HHS Public Access

Author manuscript

Med Image Comput Comput Assist Interv. Author manuscript; available in PMC 2020 October 01.

Published in final edited form as:

Med Image Comput Comput Assist Interv. 2019 October ; 11767: 786–794.

doi:10.1007/978-3-030-32251-9_86.

Wavelet-based Semi-supervised Adversarial Learning for Synthesizing Realistic 7T from 3T MRI

Liangqiong Qu, **Shuai Wang**, **Pew-Thian Yap**, **Dinggang Shen**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Abstract

Ultra-high field 7T magnetic resonance imaging (MRI) scanners produce images with exceptional anatomical details, which can facilitate diagnosis and prognosis. However, 7T MRI scanners are often cost prohibitive and hence inaccessible. In this paper, we propose a novel wavelet-based semi-supervised adversarial learning framework to synthesize 7T MR images from their 3T counterparts. Unlike most learning methods that rely on supervision requiring a significant amount of 3T-7T paired data, our method applies a semi-supervised learning mechanism to leverage unpaired 3T and 7T MR images to learn the 3T-to-7T mapping when 3T-7T paired data are scarce. This is achieved via a cycle generative adversarial network that operates in the joint spatial-wavelet domain for the synthesis of multi-frequency details. Extensive experimental results show that our method achieves better performance than state-of-the-art methods trained using fully paired data.

1 Introduction

The first 7T magnetic resonance (MR) scanner was approved for clinical use by the United States Food and Drug Administration in 2017. Compared to routine 3T MRI, 7T MRI typically affords greater anatomical details and higher signal-to-noise ratio, potentially facilitating diagnosis and prognosis [8]. However, 7T MR scanners are cost-prohibitive and hence less accessible at the clinics. For example, to date, there are less than 100 7T MR scanners compared with more than 20,000 3T MR scanners worldwide. Synthesizing 7T MR images from their 3T counterparts is thus valuable for both clinical and research applications.

In the past decades, a large number of learning based methods have been proposed for 7T MR image synthesis, including sparse learning, random forest, and deep learning. For example, Bahrami *et al.* [4] proposed a convolutional neural network taking into account appearance and anatomical features to learn the nonlinear 3T-to-7T mapping. Zhang *et al.* [9] utilized two multi-stage linear regression streams in both spatial and frequency domains for 7T MR image synthesis. While effective, these methods often share the following limitations. (i) Existing methods usually require a large amount of paired data (i.e., 3T and 7T MR images of the same subject) for training. However, acquiring paired data can be

ptyap@med.unc.edu, dgshen@med.unc.edu.

challenging and they are hence not always available in large quantities. (ii) Existing methods often fail to capture sufficient anatomical details and produce blurred results.

In this paper, we propose a wavelet-based semi-supervised adversarial network to synthesize 7T T1-weighted MR images from their 3T counterparts. Our method, called SemiWave, learns a two-way mapping by a small number of paired data together with a large amount of unpaired data. To learn from unpaired data, we couple the primal 3T-to-7T mapping task with a dual 7T-to-3T mapping task and impose a cycle-consistent loss [10] to regularize these two tasks, i.e., a synthetic 7T MR image generated by 3T-to-7T mapping should be mapped back to the original 3T MR image via a 7T-to-3T mapping. Specifically, we further transform these two mapping tasks from the spatial domain to a flexible wavelet coefficient predication task in a joint spatial-wavelet domain. Emphasis on the prediction of low-frequency wavelet coefficients enforces consistency in global information, whereas emphasis on high-frequency wavelet coefficients promotes reconstruction with better anatomical details. A flexible wavelet loss with adaptive weights on wavelet coefficients of different bands is introduced to help generate realistic 7T MR images.

As shown in Fig. 1, SemiWave consists of four main modules: (i) a standard adversarial loss module, (ii) a standard pair-wise reconstruction loss module, (iii) a cycle consistency loss module, and (iv) a wavelet loss module. Specifically, the cycle consistency loss is introduced to alleviate the need of large amounts of paired data and the wavelet loss is utilized to facilitate effective reconstruction with multi-frequency details. Extensive experimental results demonstrate that the proposed method is capable of generating realistic 7T MR images with greater anatomical details, outperforming the competing methods trained with the fully paired images. The main contributions are summarized below:

1. We propose a novel semi-supervised cycle generative adversarial network to leverage unpaired data to improve the performance of 7T MR image synthesis when 3T-7T paired data are scarce.
2. We transform 7T MR image synthesis task in the spatial domain to a flexible wavelet coefficient prediction task in the joint spatial-wavelet domain, facilitating effective synthesis of multi-frequency details.
3. We show that our semi-supervised method achieves better performance than competing state-of-the-art methods trained with fully paired data.

