Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data

Elisa Moya-Sáez*, Óscar Peña-Nogales, Rodrigo de Luis-García, Carlos Alberola-López

Laboratorio de Procesado de Imagen, Universidad de Valladolid, Valladolid, Spain

ARTICLE INFO

ABSTRACT

Background and Objective: Synthetic magnetic resonance imaging (MRI) is a low cost procedure that serves as a bridge between qualitative and quantitative MRI. However, the proposed methods require very specific sequences or private protocols which have scarcely found integration in clinical scanners. We propose a learning-based approach to compute T1, T2, and PD parametric maps from only a pair of T1- and T2-weighted images customarily acquired in the clinical routine.

Methods: Our approach is based on a convolutional neural network (CNN) trained with synthetic data; specifically, a synthetic dataset with 120 volumes was constructed from the anatomical brain model of the BrainWeb tool and served as the training set. The CNN learns an end-to-end mapping function to transform the input T1- and T2-weighted images to their underlying T1, T2, and PD parametric maps. Then, conventional weighted images unseen by the network are analytically synthesized from the parametric maps. The network can be fine tuned with a small database of actual weighted images and maps for better performance.

Results: This approach is able to accurately compute parametric maps from synthetic brain data achieving normalized squared error values predominantly below 1%. It also yields realistic parametric maps from actual MR brain acquisitions with T1, T2, and PD values in the range of the literature and with correlation values above 0.95 compared to the T1 and T2 maps obtained from relaxometry sequences. Further, the synthesized weighted images are visually realistic; the mean square error values are always below 9% and the structural similarity index is usually above 0.90. Network fine tuning with actual maps improves performance, while training exclusively with a small database of actual maps shows a performance degradation.

Conclusions: These results show that our approach is able to provide realistic parametric maps and weighted images out of a CNN that (a) is trained with a synthetic dataset and (b) needs only two inputs, which are in turn obtained from a common full-brain acquisition that takes less than 8 min of scan time. Although a fine tuning with actual maps improves performance, synthetic data is crucial to reach acceptable performance levels. Hence, we show the utility of our approach for both quantitative MRI in clinical viable times and for the synthesis of additional weighted images to those actually acquired.

1. Introduction

Magnetic resonance imaging (MRI) is widely used in the clinical routine due to its capability of providing meaningful anatomical and functional information. Images are obtained by applying a pulse sequence with a selection of parameters; their combination gives rise to a specific weighting of the tissue magnetic properties, which, in turn, produce an image with an associated contrast. Hereinafter, each of the different contrasts is referred as an image modality. An MRI scan protocol consists of a number of sequences that provide different image modalities, which are intended to provide complementary information for diagnosis [1].

These so-called weighted images are of a qualitative nature and are routinely used by radiologists for diagnosis. Quantitative MRI, on the other side, consists in finding the tissue magnetic properties themselves (the so-called T1, T2, and PD maps, which are jointly

---

* Corresponding author.
E-mail address: emoysae@lpi.tel.uva.es (E. Moya-Sáez).
URL: http://www.lpi.tel.uva.es (E. Moya-Sáez)

referred to as parametric maps). These parametric maps are more robust than weighted images due to their lower sensitivity to MRI hardware and sequence configurations [2]. In addition, the dependence of these maps with biophysical tissue properties plays an important role in tissue characterization in healthy and diseased stages, for pathologies such as epilepsy [3] or multiple sclerosis [4], and for tumor detection [5]. Although, as indicated above, radiologists are not used to performing diagnosis solely on the basis of T1, T2, and PD parametric maps [6], quantitative MRI is gaining popularity [7].

Synthetic MRI has recently entered the field and it somehow serves as a bridge between qualitative and quantitative MRI [8]. It has been conventionally defined as a three-step procedure [9] where 1) a set of weighted images are acquired, 2) quantitative parametric maps are computed from them and 3) any weighted image can be synthesized using the equations that describe MR intensity as a function of sequence parameters and parametric maps. Thus, this low-cost procedure could retrospectively enhance patient throughput and facilitate radiologists routine.

