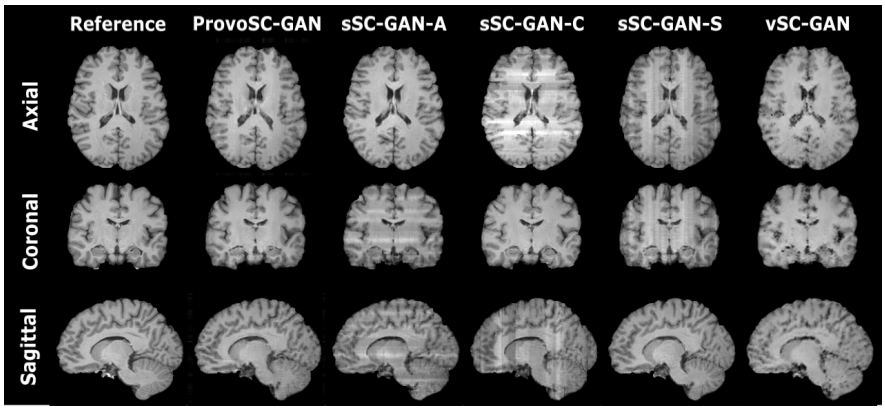


Figure 4.9: Progressive volumetrization was performed on recently proposed SC-GAN architecture. Representative results for T_1-weighted image synthesis from T_2- and PD-weighted images in the IXI dataset are displayed. Results are shown for progressively volumetrized (ProvoSC-GAN), cross-sectional (sSC-GAN), and volumetric (vSC-GAN) models, along with the ground truth target images (first column). The first row displays results for the axial orientation, the second row for the coronal orientation, and the third row for the sagittal orientation. Overall, ProvoSC-GAN improves delineation of structural details compared to vSC-GAN, and enhances contextual consistency in the longitudinal dimensions compared to sSC-GAN models.

Chapter 5

Conclusion

Here, we introduced a progressively volumetrized deep generative model (ProvoGAN) for accelerated MRI that decomposes complex volumetric image recovery tasks into a series of cross-sectional mappings task-optimally ordered across individual rectilinear orientations. This progressive decomposition empowers ProvoGAN to learn both global contextual priors and fine-structural details in each orientation with enhanced data efficiency. Comprehensive evaluations on brain and knee MRI datasets illustrate the superior performance of ProvoGAN against state-of-the-art volumetric and cross-sectional models. Compared to volumetric models, ProvoGAN better captures fine structural details while at the same time maintaining lower instantaneous model complexity. As subtasks in ProvoGAN take single cross-sections as separate training samples, the effective size of the training set is expanded. Therefore, for a given model complexity, ProvoGAN demands an order of magnitude lower memory load than volumetric models. Compared to cross-sectional models, ProvoGAN mitigates discontinuity artifacts across the longitudinal dimensions and extends reliable capture of structural details from transverse onto longitudinal dimensions. Importantly, ProvoGAN offers this advanced recovery performance for the same budget of model complexity and memory load as cross-sectional models, albeit at the expense of a three-times prolonged training procedure due to sequential learning.

Several recent studies in medical image processing have focused on improving learning behavior in volumetric models. An earlier group of studies proposed spatially-focused 3D models to process volumetric patches during MRI recovery [7, 10, 11, 87, 90–94]. Patch-based models that restrict the spatial extent of network inputs-outputs can reduce model size to offer performance improvements. That said, a compact 3D patch incorporates context along the longitudinal axis at the expense of narrowing coverage in the in-plane dimensions. Since patches are processed independently, the predicted volumes might also manifest discontinuity artifacts. These limitations can undercut potential benefits of patch-based processing for 3D models. Later studies proposed hybrid models to bridge 2D and 3D models in an effort to combine their strengths [65–68]. Among hybrid methods are fusion models that aggregate the outputs of parallel 2D models in multiple orientations [65, 66]. Fusion models employ a cascade of 2D and 3D processing, so they incur high computational complexity, and sensitivity to fine structural details might be limited by the aggregation process across orientations. An alternative approach is transfer learning from 2D onto 3D models to facilitate model training [67, 68]. A full-scale 3D model is leveraged in transfer learning methods that lead to elevated model complexity, and a similar computational footprint to conventional 3D models. In contrast, ProvoGAN is composed of a sequence of 2D models, without any 3D module, resulting in substantially lower model complexity and computational load.

An alternative approach to volumetrization in medical imaging tasks has been to revise cross-sectional models to help them better incorporate spatial context. In [95], enhanced spatial consistency during cardiac image segmentation was aimed by performing cross-sectional mapping on short-axis images sequentially across neighboring cross-sections. The segmentation map from the earlier cross-section was used to initialize the map for the current cross-section [95]. While benefits were demonstrated over 2D processing, this approach limits accumulation of contextual information to a single direction and to neighboring cross-sections. A different strategy for cardiac MRI segmentation was to perform cross-sectional mapping in short-axis orientation while latent representations captured via an autoencoder on a multitude of view orientations were fused

at intermediate layers [96]. The complexity of the resulting models scales with the number of additional views included, and this promising approach might be limited in applications where a multitude of different views on the same anatomy are unavailable. In [97–99], MR images at three rectilinear views that span across a target voxel were incorporated as inputs to a cross-sectional model during segmentation or classification tasks. Classifying a center voxel by fusing information across orientations might limit flow of contextual information from nearby voxels not covered by the input image views.