2 Method

Our goal is to learn the mapping from 3T MR images in domain X to 7T MR images in domain Y, given a set of training samples $\{x_i\}_{i=1}^{N} \subset X$ and $\{y_i\}_{i=1}^{M} \subset Y$, where $X = X' \cup X''$ and $Y = Y' \cup Y''$. The training samples in X' and Y' are paired 3T-7T MR images, whereas the samples in X'' and Y'' are unpaired. To learn from unpaired data, we couple the primal 3T-to-7T mapping task with a dual 7T-to-3T mapping task. As shown in Fig. 1, SemiWave consists of two mapping functions: a primal 3T-to-7T mapping ($G: x \rightarrow y$) and a dual 7T-to-3T mapping ($F: y \rightarrow x$), two associated adversarial discriminators (D_{3T} and D_{7T}), and one wavelet coefficient extractor (E_w). Discriminator D_{7T} tries to distinguish

between the real 7T MR image y and the synthesized 7T MR image $G(x)$. Similarly, D_{3T} tries to distinguish between x and $F(y)$. The wavelet coefficient extractor E_w computes the wavelet coefficients of four frequency bands. Four losses are involved to regularize these two mappings: adversarial loss, pair-wise reconstruction loss, cycle consistency loss, and wavelet loss.

2.1 Adversarial Loss

Adversarial loss is applied to both mapping functions, aiming to match the distribution of the synthetic images to the distribution of the target images. For generator G and its corresponding discriminator D_{7T}, the adversarial loss is defined as

$$\mathscr{L}_{\text{gan}}(G, D_{7T}) = \mathbb{E}_{y \in Y}[\log D_{7T}(y)] + \mathbb{E}_{x \in X}[1 - \log D_{7T}(G(x))], \tag{1}$$

where G tries to generate a synthetic 7T MR image $G(x)$ that resembles real 7T MR image y in Y, whereas D_{7T} aims to distinguish between a synthetic 7T MR image $G(x)$ and a real image y. G tries to minimize this objective function, whereas D_{7T} aims to maximize it, i.e., $\arg\min_G \max_{D_{7T}} \mathscr{L}_{\text{gan}}(G, D_{7T})$. Similarly, adversarial loss is also applied to generator F and its discriminator D_{3T} as

$$\mathscr{L}_{\text{gan}}(F, D_{3T}) = \mathbb{E}_{x \in X}[\log D_{3T}(x)] + \mathbb{E}_{y \in Y}[1 - \log D_{3T}(F(y))]\,. \tag{2}$$

2.2 Pair-Wise Reconstruction Loss

Networks with adversarial loss are often affected by model collapse and training instability. We incorporate a pair-wise reconstruction loss to impose additional constraints on the generators with a small number of paired images. Specifically, the generator is tasked not only to fool the discriminator but also to minimize the pixel-wise intensity difference between synthetic and real images.

Given a set of paired 3T and 7T MR images $\{X', Y'\}$, the pair-wise reconstruction loss is defined as

$$\mathscr{L}_{\text{p\_s}}(G, F) = \mathbb{E}_{x \in X', y \in Y'}[\|y - G(x)\|_1 + \|x - F(y)\|_1]\,. \tag{3}$$

We choose L1 distance rather than L2 to encourage image sharpness.