In this paper we propose a novel method for synthetic MRI in which step 2 of the procedure described above is based on deep learning. Its main novelty stems from the fact that the proposed training is performed mainly on the basis of synthetic data. The inputs to our method are two conventional sequences customarily used in any brain acquisition protocol (step 1). A number of weighted images unseen by the network can be then satisfactorily obtained out of the parametric maps we provide (step 3). Thus, the proposed method is not limited to a number of predefined weighted images that have entered the learning process, but is able to generalize to any image modality.

2. Related work

2.1. Quantitative MRI: classical relaxometry techniques

In the classical methods to obtain parametric maps —steps 1) and 2) of synthetic MRI— a set of weighted images with different sequence parameters is customarily acquired; from these images, the map is computed by voxelwise fitting a known relaxation model. These techniques are termed as relaxometry. In order to perform this fitting, different estimation-based procedures have been proposed [10–12]. Also, approaches that perform the fitting by deep learning have been recently developed [13]. However, the acquisitions needed to obtain a sufficiently large set of weighted images are time consuming, a fact that limits their utility in clinical routine. In addition, most of the proposed methods only provide information of a single parameter at a time [12–14], so additional sequences are needed to obtain the three parametric maps.

To overcome these limitations, MR fingerprinting (MRF) [15] is able to estimate parametric maps within a short scan time. However, the specific spiral acquisition needed has rarely found integration into clinical scan protocols [16].

2.2. Medical image translation

Various deep learning approaches have been proposed that learn the mapping between different pairs of images [17–19], a procedure that we will refer to as medical image translation. These approaches share the same objective as step 3) of synthetic MRI, although their input is a set of weighted images as opposed to parametric maps. Hence, they are not flexible as to which modalities can be generated, since most of them are tailored for a specific application where, given some input image modalities, new predefined image modalities are synthesized. For example, in [18] the authors synthesize T1-weighted images from T2-weighted images (hereinafter, T1w and T2w).

These methodologies limit themselves to the image modalities used in the learning stage since the potential of parametric maps to synthesize any weighted image is not employed. This design choice may be the consequence of the nonexistence of large datasets that contain both weighted images and the corresponding parametric maps due to their long acquisition times. At first glance, this would be mandatory since deep learning model training requires extensive datasets [20].

2.3. Synthetic MRI methods

Different methods that fit the conventional three-step definition of synthetic MRI [9] have been proposed. Gulani et al. [21] proposed an steady-state precession (IR-TrueFISP) sequence in which a series of different IR time-delayed TrueFISP images are acquired to quantify the parametric maps. Then T1w, T2w, PD-weighted (PDw), and T2w fluid attenuated inversion recovery (T2w FLAIR) images were synthesized from these maps. Warntjes et al. [22] proposed a multiecho acquisition of a saturation-recovery turbo spin-echo readout (QRAPMASTER)[1] for the quantification of T1, T2, PD, and B1 inhomogeneity parametric maps. After quantification, T1w, T2w, and T2w FLAIR images were synthesized [24]. Finally, Cheng et al. [25] suggested a multipathway multiecho (MPME) sequence using an unbalanced steady-state sequence with two different flip angles and resolution scans to quantify T1, T2, T2*, B1 inhomogeneity, and B0 inhomogeneity parametric maps. Then, the authors showed the synthesis of T1w, T2w, PDw, T2w FLAIR, and magnetization prepared rapid gradient echo (MPRAGE) images using a neural network. Nevertheless, this latter method suffers from noise amplification due to the multiple processing steps which leads to somewhat noisy maps and synthesized images.

It is important to note, however, that all of these methods require very specific sequences or private protocols scarcely available in clinical scanners. Also, these quantitative sequences are focused on obtaining the parametric maps, but they are not valuable by themselves for diagnosis purposes in the clinical routine.

2.4. Our contributions

In this work we propose a joint synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images from two common clinical routine sequences, through a convolutional neural network (CNN).

Our main contributions are summarized as follows:

- A novel three-from-two approach for the computation of T1, T2, and PD parametric maps from only a pair of weighted images, namely a T1w and a T2w. The mapping between the weighted images and the parametric maps is carried out by means of a CNN. The T1w and T2w input sequences are customarily used in the clinical routine.
- A new training strategy based on a synthetic dataset generated from the BrainWeb anatomical brain model is proposed. This way, we overcome the need of large datasets with quantitative parametric maps.
- Realistic parametric maps with values in the range of the literature for 3T scanners are computed from actual MR brain acquisitions. These achieve high correlation values compared to parametric maps obtained from relaxometry sequences, hereinafter referred to as *silver standard*.
- The synthesis of multiple realistic weighted images from these computed parametric maps both for modalities previously seen

---

[1] QRAPMASTER is nowadays referred to as a multidynamic multiecho (MDME) sequence [23].