Several technical lines of development can be taken to further improve the performance and reliability of progressive volumetrization. In this study, ProvoGAN was independently demonstrated for mainstream MRI reconstruction and synthesis tasks. ProvoGAN can also be adopted for a joint reconstruction-synthesis task to further improve the utility and practicality of accelerated multi-contrast MRI protocols [60,100,101]. Here ProvoGAN was trained using a fully-supervised learning framework, which assumes the availability of datasets containing high-quality ground truth target images. However, compiling large datasets with high-quality references might prove difficult due to various concerns such as patient motion or examination costs [102]. An alternative would be to train ProvoGAN in a self-supervised setting for reconstruction tasks [103–107] or in a semi-supervised setting for synthesis tasks [101] to alleviate dependency on high-quality training datasets. Another avenue of development concerns the generalization of ProvoGAN to work on nonrectilinear orientations [108–110]. While ProvoGAN was mainly demonstrated for rectilinear acquisitions in this work, similar decompositions can be viable for nonrectilinear sampling schemes in MRI such as radial and spiral acquisitions. Additionally, the number of total progressions in ProvoGAN can be adaptively modified together with the specific ordering of the orientations used in the progressions to enhance task-optimal recovery performance. Instead of performing a separate sequential training of each progression, an end-to-end training of the whole network can also be performed for improved performance by leveraging advanced model parallelism techniques [111].

In this work, we demonstrated the proposed progressive volumetrization approach via a data-driven deep generative model that performs recovery in the

image domain. Although image-to-image learning of deep models proves popular in MRI recovery tasks [57, 112–116], there are other successful approaches to MRI processing based on k-space-to-k-space learning [117], k-space-to-image learning [44, 118–120], or model-based learning with unrolled network architectures [121–125]. In principle, ProvoGAN can also be implemented to volumetrize models based on these recent powerful approaches. Thus, it remains important future work to investigate the potential benefits of progressive volumetrization to the contextual sensitivity of a broader family of recovery methods [39, 112, 126–137].

In summary, here we introduced a progressive volumetrization framework for deep network models to process 3D imaging datasets. The superior learning behavior of ProvoGAN was demonstrated for inverse problem solutions in two mainstream MRI tasks, reconstruction and synthesis. Additional to structural MRI, proposed framework can be applied for other MRI applications such as functional [138–148, 148–150] and diffusion MRI [151, 152]. Yet, our framework can be adopted to other imaging modalities and tasks with minimal effort [33, 95, 98, 99, 153–162]. As the key idea of subtasking across cross-sectional orientations is domain general, ProvoGAN has further implications for computer vision applications that rely on 3D processing such as style transfer, semantic segmentation and video processing [163, 164].

Bibliography

[1] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology*, vol. 58, pp. R97–R129, jun 2013.

[2] G. Wang, J. C. Ye, and B. D. Man, "Deep learning for tomographic image reconstruction," *Nature Machine Intelligence*, vol. 2, pp. 737–748, 2020.

[3] J. C. Ye, "Compressed sensing MRI: a review from signal processing perspective," *BMC Biomedical Engineering*, vol. 1, 03 2019.

[4] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 2002.

[5] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–62, 1999.

[6] M. Lustig, D. Donoho, J. Santos, and J. Pauly, "Compressed Sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[7] A. Jog, A. Carass, S. Roy, D. L Pham, and J. Prince, "MR image synthesis by contrast learning on neighborhood ensembles," *Medical Image Analysis*, vol. 24, pp. 63–76, 05 2015.

[8] M. Yurt, S. U. Dar, A. Erdem, E. Erdem, K. K. Oguz, and T. Çukur, "mustGAN: Multi-stream generative adversarial networks for MR image synthesis," *Medical Image Analysis*, vol. 70, p. 101944, 2021.

[9] H. Van Nguyen, K. Zhou, and R. Vemulapalli, "Cross-domain synthesis of medical images using efficient location-sensitive deep network," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, (Cham), pp. 677–684, Springer International Publishing, 2015.

[10] S. Roy, Y.-Y. Chou, A. Jog, J. Butman, and D. L. Pham, "Patch based synthesis of whole head MR images: application to EPI distortion correction," in *Simulation and Synthesis in Medical Imaging* (S. A. Tsaftaris, A. Gooya, A. F. Frangi, and J. L. Prince, eds.), vol. 9968, pp. 146–156, 10 2016.

[11] S. Roy, A. Carass, and J. L. Prince, "Magnetic resonance image example-based contrast synthesis," *IEEE Transactions on Medical Imaging*, vol. 32, no. 12, pp. 2348–2363, 2013.

[12] M. Shahdloo, E. Ilicak, M. Tofighi, E. U. Saritas, A. E. Çetin, and T. Çukur, "Projection onto epigraph sets for rapid self-tuning compressed sensing mri," *IEEE transactions on medical imaging*, vol. 38, no. 7, pp. 1677–1689, 2018.

[13] T. Çukur, J. M. Santos, J. M. Pauly, and D. G. Nishimura, "Variable-density parallel imaging with partially localized coil sensitivities," *IEEE transactions on medical imaging*, vol. 29, no. 5, pp. 1173–1181, 2010.

[14] E. Biyik, E. Ilicak, and T. Cukur, "Reconstruction by calibration over tensors for multi-coil multi-acquisition balanced ssfp imaging," *Magnetic resonance in medicine*, vol. 79, no. 5, pp. 2542–2554, 2018.