2.3 Cycle Consistency Loss

Training an effective 7T MR image synthesis model with adversarial loss and pair-wise reconstruction loss often requires a significant amount of 3T-7T paired data. However, acquiring paired data is often challenging due to factors such as prolonged scanning time, patient discomfort, and costs. Acquisition of unpaired data is relatively straightforward. In light of this, we design a semi-supervised learning mechanism to leverage both the paired and unpaired data to learn the 3T-to-7T mapping. Specifically, to learn from extra unpaired data, a cycle consistency constraint is enforced on G and F. Following [10], the cycle consistency constraint encourages similarity between each synthetic 7T MR image $G(x)$ generated by the 3T-to-7T mapping G with the original 3T image x when mapped back via

the 7T-to-3T mapping F, i.e., $F(G(x)) \approx x$. Similarly, $G(F(y)) \approx y$. The cycle consistency loss is defined as

$$\mathscr{L}_{\text{cyc\_s}}(G, F) = \mathbb{E}_{x \in X}[\|x - F(G(x))\|_1] + E_{y \in Y}[\|y - G(F(y))\|_1]. \quad (4)$$

2.4 Wavelet Loss

The pair-wise reconstruction loss and cycle consistency loss aim at keeping the pixel/voxel-wise intensity consistency between synthetic and real images. These L1/L2 driven pixel-wise losses tend to produce over-smoothed outputs and fail to capture anatomical details [5].

We thus introduce a wavelet loss to encourage the network to generate images with realistic details. Specifically, a wavelet coefficient extractor E_w is first utilized to decompose the image into its wavelet coefficients in four bands. Given a 7T MR image y, we denote its wavelet coefficients in four bands as $\{y_{\text{w}}^{\text{A}}, y_{\text{w}}^{\text{H}}, y_{\text{w}}^{\text{V}}, y_{\text{w}}^{\text{D}}\}$. The approximation coefficients y_{w}^{A} capture the global topology information, and the detail coefficients $\{y_{\text{w}}^{\text{H}}, y_{\text{w}}^{\text{V}}, y_{\text{w}}^{\text{D}}\}$ store the high-frequency details. The wavelet loss $\mathscr{L}_{p\_w}(G)$ between the generated image $G(x)$ and real image y is defined as

$$\begin{aligned}\mathscr{L}_{\text{p\_w}}(G) = \mathbb{E}_{x \in X', y \in Y'}[\lambda_1 \|y_{\text{w}}^{\text{A}} - (G(x))_{\text{w}}^{\text{A}}\|_1 + \lambda_2(\|y_{\text{w}}^{\text{H}} - (G(x))_{\text{w}}^{\text{H}}\|_1 + \|y_{\text{w}}^{\text{V}} \\ - (G(x))_{\text{w}}^{\text{V}}\|_1) + \lambda_3 \|y_{\text{w}}^{\text{D}} - (G(x))_{\text{w}}^{\text{D}}\|_1],\end{aligned} \quad (5)$$

where λ_1, λ_2, and λ_3 are the weights for balancing the contributions of wavelet coefficients of different bands. Emphasis on the reconstruction of high-frequency wavelet coefficients helps recover better anatomical details, whereas emphasis on the prediction of low-frequency wavelet coefficients enforces consistency in global information. During training, we gradually increase the weights λ_2 and λ_3 in order to generate more realistic 7T MR images with greater anatomical details.

Similarly, we also apply wavelet loss $\mathscr{L}_{\text{p\_w}}(F)$ between the synthetic image $F(y)$ and real image x, and the wavelet loss $\mathscr{L}_{\text{cyc\_w}}(G, F)$ between real images $\{x, y\}$ and reconstructed images $\{F(G(x)), G(F(y))\}$.

2.5 Overall Objective Function

The overall objective function is defined as

$$\mathscr{L}(G, F, D_{3\text{T}}, D_{7\text{T}}) = \mathscr{L}_{\text{gan}} + \alpha\mathscr{L}_{\text{p\_s}} + \beta\mathscr{L}_{\text{cyc\_s}} + \gamma\mathscr{L}_{\text{p\_w}} + \delta\mathscr{L}_{\text{cyc\_w}}, \quad (6)$$

where $\mathscr{L}_{\text{p\_s}}$ and $\mathscr{L}_{\text{p\_w}}$ are only valid for paired data $\{X', Y'\}$, and the others are valid for all the training data $\{X, Y\}$. Parameters $\alpha, \beta, \gamma, \delta$ balance the contributions of the different losses.

2.6 Network Architecture

The detailed network architecture of SemiWave is shown in Fig. 1. Similar to [10], our generators G and F contain two stride-2 convolutions, six residual blocks, and two

fractionally-strided convolutions with a stride of 1/2. Instance normalization is applied after each convolutional layers. For discriminators D_{7T} and D_{3T}, a 70 × 70 PatchGans scheme [6, 10] is applied to classify whether 70 × 70 overlapping patches are real or fake. The wavelet coefficient extractor E_w is based on one-level Haar wavelet packet decomposition [1]. It is composed of a stride-2 convolution with its filter weights determined by Haar-based filters.