Synthetic dataset generation used for training and testing of the CNN

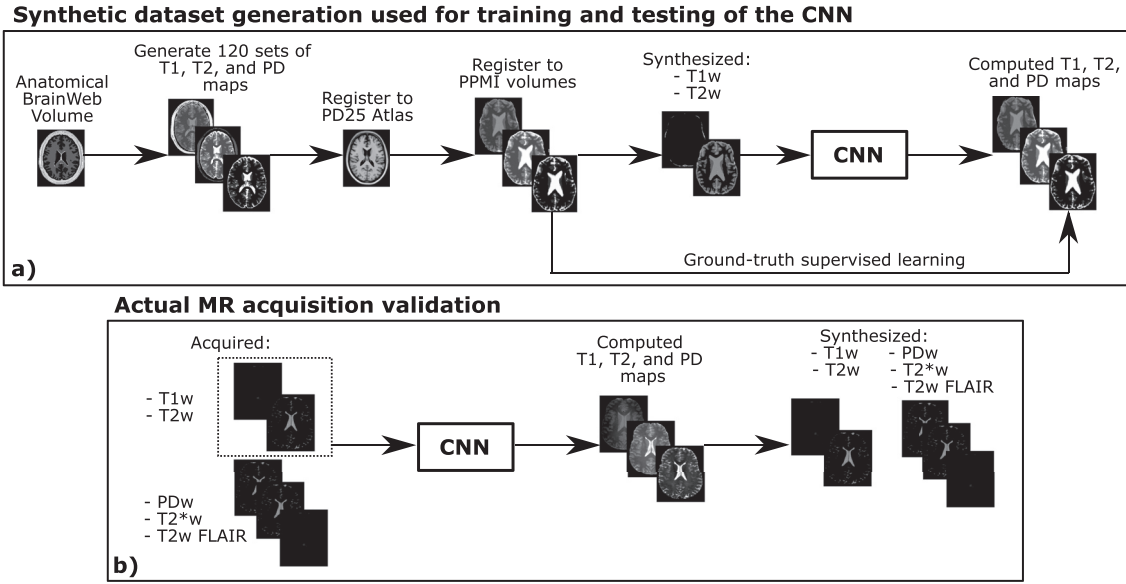**Actual MR acquisition validation**

Fig. 1. Pipeline of the proposed training and validation approaches. a) Synthetic dataset generation used for training and testing of the supervised convolutional neural network (CNN). b) Validation of the CNN with actual MR brain acquisitions.

by the network and for other modalities not used in the learning stage. Quantitative and qualitative comparisons between the synthesized and the acquired weighted images are provided.

- The performance of our approach to synthesize additional image modalities with different sequence parameters (i.e., TE, TR, TI) is compared with other state-of-the-art methods.

3. Methods

In this manuscript we propose a joint synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images from only a pair of inputs, namely a T1w and a T2w. To this end, we adapt the CNN proposed by Chartsias et al. [17] and train it with a synthetic dataset where the T1w and the T2w images were synthesized from their corresponding parametric maps. The proposed dataset generation, neural network training, and validation procedures are shown in Fig. 1.

3.1. Sequences

Any weighted MR image can be analytically synthesized if the parametric maps needed to feed the pulse sequence are known. For simple cases, this synthesis can be performed using the well-known equations that describe MR intensity as a function of sequence parameters, such as echo time (TE), repetition time (TR), inversion time (TI) and flip angle (α), in relation to the involved parametric maps. In more complicated cases, more sophisticated methods are needed [26]. In this manuscript, based on the T1, T2, and PD maps, we synthesise weighted images corresponding to the sequences magnetization prepared rapid gradient echo (MPRAGE), spin echo (SE), gradient echo (GRE), and inversion recovery spin echo (IR-SE), with respective Eqs. (1)–(4):