[15] E. Ilicak, L. K. Senel, E. Biyik, and T. Çukur, "Profile-encoding reconstruction for multiple-acquisition balanced steady-state free precession imaging," *Magnetic resonance in medicine*, vol. 78, no. 4, pp. 1316–1329, 2017.

[16] E. Kopanoglu, A. Güngör, T. Kilic, E. U. Saritas, K. K. Oguz, T. Çukur, and H. E. Güven, "Simultaneous use of individual and joint regularization

terms in compressive sensing: Joint reconstruction of multi-channel multi-contrast mri acquisitions," *NMR in Biomedicine*, vol. 33, no. 4, p. e4247, 2020.

[17] L. K. Senel, T. Kilic, A. Gungor, E. Kopanoglu, H. E. Guven, E. U. Saritas, A. Koc, and T. Cukur, "Statistically segregated k-space sampling for accelerating multiple-acquisition mri," *IEEE transactions on medical imaging*, vol. 38, no. 7, pp. 1701–1714, 2019.

[18] E. Ilicak, S. Cetin, E. Bulut, K. K. Oguz, E. U. Saritas, G. Unal, and T. Çukur, "Targeted vessel reconstruction in non-contrast-enhanced steady-state free precession angiography," *NMR in Biomedicine*, vol. 29, no. 5, pp. 532–544, 2016.

[19] T. Çukur, "Accelerated phase-cycled ssfp imaging with compressed sensing," *IEEE transactions on medical imaging*, vol. 34, no. 1, pp. 107–115, 2014.

[20] T. Çukur, M. Lustig, E. U. Saritas, and D. G. Nishimura, "Signal compensation and compressed sensing for magnetization-prepared mr angiography," *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1017–1027, 2011.

[21] T. Çukur, J. H. Lee, N. K. Bangerter, B. A. Hargreaves, and D. G. Nishimura, "Non-contrast-enhanced flow-independent peripheral mr angiography with balanced ssfp," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 61, no. 6, pp. 1533–1539, 2009.

[22] T. Çukur and D. G. Nishimura, "Multiple repetition time balanced steady-state free precession imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 62, no. 1, pp. 193–204, 2009.

[23] T. Çukur, M. Lustig, and D. G. Nishimura, "Multiple-profile homogeneous image combination: application to phase-cycled ssfp and multicoil imaging," *Magnetic resonance in medicine*, vol. 60, no. 3, pp. 732–738, 2008.

[24] T. Çukur, N. K. Bangerter, and D. G. Nishimura, "Enhanced spectral shaping in steady-state free precession imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1216–1223, 2007.

[25] B. Gözcü, R. K. Mahabadi, Y.-H. Li, E. Ilıcak, T. Cukur, J. Scarlett, and V. Cevher, "Learning-based compressive mri," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1394–1406, 2018.

[26] M. Aksoy, C. Forman, M. Straka, T. Çukur, J. Hornegger, and R. Bammer, "Hybrid prospective and retrospective head motion correction to mitigate cross-calibration errors," *Magnetic resonance in medicine*, vol. 67, no. 5, pp. 1237–1251, 2012.

[27] E. Bıyık, K. Keskin, S. UH Dar, A. Koç, and T. Çukur, "Factorized sensitivity estimation for artifact suppression in phase-cycled bssfp mri," *NMR in Biomedicine*, vol. 33, no. 4, p. e4228, 2020.

[28] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[29] S. Roy, A. Carass, and J. Prince, "A compressed sensing approach for MR tissue contrast synthesis," in *Information Processing in Medical Imaging*, pp. 371–383, 2011.

[30] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Medical Image Analysis*, p. 101552, 2019.

[31] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[32] Y. Choi, Y. Uh, J. Yoo, and J.-W. Ha, "StarGAN v2: diverse image synthesis for multiple domains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8188–8197, 2020.

[33] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D deep learning on medical images: A review," *Sensors*, vol. 20, p. 5097, sep 2020.

[34] H. Lan, A. W. Toga, and F. Sepehrband, "SC-GAN: 3D self-attention conditional GAN with spectral normalization for multi-modal neuroimaging synthesis," *bioRxiv*, 2020.

[35] M. O. Malavé, C. A. Baron, S. P. Koundinyan, C. M. Sandino, F. Ong, J. Y. Cheng, and D. G. Nishimura, "Reconstruction of undersampled 3D non-Cartesian image-based navigators for coronary MRA using an unrolled deep learning model," *Magnetic Resonance in Medicine*, vol. 84, no. 2, pp. 800–812, 2020.

[36] H. Yang, X. Lu, S.-H. Wang, Z. Lu, J. Yao, Y. Jiang, and P. Qian, "Synthesizing multi-contrast MR images via novel 3D conditional variational auto-encoding GAN," *Mobile Networks and Applications*, 2020.

[37] B. Yu, L. Zhou, L. Wang, J. Fripp, and P. Bourgeat, "3D cGAN based cross-modality MR image synthesis for brain tumor segmentation," in *IEEE 15th International Symposium on Biomedical Imaging*, pp. 626–630, IEEE, apr 2018.

[38] B. Yu, L. Zhou, L. Wang, Y. Shi, J. Fripp, and P. Bourgeat, "Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 1750–1762, 2019.

[39] C. Chen, C. Raymond, B. Speier, X. Jin, T. F. Cloughesy, D. Enzmann, B. M. Ellingson, and C. W. Arnold, "Synthesizing MR image contrast enhancement using 3D high-resolution ConvNets," *arXiv preprint arXiv:2104.01592*, 2021.