3 Experiments

3.1 Dataset

We utilized 15 pairs of 3T and 7T T1 weighted MR brain images in our study. The 3T MR images were acquired using a Siemens Magnetom Trio 3T scanner with voxel size 1 × 1 × 1mm^3, TR = 1990ms, and TE = 2.16ms. The 7T MR images were acquired using a Siemens Magnetom 7T whole-body MR scanner, with voxel size 0.65 × 0.65 × 0.65 mm^3, TR = 6000 ms, and TE = 2.95 ms. These images were linearly aligned and skull-stripped to remove non-brain voxels. After skull stripping, the intensity values of each image were linearly scaled to [−1,1].

We adopt leave-one-out cross validation for performance evaluation. One pair of 3T and 7T MR images were used for testing, and the remaining 14 pairs were used for training. Specifically, we divided the training set into two subsets: four pairs for paired training and the remaining ten pairs were randomly shuffled for unpaired training.

3.2 Implementation Details

SemiWave was implemented with Pytorch and optimized with Adam [7]. In our implementation, the negative log likelihood objectives in (1) and (2) were replaced by a more stable least-square loss [10]. In order to stabilize the training, we set the batch size to 2 with one batch from paired data and another from unpaired data. The learning rate was set to 0.0002 for the first 100 epochs and linearly decayed to zero in subsequent 100 epochs. Parameters α, γ were set to 30, and parameters β, δ were set to 10. Parameters λ_2 and λ_3 were initially set to be small and were gradually increased for greater contribution from high-frequency wavelet coefficients. Specifically, we linearly increased parameters λ_1, λ_2, λ_3 from 1, 1, 1 to 1, 3, 5 for the first 50 epochs and then fixed them for the following epochs. In the training phase, we extracted several consecutive axial slices from the 3D images as the training images. Horizontal flipping, randomly scaling, and random rotation were applied to augment the training data.

3.3 Prediction Performance

To validate the effectiveness of the proposed method, we compared our method with four state-of-the-art fully-supervised 7T MR image synthesis methods: MCCA [4], RF [2], CAAF [3] and DDCR [9]. We evaluated the performance of these methods with two commonly accepted metrics: peak signal-noise ratio (PSNR), and structural similarity (SSIM). Table 1 shows the quantitative comparison results of the different 7T MR image synthesis methods. Even with only 28.5% paired data, SemiWave still achieves superior performance compared with the competing fully-supervised methods. Specifically, with the cycle consistency loss and the wavelet loss, SemiWave improves the SSIM value from

0.8580 given by [9] to 0.8740. The visual comparison results in Fig. 2 also demonstrate the ability of SemiWave in generating realistic 7T MR images with greater anatomical details.

3.4 Ablation Study

We investigated the contributions of the three main loses of SemiWave, i.e., pair-wise reconstruction loss ($\mathscr{L}_{p\_s}$ and $\mathscr{L}_{p\_w}$), cycle consistency loss ($\mathscr{L}_{cyc\_s}$ and $\mathscr{L}_{cyc\_w}$), and wavelet loss ($\mathscr{L}_{p\_w}$ and $\mathscr{L}_{cyc\_w}$). We removed each of these losses from SemiWave and train variant models using a strategy similar to SemiWave. Figure 3 and Table 2 compare SemiWave and its variants.

Without pair-wise reconstruction loss, the performance drops dramatically, i.e., the PSNR value drops from 27.85 to 26.72. With the cycle consistency loss, SSIM is improved from 0.8697 to 0.8740. With wavelet loss, realistic images with greater anatomical details can be generated, as depicted in Fig. 3 and Table 2.

4 Conclusion

In this paper, we propose a novel semi-supervised cycle generative adversarial network called SemiWave for 7T MR image synthesis, considering information in both spatial and wavelet domains. Adversarial loss, pair-wise loss, cycle consistency loss, and wavelet loss are incorporated to capture both global and local information. Comprehensive qualitative and quantitative experiments show that our method generates realistic 7T MR images with greater anatomical details than competing fully-supervised methods. Future work will focus on boosting the performance of SemiWave with large-scale unpaired images and extending SemiWave to tackle general cross-modality image synthesis problems.