$$m_{\text{MPRAGE}}(\mathbf{x}) = PD(\mathbf{x}) \frac{1 - 2e^{-T_I/T_1(\mathbf{x})} + e^{-T_R/T_1(\mathbf{x})}}{1 + \cos(\alpha)e^{-T_R/T_1(\mathbf{x})}} \sin(\alpha)e^{-T_E/T_2(\mathbf{x})} \quad (1)$$

$$m_{\text{SE}}(\mathbf{x}) = PD(\mathbf{x})\left[1 - 2e^{-(T_R - T_E/2)/T_1(\mathbf{x})} + e^{-T_R/T_1(\mathbf{x})}\right]e^{-T_E/T_2(\mathbf{x})} \quad (2)$$

$$m_{\text{GRE}}(\mathbf{x}) = PD(\mathbf{x}) \frac{1 - e^{-T_R/T_1(\mathbf{x})}}{1 - \cos(\alpha)e^{-T_R/T_1(\mathbf{x})}} \sin(\alpha)e^{-T_E/T_2(\mathbf{x})} \quad (3)$$

Table 1
Ranges of the T1, T2, and PD values for each of the three considered tissues in the anatomical brain model (WM: white matter, GM: grey matter, and CSF: cerebrospinal fluid) for the generation of 120 synthetic brain volumes. The specific value for each parameter in each tissue and volume is selected from a uniform distribution within these ranges. T1 and T2 values are given in seconds.

	T1 (s)	T2 (s)	PD
WM	0.80–1.10	0.055–0.075	0.65–0.72
GM	1.40–1.60	0.075–0.120	0.77–0.82
CSF	4.50–4.80	1.200–1.600	1.20–1.30

$$m_{\text{IR-SE}}(\mathbf{x})$$
$$= PD(\mathbf{x})\left[1 - 2e^{-T_I/T_1(\mathbf{x})} + 2e^{-(T_R - T_E/2)/T_1(\mathbf{x})} - e^{-T_R/T_1(\mathbf{x})}\right]e^{-T_E/T_2(\mathbf{x})}$$
$$(4)$$

with $\mathbf{x}$ the voxel location defined on some domain $\mathcal{X}$. Equations (1) and (2) are employed for the generation of the synthetic dataset used for training and testing the CNN. On the other hand, the four of them are employed for the validation of the proposed approach with actual MR brain acquisitions.

3.2. Proposed approach

3.2.1. Synthetic dataset generation

We create a synthetic dataset with 120 brain volumes starting from the anatomical model of a normal brain obtained with Brain-Web [27]. The pipeline to create the synthetic dataset is described in the next four steps (see Fig. 1a).

First, we created 120 different sets of T1, T2, and PD maps from the BrainWeb anatomical model by giving uniformly distributed random values to the white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF) of each map, one value for each label. The exact ranges of T1, T2, and PD values defined for each parameter in each tissue are shown in Table 1. Note that these ranges are within those reported in the literature for 3T MR scanners [28]. Also, additive Gaussian noise was added to each volume with distribution $\mathcal{N}(\mu = 0, \sigma = 0.01)$.

Table 2

Parameters of the acquisitions in *Brain MRI* and *Quantitative Brain MRI* protocols. The parameters are the echo time (TE), the repetition time (TR), the inversion time (TI), the flip angle (alpha), the echo train length (ETL), the SENSE acceleration factor (SENSE), the number of signal averages (NSA), the number of slices (and its orientation), the slice thickness, the in-plane resolution, the field-of-view (FOV), and the approximate acquisition scan time.

	Brain MRI				Quantitative Brain MRI		
	T1w(MPRAGE)	T2*w(GRE)	T2w FLAIR(IR-TSE)	PDw/T2w(TSE)	T1(VFA GRE)	T2(multiecho TSE)	PD(GRE)
TE (ms)	3	20	100	30/85	2	17, 46, 75, 104, 133, 162	2
TR (ms)	6.44	746.99	11000	4000	18	1000	50
TI (ms)	900	-	2800	-	-	-	-
α (°)	10	20	-	-	2, 3, 4, 5, 7, 9, 11, 14, 17, 19, 22	-	5
ETL	192	1	12	10	1	30	1
SENSE	1.8/1/1.2	2	2	2	2	2	2
NSA	1	1	1	1	1	1	1
# Slices	170	27	27	50	150	150	150
	(sagittal)	(axial)	(axial)	(axial)	(axial)	(axial)	(axial)
Thickness (mm)	1.2	5	5	3	1.5	1.5	1.5
Resolution (mm^2)	1.25 × 1.25	0.94 × 1.25	0.94 × 1.25	1.02 × 1.36	1.50 × 1.50	1.50 × 1.50	1.50 × 1.50
FOV (mm)	240 × 240	240 × 240	240 × 240	260 × 195	240 × 240	240 × 240	240 × 240
Scan time (min)	∼ 4:00	2:30 - 4:00	2:30 - 4:30	2:30 - 4:30	∼ 17:00	∼ 18:00	4:00 - 4:30