[40] R. R. Sood, W. Shao, C. Kunder, N. C. Teslovich, J. B. Wang, S. J. Soerensen, N. Madhuripan, A. Jawahar, J. D. Brooks, P. Ghanouni, *et al.*, "3D registration of pre-surgical prostate mri and histopathology images via

super-resolution volume reconstruction," *Medical Image Analysis*, vol. 69, p. 101957, 2021.

[41] T. Küstner, N. Fuin, K. Hammernik, A. Bustin, H. Qi, R. Hajhosseiny, P. G. Masci, R. Neji, D. Rueckert, R. M. Botnar, *et al.*, "CINENet: deep learning-based 3D cardiac CINE MRI reconstruction with multi-coil complex-valued 4D spatio-temporal convolutions," *Scientific reports*, vol. 10, no. 1, pp. 1–13, 2020.

[42] H. El-Rewaidy, U. Neisius, J. Mancio, S. Kucukseymen, J. Rodriguez, A. Paskavitz, B. Menze, and R. Nezafat, "Deep complex convolutional network for fast reconstruction of 3D late gadolinium enhancement cardiac MRI," *NMR in Biomedicine*, vol. 33, no. 7, p. e4312, 2020.

[43] C. K. Chong and E. T. W. Ho, "Synthesis of 3D mri brain images with shape and texture generative adversarial deep neural networks," *IEEE Access*, vol. 9, pp. 64747–64760, 2021.

[44] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magnetic Resonance in Medicine*, vol. 81, no. 1, pp. 439–453, 2019.

[45] A. Chartsias, T. Joyce, M. V. Giuffrida, and S. A. Tsaftaris, "Multimodal MR synthesis via modality-invariant latent representation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 3, pp. 803–814, 2018.

[46] J. Y. Cheng, F. Chen, M. T. Alley, J. M. Pauly, and S. S. Vasanawala, "Highly scalable image reconstruction using deep neural networks with bandpass filtering," *arXiv preprint arXiv:1805.03300v2*, 2018.

[47] D. Lee, W.-J. Moon, and J. C. Ye, "Assessing the importance of magnetic resonance contrasts using collaborative generative adversarial networks," *Nature Machine Intelligence*, vol. 2, pp. 34–42, 01 2020.

[48] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanawala, G. Zaharchuk, L. Xing, and J. M. Pauly, "Deep generative adversarial neural networks for

compressive sensing MRI," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 167–179, 2019.

[49] M. Mardani, H. Monajemi, V. Papyan, S. Vasanawala, D. Donoho, and J. M. Pauly, "Recurrent generative adversarial networks for proximal learning and automated compressive image recovery," *arXiv preprint arXiv:1711.10046*, 2017.

[50] B. Zhan, D. Li, X. Wu, J. Zhou, and Y. Wang, "Multi-modal MRI image synthesis via GAN with multi-scale gate mergence," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, 2021.

[51] T. Joyce, A. Chartsias, and S. A. Tsaftaris, "Robust multi-modal MR image synthesis," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 347–355, 2017.

[52] Y. Chen, D. Firmin, and G. Yang, "Wavelet improved GAN for MRI reconstruction," in *Medical Imaging 2021: Physics of Medical Imaging*, vol. 11595, p. 1159513, International Society for Optics and Photonics, 2021.

[53] G. Li, J. Lv, and C. Wang, "A modified generative adversarial network using spatial and channel-wise attention for CS-MRI reconstruction," *IEEE Access*, 2021.

[54] B. Zhan, D. Li, Y. Wang, Z. Ma, X. Wu, J. Zhou, and L. Zhou, "LR-cGAN: latent representation based conditional generative adversarial network for multi-modality MRI synthesis," *Biomedical Signal Processing and Control*, vol. 66, p. 102457, 2021.

[55] T. Wang, Y. Lei, Y. Fu, J. F. Wynne, W. J. Curran, T. Liu, and X. Yang, "A review on medical imaging synthesis using deep learning and its clinical applications," *Journal of Applied Clinical Medical Physics*, vol. 22, no. 1, pp. 11–36, 2021.

[56] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine and Biology*, vol. 63, no. 13, p. 135007, 2018.

[57] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction with cyclic loss in generative adversarial networks," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.

[58] J. Schlemper, J. Caballero, J. V. Hajnal, A. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for MR image reconstruction," in *International Conference on Information Processing in Medical Imaging*, pp. 647–658, 2017.

[59] S. U. H. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, "Image synthesis in multi-contrast MRI with conditional generative adversarial networks," *IEEE Transactions on Medical Imaging*, vol. 90, no. 312, pp. 2375–2388, 2019.

[60] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, B. Tınaz, and T. Çukur, "Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1072–1087, 2020.

[61] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3055–3071, 2017.

[62] Y. Han, J. Yoo, H. H. Kim, H. J. Shin, K. Sung, and J. C. Ye, "Deep learning with domain adaptation for accelerated projection-reconstruction MR," *Magnetic Resonance in Medicine*, vol. 80, no. 3, pp. 1189–1205, 2018.

[63] D. Lee, J. Kim, W. Moon, and J. C. Ye, "Collagan: Collaborative GAN for missing image data imputation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2482–2491, June 2019.