Acknowledgement.

This work was supported in part by NIH grant EB006733.

References

1. Akansu AN, Haddad PA, Haddad RA: Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets. Academic Press, Orlando (2001)
2. Bahrami K, Shi F, Rekik I, Gao Y, Shen D: 7T-guided super-resolution of 3T MRI. Med. Phys 44(5), 1661–1677 (2017) [PubMed: 28177548]
3. Bahrami K, Shi F, Rekik I, Shen D: Convolutional neural network for reconstruction of 7T-like images from 3T MRI using appearance and anatomical features In: Carneiro G, et al. (eds.) LABELS/DLMIA-2016. LNCS, vol. 10008, pp. 39–47. Springer, Cham (2016). 10.1007/978-3-319-46976-8_5
4. Bahrami K, Shi F, Zong X, Shin HW, An H, Shen D: Hierarchical reconstruction of 7T-like images from 3T MRI using multi-level CCA and group sparsity In: Navab N, Hornegger J, Wells WM, Frangi AF (eds.) MICCAI 2015. LNCS, vol. 9350, pp. 659–666. Springer, Cham (2015). 10.1007/978-3-319-24571-3_79
5. Huang H, He R, Sun Z, Tan T: Wavelet-SRNet: a wavelet-based CNN for multi-scale face super resolution. In: Proceedings IEEE ICCV, pp. 1689–1697 (2017)
6. Isola P, Zhu JY, Zhou T, Efros AA: Image-to-image translation with conditional adversarial networks. In: Proceedings IEEE CVPR, pp. 1125–1134 (2017)

7. Kingma DP, Ba J: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
8. Van der Kolk AG, Hendrikse J, Zwanenburg JJ, Visser F, Luijten PR: Clinical applications of 7T MRI in the brain. Eur. J. Radiol 82(5), 708–718 (2013) [PubMed: 21937178]
9. Zhang Y, Cheng J-Z, Xiang L, Yap P-T, Shen D: Dual-domain cascaded regression for synthesizing 7T from 3T MRI In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-Lopez C, Fichtinger G (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 410–417. Springer, Cham (2018). 10.1007/978-3-030-00928-1_47
10. Zhu JY, Park T, Isola P, Efros AA: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings IEEE CVPR, pp. 2223–2232 (2017)

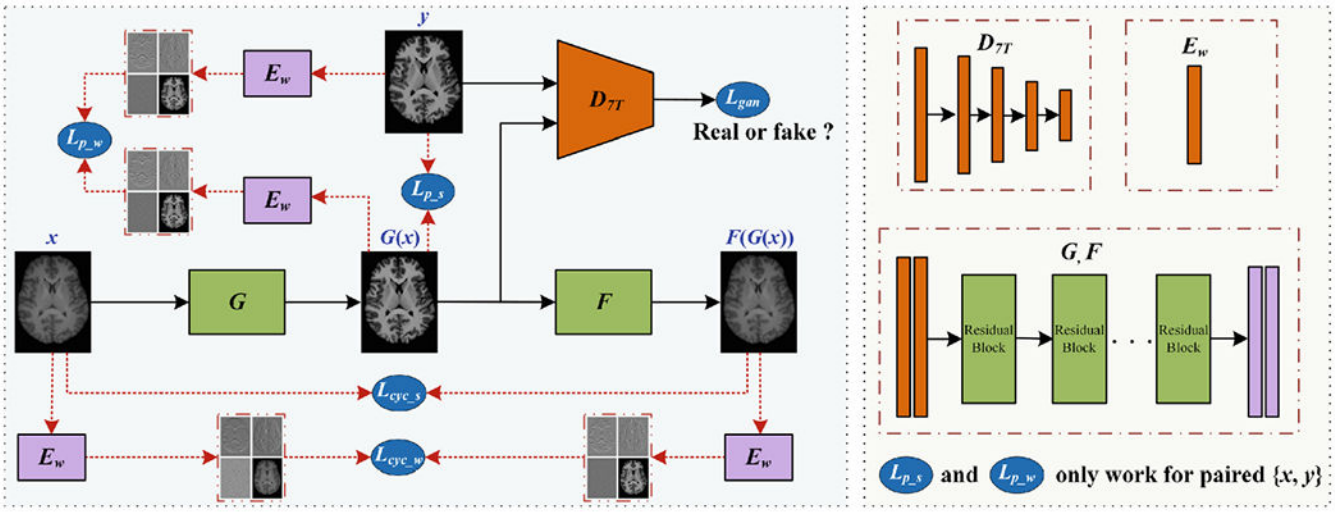

Fig. 1.