Second, in order to introduce spatial variability across brain volumes (i.e., brains with different anatomical features), the maps were affine-registered to the PD25 atlas [29] using the FLIRT tool of FSL (Oxford, UK) [30]. Then, each set of maps was non-linearly registered to one out of the 120 different T1w volumes selected from the PPMI database (www.ppmi-info.org), with the FNIRT tool of FSL as described in [31]. The size of each of these maps is of 240 × 176 with 256 slices. Subsequently, all sets of parametric maps were skull-stripped.

Third, for each set of skull-stripped T1, T2, and PD maps, a pair of weighted images was analytically synthesized. A T1w image was synthesized as a MPRAGE acquisition (Eq. (1)) with TE = 3 ms, TR = 6.44 ms, TI = 900 ms, and α = 10°. A T2w image was synthesized as a SE acquisition (Eq. (2)) with TE = 85 ms and TR = 4000 ms. These particular sequences and parameter sets were chosen to match the actual MR brain acquisitions described in Table 2. Note that the weighted images have the same dimensions as the parametric maps (i.e., 240 × 176 with 256 slices).

Finally, we normalized the T1w and T2w images by dividing each of them by its average intensity without considering the background. This facilitates convergence of the CNN during training without altering image properties due to their qualitative nature.

3.2.2. The network

Network training with the synthetic dataset. The aforementioned synthetic dataset was used to train an adapted version of the CNN described in [17]; our adaptation pursued to perform an end-to-end mapping function to transform the input T1w and T2w images to their corresponding set of T1, T2, and PD parametric maps (see Fig. 1a). Specifically, the weighted images were input to two encoders which embed these inputs into multi-channel latent spaces with the same image size as the inputs. Note that the CNN processes the inputs as 2D slices. The number of channels used is 16. Then, the latent representations of the input are fused into a single 16-channel representation using a maximum pixelwise function between each pair of corresponding channels. This fused latent representation is next input to three decoders to obtain the three desired parametric maps. Supervised training was carried out using the cost function proposed in [17]. This cost function minimizes 1) the mean absolute error (MAE) between the ground-truth parametric maps and the output's decoders (i.e. the synthesized parametric maps), and 2) the mean pixelwise variance between latent representations. The model was trained through a mini-batch approach with a batch size of 8 images using Adam optimizer [32] with a learning rate of 1×10^{-5}. We performed the training with early stopping to avoid overfitting. From the 120 brain volumes of the synthetic dataset with a set of three parametric maps and two weighted images each, we used 70 for training (17920 slices), 36 for the early-stopping monitoring (9216 slices), and 14 for test (3584 slices).

The adapted CNN is coded in Python with Keras. We run the code using the TensorFlow backend on a single NVidia GeForce GTX 1070. The total learning took about 10 h of computation time. Note that once the CNN has been trained, the network computation time reduces to a few seconds.

Network testing with the synthetic dataset. We evaluated the proper parametric mapping of the network through the 14 brain volumes of the synthetic dataset remaining for testing. In addition to visual evaluation, we carried out a quantitative analysis in the parametric maps domain due to the existence of the corresponding ground-truth. The comparison between the computed and the ground-truth T1, T2, and PD parametric maps was performed with the normalized squared error (NSE) map computed as

$$\text{NSE}(\mathbf{x}) = \frac{(\text{MAP}_c(\mathbf{x}) - \text{MAP}_{GT}(\mathbf{x}))^2}{\text{MAP}_{GT}^2(\mathbf{x})} \times 100\%, \tag{5}$$

where MAP is one of the T1, T2, or PD maps, c stands for *computed* and GT for *ground-truth*. Similarly, they are also compared with the scalar metrics described later in Section 3.3.