[64] V. Sevetlidis, M. V. Giuffrida, and S. A. Tsaftaris, "Whole image synthesis using a deep encoder-decoder network," in *Simulation and Synthesis in Medical Imaging* (S. A. Tsaftaris, A. Gooya, A. F. Frangi, and J. L. Prince, eds.), (Cham), pp. 127–137, Springer International Publishing, 2016.

[65] J. Wei, Y. Xia, and Y. Zhang, "M3Net: A multi-model, multi-size, and multi-view deep neural network for brain magnetic resonance image segmentation," *Pattern Recognition*, vol. 91, pp. 366 – 378, 2019.

[66] C. Peng, W. Lin, H. Liao, R. Chellappa, and S. Zhou, "SAINT: Spatially aware interpolation network for medical slice synthesis," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (Los Alamitos, CA, USA), pp. 7747–7756, IEEE Computer Society, jun 2020.

[67] H. Shan, Y. Zhang, Q. Yang, U. Kruger, M. K. Kalra, L. Sun, W. Cong, and G. Wang, "3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1522–1534, 2018.

[68] S. Liu, D. Xu, S. K. Zhou, O. Pauly, S. Grbic, T. Mertelmeier, J. Wicklein, A. Jerebko, W. Cai, and D. Comaniciu, "3D anisotropic hybrid network: Transferring convolutional features from 2D images to 3D anisotropic volumes," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 851–858, Springer, 2018.

[69] M. Özbey, M. Yurt, S. U. H. Dar, and T. Çukur, "Three Dimensional MR Image Synthesis with Progressive Generative Adversarial Networks," in *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, (Virtual Conference),4 2020.

[70] M. Özbey, M. Yurt, S. U. H. Dar, and T. Çukur, "Three dimensional mr image synthesis with progressive generative adversarial networks," *arXiv preprint arXiv:2101.05218*, 2020.

[71] M. Yurt, M. Özbey, S. U. H. Dar, B. Tınaz, K. K. Oğuz, and T. Çukur, "Progressively volumetrized deep generative models for data-efficient contextual learning of mr image recovery," *arXiv preprint arXiv:2011.13913*, 2020.

[72] M. Yurt, M. Özbey, S. U. H. Dar, B. Tınaz, K. K. Oğuz, and T. Çukur, "Progressively Volumetrized Deep Generative Models for Data-Efficient

Contextual Learning of MR Image Recovery," in *29th Annual Meeting of International Society for Magnetic Resonance Imaging (ISMRM)*, (Virtual Conference),5 2021.

[73] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," in *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

[74] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," in *IEEE International Conference on Computer Vision (ICCV)*, pp. 2813–2821, Oct 2017.

[75] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976, July 2017.

[76] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision*, pp. 2223–2232, 2017.

[77] H. Li, J. C. Paetzold, A. Sekuboyina, F. Kofler, J. Zhang, J. S. Kirschke, B. Wiestler, and B. Menze, "DiamondGAN: Unified multi-modal generative adversarial networks for MRI sequences synthesis," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, (Cham), pp. 795–803, Springer International Publishing, 2019.

[78] S. Olut, Y. H. Sahin, U. Demir, and G. Unal, "Generative adversarial training for MRA image synthesis using multi-contrast MRI," in *PRedictive Intelligence in MEdicine (PRIME)*, (Cham), pp. 147–154, Springer International Publishing, 2018.

[79] A. Sharma and G. Hamarneh, "Missing MRI pulse sequence synthesis using multi-modal generative adversarial network," *IEEE Transactions on Medical Imaging*, vol. 39, pp. 1170–1183, 2019.

[80] Q. Yang, N. Li, Z. Zhao, X. Fan, E. I-Chao Chang, and Y. Xu, "MRI image-to-image translation for cross-modality image registration and segmentation," *arXiv preprint arXiv:1801.06940v1*, 2018.

[81] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 11 2014.

[82] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *International workshop on simulation and synthesis in medical imaging*, pp. 1–11, Springer, 2018.

[83] K. Epperson, A. M. Sawyer, M. Lustig, M. Alley, M. Uecker, P. Virtue, P. Lai, and S. Vasanawala, "Creation of fully sampled MR data repository for compressed sensing of the knee," in *2013 Meeting Proceedings. Section for Magnetic Resonance Technologists*, 2013.

[84] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[85] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[86] M. Lustig and J. M. Pauly, "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space.," *Magnetic Resonance in Medicine*, vol. 64, no. 2, pp. 457–71, 2010.

[87] A. Jog, A. Carass, S. Roy, D. L. Pham, and J. L. Prince, "Random forest regression for magnetic resonance image synthesis," *Medical Image Analysis*, vol. 35, pp. 475–488, 2017.

[88] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.

[89] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA," *Magnetic resonance in medicine*, vol. 71, no. 3, pp. 990–1001, 2014.

[90] N. Cordier, H. Delingette, M. Le, and N. Ayache, "Extended modality propagation: Image synthesis of pathological cases," *IEEE Transactions on Medical Imaging*, vol. 35, no. 12, pp. 2598–2608, 2016.

[91] Y. Huang, L. Shao, and A. F. Frangi, "Cross-modality image synthesis via weakly-coupled and geometry co-regularized joint dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 3, pp. 815 – 827, 2018.

[92] S. Roy, A. Carass, N. Shiee, D. L. Pham, and J. L. Prince, "MR contrast synthesis for lesion segmentation," in *IEEE International Symposium on Biomedical Imaging*, pp. 932–935, 2010.