Architecture of SemiWave. We show only $x \rightarrow G(x) \rightarrow F(G(x))$ for simplicity. SemiWave consists of two mapping functions ($G: x \rightarrow y$ and $F: y \rightarrow x$), an adversarial discriminator D_{7T}, and a wavelet coefficient extractor E_w. Four kinds of losses are used to regularize these two mapping functions: Adversarial loss $\mathscr{L}_{gan}$, pair-wise reconstruction loss $\mathscr{L}_{p\_s}$, cycle consistency loss $\mathscr{L}_{cyc\_s}$, and wavelet losses $\mathscr{L}_{p\_w}$ and $\mathscr{L}_{cyc\_w}$. $\mathscr{L}_{p\_s}$ and $\mathscr{L}_{p\_w}$ are for paired data and $\mathscr{L}_{gan}$, $\mathscr{L}_{cyc\_s}$ and $\mathscr{L}_{cyc\_w}$ are for both paired and unpaired data.

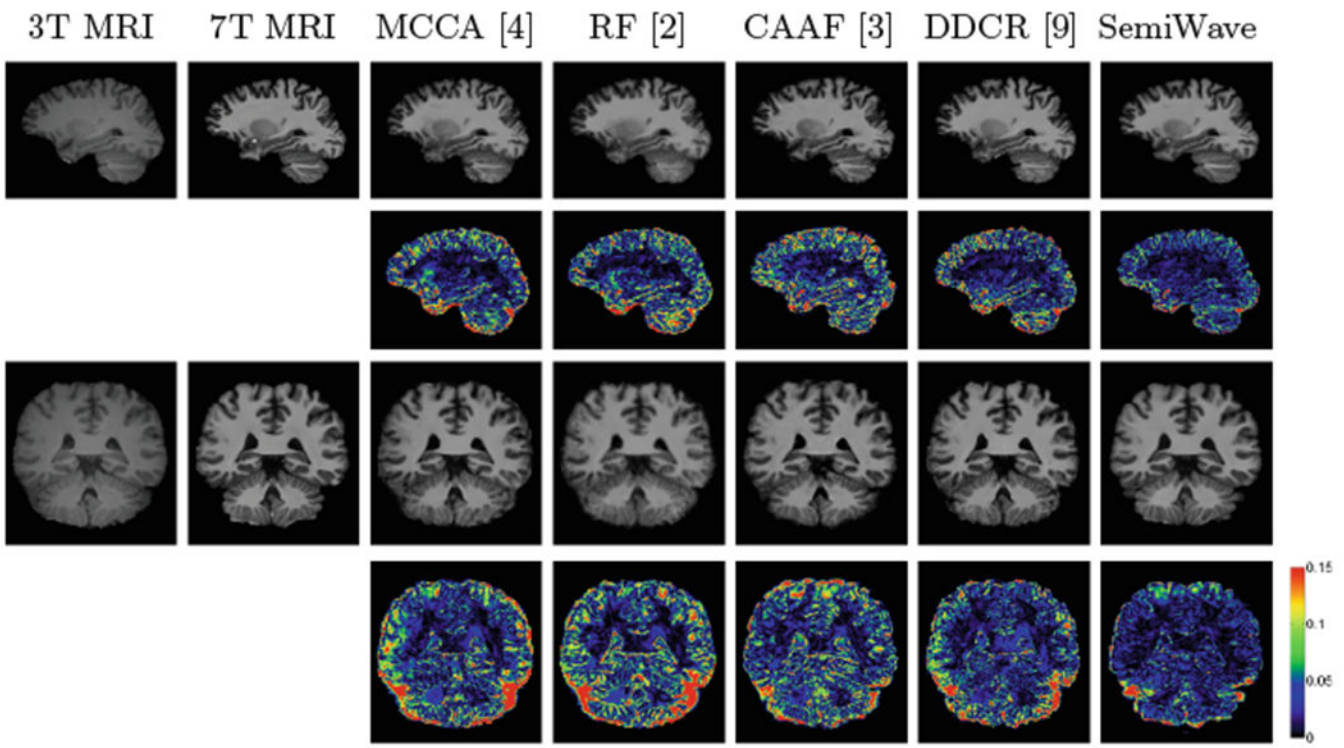

Fig. 2.