3.2.3. Validation with actual MR acquisitions

MRI acquisitions. The real data used in these experiments have the approval of the institutional review board (IRB) and the eight subjects involved —suspected of early Alzheimer disease— signed the informed written consent; the *Brain MRI* protocol that follows was acquired with a 32-channel head coil on a 3T scanner (Achieva, Philips, Best, The Netherlands). Local B0 and B1 shimming were used in order to correct for field inhomogeneities. Each study was composed of four sequences and a total of five image modalities. These sequences are: 1) 3D MPRAGE sequence to obtain T1w images; 2) GRE sequence to obtain T2*w images; 3) IR-TSE sequence to obtain T2w FLAIR images; 4) turbo spin echo (TSE) multi echo sequence to obtain PDw and T2w images.

Additionally, for validation purposes in the same MRI unit and with the same corrections, we also scanned five healthy volunteers with IRB approval and informed written consent. For these volunteers, we performed the *Quantitative Brain MRI* protocol with previous 1) and 4) sequences to input the CNN, sequence 3) only for registration purposes, and the following set of relaxometry sequences: 5) The T1 map was retrieved using a variable flip angle (VFA) of a 3D GRE sequence with 11 different flip angles. Then, NOVIFAST algorithm [12] was employed for the T1 estimation; 6) The T2 map was measured using a 3D multi echo sequence with six different TEs and a least squares estimation procedure; 7) The

PD map was obtained using a 3D GRE. We name the maps that result from 5), 6) and 7) silver standard, since they are affected by common artifacts as well as by physiological motion due to the length of the sequences. The parameters of each of previous sequences are given in Table 2.

Data preprocessing. We preprocessed the actual MR brain volumes in order to register all the image modalities to the same image space and to adapt them to the network input layer. All image modalities were affine-registered to the T2w FLAIR using FLIRT (Oxford, UK) [30]. After registration, the size of each image modality is of 256×256 with 27 slices with voxel size of $0.94 \times 1.25 \times 5$ mm as shown in Table 2. Note that this registration step is only necessary for training and validation purposes, because in production mode —once the network is fully trained— the only requirement is to have the input images with spatial alignment. All images were then skull-stripped. Subsequently, all images were cropped to 240×176 pixels which is the dimension of the network's input layer. We normalized the weighted images by dividing each of them by its average intensity without considering the background. This normalization was done in accordance with the preprocessing steps of the CNN training data. In addition, the relaxometry PD maps were normalized so that their 99th percentiles matched the maximum of the PD map from the synthetic training dataset. Finally, the 14 central slices of each actual MR brain volume were then selected to avoid slices with predominant background areas and/or very prone to artifacts.

Validation. We validated the performance of the proposed approach to compute parametric maps and to synthesize different weighted image modalities when actual T1w and T2w images are input to the network following the pipeline in Fig. 1b). Synthesis quality has been assessed both on the maps directly provided by the network output as well as on the synthesized weighted images. Quality parameters have been defined both at region of interest (ROI) level and at whole image level. Precise definitions for these parameters are provided in the next section.

3.3. Quantitative parameters for quality assessment

We drew nine circular ROIs in each subject of the *Brain MRI* protocol co-localized across the different parametric maps and weighted images enumerated in Fig. 1b). From the nine ROIs, three were located in the CSF (approximately 3 mm of radius), three in the white matter (approximately 3 mm of radius), and three in the grey matter (approximately 2 mm of radius). Let $\mathcal{X}_i^k(n)$ denote the set of voxels[2] belonging to ROI i, $1 \le i \le 3$ from tissue k, $1 \le k \le 3$ (say, 1 for CSF, 2 for GM, and 3 for WM) and subject n, $1 \le n \le 8$. As for the parametric maps provided by the network from *Brain MRI*, we define the following two parameters

$$\mu^k = \frac{1}{\sum_{n=1}^{8}\sum_{i=1}^{3}|\mathcal{X}_i^k(n)|}\sum_{n=1}^{8}\sum_{i=1}^{3}\sum_{\mathbf{x}\in\mathcal{X}_i^k(n)}\mathrm{MAP}_c^n(\mathbf{x}) \tag{6}$$