[93] R. Vemulapalli, H. V. Nguyen, and S. K. Zhou, "Unsupervised cross-modal synthesis of subject-specific scans," in *IEEE International Conference on Computer Vision*, pp. 630–638, 2015.

[94] D. H. Ye, D. Zikic, B. Glocker, A. Criminisi, and E. Konukoglu, "Modality propagation: Coherent synthesis of subject-specific scans with data-driven regularization," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 606–613, 2013.

[95] Q. Zheng, H. Delingette, N. Duchateau, and N. Ayache, "3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2137–2148, 2018.

[96] C. Chen, C. Biffi, G. Tarroni, S. Petersen, W. Bai, and D. Rueckert, "Learning shape priors for robust cardiac MR segmentation from multi-view images," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, eds.), (Cham), pp. 523–531, Springer International Publishing, 2019.

[97] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar

convolutional neural network," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013* (K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, eds.), (Berlin, Heidelberg), pp. 246–253, Springer Berlin Heidelberg, 2013.

[98] S. Wang, M. Zhou, O. Gevaert, Z. Tang, D. Dong, Z. Liu, and T. Jie, "A multi-view deep convolutional neural networks for lung nodule segmentation," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1752–1755, 2017.

[99] Y. Xie, Y. Xia, J. Zhang, Y. Song, D. D. F. Feng, M. Fulham, and W. Cai, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 991–1004, 04 2019.

[100] J. E. Iglesias, B. Billot, Y. Balbastre, A. Tabari, J. Conklin, R. G. González, D. C. Alexander, P. Golland, B. L. Edlow, B. Fischl, *et al.*, "Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution and contrast," *NeuroImage*, p. 118206, 2021.

[101] M. Yurt, S. U. H. Dar, B. Tınaz, M. Özbey, and T. Çukur, "Semi-supervised learning of mutually accelerated multi-contrast MRI synthesis without fully-sampled ground-truths," *arXiv preprint arXiv:2011.14347*, 2020.

[102] J. Liu, A. I. Aviles-Rivero, H. Ji, and C.-B. Schönlieb, "Rethinking medical image reconstruction via shape prior, going deeper and faster: Deep joint indirect registration and reconstruction," *Medical Image Analysis*, vol. 68, p. 101930, 2021.

[103] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magnetic resonance in medicine*, vol. 84, no. 6, pp. 3172–3191, 2020.

[104] O. B. Demirel, B. Yaman, L. Dowdle, S. Moeller, L. Vizioli, E. Yacoub, J. Strupp, C. A. Olman, K. Uğurbil, and M. Akçakaya, "Improved simultaneous multi-slice functional MRI using self-supervised deep learning," *arXiv preprint arXiv:2105.04532*, 2021.

[105] E. K. Cole, J. M. Pauly, S. S. Vasanawala, and F. Ong, "Unsupervised MRI reconstruction with generative adversarial networks," *arXiv preprint arXiv:2008.13065*, 2020.

[106] Y. Korkmaz, S. U. Dar, M. Yurt, M. Özbey, and T. Çukur, "Unsupervised MRI reconstruction via zero-shot learned adversarial transformers," *arXiv preprint arXiv:2105.08059*, 2021.

[107] S. U. H. Dar, M. Yurt, and T. Çukur, "A few-shot learning approach for accelerated mri via fusion of data-driven and subject-driven priors," *arXiv preprint arXiv:2103.07790*, 2021.

[108] Z. Ramzi, J.-L. Starck, and P. Ciuciu, "Density compensated unrolled networks for Non-Cartesian MRI reconstruction," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 1443–1447, IEEE, 2021.

[109] C. Sun, Y. Yang, X. Cai, M. Salerno, C. H. Meyer, D. Weller, and F. H. Epstein, "Non-cartesian slice-GRAPPA and slice-SPIRiT reconstruction methods for multiband spiral cardiac MRI," *Magnetic resonance in medicine*, vol. 83, no. 4, pp. 1235–1249, 2020.

[110] S. Motyka, L. Hingerl, B. Strasser, G. Hangel, E. Heckova, A. Agibetov, G. Dorffner, S. Gruber, S. Trattning, and W. Bogner, "k-space-based coil combination via geometric deep learning for reconstruction of non-Cartesian MRSI data," *Magnetic Resonance in Medicine*, 2021.

[111] W. Zhu, C. Zhao, W. Li, H. Roth, Z. Xu, and D. Xu, "LAMP: Large deep nets with automated model parallelism for image segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* (A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga,

S. K. Zhou, D. Racoceanu, and L. Joskowicz, eds.), (Cham), pp. 374–384, Springer International Publishing, 2020.

[112] S. Wang, T. Xiao, Q. Liu, and H. Zheng, "Deep learning for fast MR imaging: a review for learning reconstruction from incomplete k-space data," *Biomedical Signal Processing and Control*, vol. 68, p. 102579, 2021.

[113] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310–1321, 2018.

[114] X. Dai, Y. Lei, Y. Fu, W. J. Curran, T. Liu, H. Mao, and X. Yang, "Multimodal MRI synthesis using unified generative adversarial networks," *Medical Physics*, vol. 47, no. 12, pp. 6343–6354, 2020.

[115] E. Cole, J. Cheng, J. Pauly, and S. Vasanawala, "Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications," *Magnetic resonance in medicine*, vol. 86, no. 2, pp. 1093–1109, 2021.