7T MR images synthesized using SemiWave and four other methods shown in sagittal and coronal views, along with the prediction error maps.

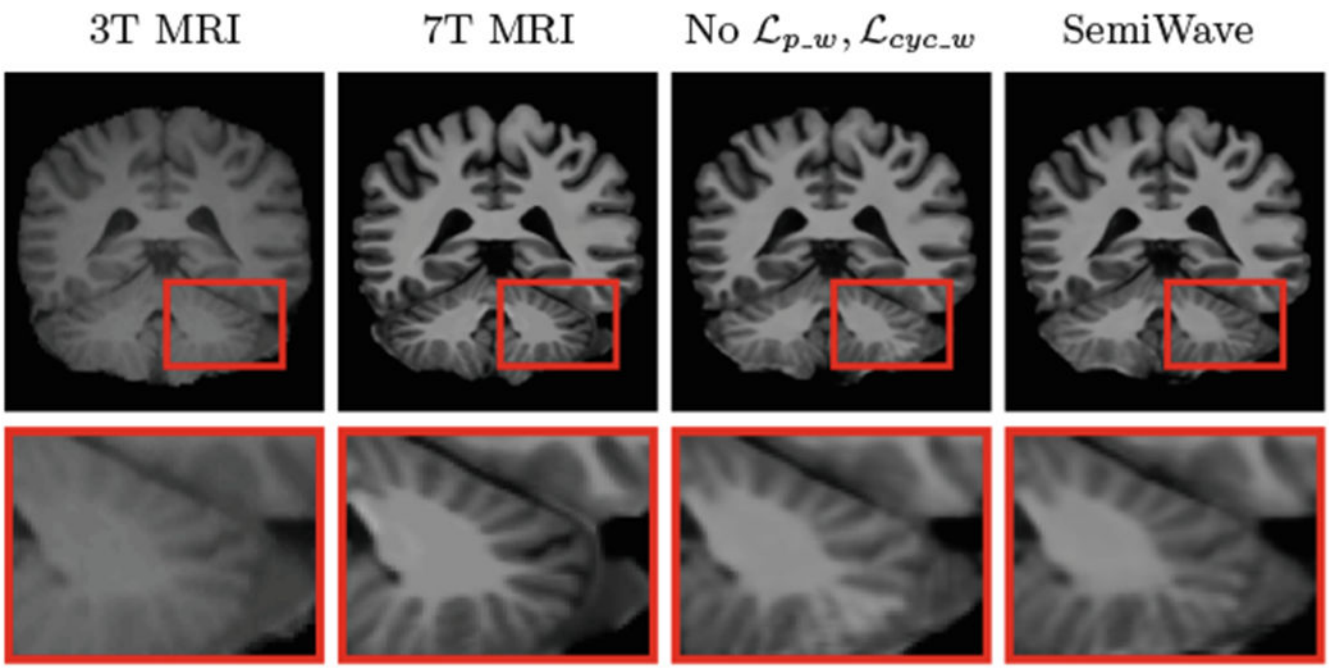

Fig. 3.
Ablation study: Effectiveness of wavelet loss.

Table 1.

Mean PSNR and SSIM for different 7T MR image synthesis methods. All the competing methods were trained with the fully paired data. SemiWave was trained with around 28.5% paired data and 71.5% unpaired data.

	MCCA [4]	RF [2]	CFFA [3]	DDCR [9]	SemiWave
PSNR	25.52	26.55	27.05	27.51	**27.85**
SSIM	0.4840	0.5728	0.8406	0.8580	**0.8740**

Table 2.

Ablation study: Effectiveness of different loses in SemiWave.

	Without $\mathscr{L}_{p\_s}$, $\mathscr{L}_{p\_w}$	**Without $\mathscr{L}_{cyc\_s}$, $\mathscr{L}_{cyc\_w}$**	**Without $\mathscr{L}_{p\_w}$, $\mathscr{L}_{cyc\_w}$**	**SemiWave**
PSNR	26.72	27.42	27.60	**27.85**
SSIM	0.8417	0.8697	0.8702	**0.8740**