[116] G. Wang, E. Gong, S. Banerjee, D. Martin, E. Tong, J. Choi, H. Chen, M. Wintermark, J. M. Pauly, and G. Zaharchuk, "Synthesize high-quality multi-contrast magnetic resonance imaging from multi-echo acquisition using multi-task deep generative model," *IEEE transactions on medical imaging*, vol. 39, no. 10, pp. 3089–3099, 2020.

[117] Y. Han, L. Sunwoo, and J. C. Ye, "k-space deep learning for accelerated MRI," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 377–386, 2020.

[118] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, "KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images," *Magnetic resonance in medicine*, vol. 80, no. 5, pp. 2188–2201, 2018.

[119] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.

[120] S. Wang, Z. Ke, H. Cheng, S. Jia, L. Ying, H. Zheng, and D. Liang, "DIMENSION: dynamic MR imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training," *NMR in Biomedicine*, p. e4131, 2019.

[121] J. Sun, H. Li, Z. Xu, *et al.*, "Deep admm-net for compressive sensing mri," *Advances in neural information processing systems*, vol. 29, 2016.

[122] J. Zhang and B. Ghanem, "ISTA-Net: interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1828–1837, 2018.

[123] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 394–405, 2018.

[124] J. Duan, J. Schlemper, C. Qin, C. Ouyang, W. Bai, C. Biffi, G. Bello, B. Statton, D. P. O'regan, and D. Rueckert, "VS-Net: variable splitting network for accelerated parallel MRI reconstruction," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 713–722, Springer, 2019.

[125] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 3, pp. 521–538, 2018.

[126] H. Yang, X. Lu, S.-H. Wang, Z. Lu, J. Yao, Y. Jiang, and P. Qian, "Synthesizing multi-contrast mr images via novel 3D conditional variational auto-encoding GAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 415–424, 2021.

[127] O. Dalmaz, M. Yurt, and T. Çukur, "ResViT: residual vision transformers for multi-modal medical image synthesis," *arXiv preprint arXiv:2106.16031*, 2021.

[128] N. Tavaf, A. Torfi, K. Ugurbil, and P.-F. Van de Moortele, "GRAPPA-GANs for parallel MRI reconstruction," *arXiv preprint arXiv:2101.03135*, 2021.

[129] Y. Hu, Y. Xu, Q. Tian, F. Chen, X. Shi, C. J. Moran, B. L. Daniel, and B. A. Hargreaves, "RUN-UP: accelerated multishot diffusion-weighted MRI reconstruction using an unrolled network with U-Net as priors," *Magnetic Resonance in Medicine*, vol. 85, no. 2, pp. 709–720, 2021.

[130] T. Cukur, M. Lustig, and D. G. Nishimura, "Improving non-contrast-enhanced steady-state free precession angiography with compressed sensing," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 61, no. 5, pp. 1122–1131, 2009.

[131] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, "A transfer-learning approach for accelerated mri using deep neural networks," *Magnetic Resonance in Medicine*, vol. 84, no. 2, pp. 663–685, 2020.

[132] K. Keskin, U. Yılmaz, and T. Çukur, "Constrained ellipse fitting for efficient parameter mapping with phase-cycled bssfp mri," *arXiv preprint arXiv:2106.03239*, 2021.

[133] M. Özbey and T. Çukur, "Multi-image reconstruction in multi-contrast mri," in *2021 29th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, 2021.

[134] İ. Kaftan, T. Çukur, *et al.*, "Synergistic reconstruction-synthesis of multi-contrast mri using transfer learning method," in *2021 29th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, 2021.

[135] M. Özbey and T. Çukur, "T1-weighted contrast-enhanced synthesis for multi-contrast mri segmentation," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, 2020.

[136] T. Sanchez, B. Gözcü, R. B. van Heeswijk, A. Eftekhari, E. Ilıcak, T. Çukur, and V. Cevher, "Scalable learning-based sampling optimization for compressive dynamic mri," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8584–8588, IEEE, 2020.

[137] S. Türk, T. Cukur, C. HEPDURGUN, S. Dar, A. Karabulut, I. TAMSEL, U. Aydingoz, M. ARGIN, *et al.*, "Deep learning algorithm for sacroiliac joint evaluation: Grading with radiographs," European Congress of Radiology-ECR 2020, 2020.

[138] E. Çelik, S. U. H. Dar, Ö. Yılmaz, Ü. Keleş, and T. Çukur, "Spatially informed voxelwise modeling for naturalistic fmri experiments," *NeuroImage*, vol. 186, pp. 741–757, 2019.

[139] M. Shahdloo, E. Çelik, and T. Çukur, "Biased competition in semantic representation during natural visual search," *Neuroimage*, vol. 216, p. 116383, 2020.

[140] Ö. Yılmaz, E. Çelik, and T. Çukur, "Informed feature regularization in voxelwise modeling for naturalistic fmri experiments," *European Journal of Neuroscience*, vol. 52, no. 5, pp. 3394–3410, 2020.

[141] S. Tokgoz, D. Aydogdu, B. Ilhan, Y. Sahin, N. Bariseri, B. M. Ozturkler, and T. Çukur, "Musical mirror-symmetrical movement tasks: comparison of rhythm versus melody-playing," *NeuroReport*, vol. 31, no. 7, pp. 523–529, 2020.

[142] T. Çukur, A. G. Huth, S. Nishimoto, and J. L. Gallant, "Functional subdomains within human ffa," *Journal of Neuroscience*, vol. 33, no. 42, pp. 16748–16766, 2013.

[143] T. Çukur, A. G. Huth, S. Nishimoto, and J. L. Gallant, "Functional subdomains within scene-selective cortex: parahippocampal place area, retrosplenial complex, and occipital place area," *Journal of Neuroscience*, vol. 36, no. 40, pp. 10257–10273, 2016.

[144] T. Cukur, S. Nishimoto, A. G. Huth, and J. L. Gallant, "Attention during natural vision warps semantic representation across the human brain," *Nature neuroscience*, vol. 16, no. 6, pp. 763–770, 2013.

[145] T. Zhang, J. S. Gao, T. Çukur, and J. L. Gallant, "Voxel-based state space modeling recovers task-related cognitive states in naturalistic fmri experiments," *Frontiers in neuroscience*, vol. 14, p. 1103, 2021.

[146] I. Kiremitçi, Ö. Yilmaz, E. Çelik, M. Shahdloo, A. G. Huth, and T. Çukur, "Attentional modulation of hierarchical speech representations in a multitalker environment," *bioRxiv*, pp. 2020–12, 2021.

[147] E. Çelik, U. Keles, İ. Kiremitçi, J. L. Gallant, and T. Çukur, "Cortical networks of dynamic scene category representation in the human brain," *Cortex*, vol. 143, pp. 127–147, 2021.

[148] M. Shahdloo, E. Çelik, B. A. Urgen, J. L. Gallant, and T. Çukur, "Task-dependent warping of semantic representations during search for visual action categories," *bioRxiv*, 2021.

[149] L. K. Şenel, İ. Utlu, V. Yücesoy, A. Koç, and T. Çukur, "Generating semantic similarity atlas for natural languages," in *2018 IEEE Spoken Language Technology Workshop (SLT)*, pp. 795–799, IEEE, 2018.

[150] L. K. Şenel, I. Utlu, V. Yücesoy, A. Koc, and T. Cukur, "Semantic structure and interpretability of word embeddings," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1769–1779, 2018.

[151] E. U. Saritas, D. Lee, T. Cukur, A. Shankaranarayanan, and D. G. Nishimura, "Hadamard slice encoding for reduced-fov diffusion-weighted imaging," *Magnetic resonance in Medicine*, vol. 72, no. 5, pp. 1277–1290, 2014.

[152] S. G. Kafali, T. Çukur, and E. U. Saritas, "Phase-correcting non-local means filtering for diffusion-weighted imaging of the spinal cord," *Magnetic resonance in medicine*, vol. 80, no. 3, pp. 1020–1035, 2018.

[153] H. Chung, E. Cha, L. Sunwoo, and J. C. Ye, "Two-stage deep learning for accelerated 3D time-of-flight MRA without matched training data," *arXiv preprint arXiv 2008.01362*, 2020.

[154] B. Dewey, C. Zhao, J. Reinhold, A. Carass, K. Fitzgerald, E. Sotirchos, S. Saidha, J. Oh, D. Pham, P. Calabresi, P. van Zijl, and J. Prince, "DeepHarmony: A deep learning approach to contrast harmonization across scanner changes," *Magnetic Resonance Imaging*, vol. 64, pp. 160–170, Dec. 2019.

[155] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional recurrent neural networks for dynamic MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 280–290, 2019.

[156] C. Zhao, B. E. Dewey, D. L. Pham, P. A. Calabresi, D. S. Reich, and J. L. Prince, "SMORE: a self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2020.

[157] S. K. Zhou, H. Greenspan, C. Davatzikos, J. S. Duncan, B. van Ginneken, A. Madabhushi, J. L. Prince, D. Rueckert, and R. M. Summers, "A review of deep learning in medical imaging: Image traits, technology trends, case studies with progress highlights, and future promises," 2020.

[158] D. Narnhofer, A. Effland, E. Kobler, K. Hammernik, F. Knoll, and T. Pock, "Bayesian uncertainty estimation of learned variational MRI reconstruction," *arXiv preprint arXiv:2102.06665*, 2021.

[159] A. Ozaslan, A. Alacaoglu, O. Demirel, T. Çukur, and E. Saritas, "Fully automated gridding reconstruction for non-cartesian x-space magnetic particle imaging," *Physics in Medicine & Biology*, vol. 64, no. 16, p. 165018, 2019.

[160] B. Aşkın, A. Güngör, D. A. Soydan, C. B. Top, and T. Cukur, "A cnn based super-resolution technique for magnetic particle imaging system matrix," in *2021 29th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, 2021.

[161] O. B. Demirel, T. Kilic, T. Çukur, and E. U. Saritas, "Anatomical measurements correlate with individual magnetostimulation thresholds for khz-range homogeneous magnetic fields," *Medical physics*, vol. 47, no. 4, pp. 1836–1844, 2020.

[162] S. Ilbey, C. B. Top, A. Güngör, T. Çukur, E. U. Saritas, and H. E. Güven, "Fast system calibration with coded calibration scenes for magnetic particle imaging," *IEEE transactions on medical imaging*, vol. 38, no. 9, pp. 2070–2080, 2019.

[163] K. Aberman, Y. Weng, D. Lischinski, D. Cohen-Or, and B. Chen, "Unpaired motion style transfer from video to animation," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 4, pp. 64–1, 2020.

[164] J. Chen, S. Li, D. Liu, and W. Lu, "Indoor camera pose estimation via style-transfer 3D models," *Computer-Aided Civil and Infrastructure Engineering*, pp. 1–19, 2021.