$$s^k = \sqrt{\frac{1}{\sum_{n=1}^{8}\sum_{i=1}^{3}|\mathcal{X}_i^k(n)|}\sum_{n=1}^{8}\sum_{i=1}^{3}\sum_{\mathbf{x}\in\mathcal{X}_i^k(n)}\left(\mathrm{MAP}_c^n(\mathbf{x})-\mu^k\right)^2} \tag{7}$$

with $\mathrm{MAP}_c^n(\mathbf{x})$ a computed parametric map evaluated at point $\mathbf{x}$ and $|\cdot|$ denotes the cardinality of a set.

For the particular case of the subjects that underwent the *Quantitative Brain MRI* protocol, we drew 12 circular ROIs in each tissue co-localized across the different parametric maps and

---

[2] ROIs have been delineated in 2D, so the third component $\forall \mathbf{x} \in \mathcal{X}_i^k(n)$ coincides.

weighted images. For its parametric maps we define:

$$\mu_i^k(n)_L = \frac{1}{|\mathcal{X}_i^k(n)|}\sum_{\mathbf{x}\in\mathcal{X}_i^k(n)}\mathrm{MAP}_L^n(\mathbf{x}) \tag{8}$$

with $\mathrm{MAP}_L^n(\mathbf{x})$ a parametric map of the nth healthy subject evaluated at point $\mathbf{x}$ and L is a label that takes the values c for the MAP computed by the network and Silver for the silver standard relaxometry maps; k follows the same convention as in Eqs. (6) and (7), $1 \le i \le 12$, and $1 \le n \le 5$.

As for the weighted images of *Brain MRI*, we define:

$$\mu_i^k(n) = \frac{1}{|\mathcal{X}_i^k(n)|}\sum_{\mathbf{x}\in\mathcal{X}_i^k(n)}m^n(\mathbf{x}) \tag{9}$$

$$s_i^k(n) = \sqrt{\frac{1}{|\mathcal{X}_i^k(n)|}\sum_{\mathbf{x}\in\mathcal{X}_i^k(n)}\left(m^n(\mathbf{x})-\mu_i^k(n)\right)^2} \tag{10}$$

$$s(n) = \frac{1}{9}\sum_{i=1}^{3}\sum_{k=1}^{3}s_i^k(n) \tag{11}$$

with $m^n(\mathbf{x})$ an image (either computed or acquired) of the nth subject evaluated at point $\mathbf{x}$. Then, the following samples (per tissue k, $1 \le k \le 3$) are created:

1. Intensity values $\mu_i^k(n)$, $1 \le i \le 3$, $1 \le n \le 8$.
2. Contrast:

$$c_{ij}^k(n) = \frac{\mu_i^k(n) - \mu_j^k(n)}{\mu_i^k(n) + \mu_j^k(n)}, \tag{12}$$

$1 \le i, j \le 3, i \ne j, 1 \le n \le 8$.
3. Contrast-to-noise ratio (CNR):

$$\mathrm{CNR}_{ij}^k(n) = \frac{\mu_i^k(n) - \mu_j^k(n)}{s(n)}, \tag{13}$$

$1 \le i, j \le 3, i \ne j, 1 \le n \le 8$.
4. Signal-to-noise ratio (SNR):

$$\mathrm{SNR}_i^k(n) = \frac{\mu_i^k(n)}{s(n)}, \tag{14}$$

$1 \le i \le 3, \ 1 \le n \le 8$.

In addition, in each subject of *Brain MRI* we also drew a rectangular ROI measuring approximately 70.50 mm $\times$ 33.75 mm, which was chosen to encompass the occipital region of the brain. The number of pixels of this rectangular ROI was of 2025.

At a whole image level, we used four well-known metrics commonly used in medical image translation methods. These metrics are the mean squared error (MSE), the structural similarity index (SSIM), the peak signal-to-noise ratio (PSNR), and the correlation coefficient (CORR) defined as follows:

$$\overline{m^n} = \frac{1}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}m^n(\mathbf{x})$$

$$c_{m_1^n m_2^n} = \frac{1}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}\left(m_1^n(\mathbf{x})-\overline{m_1^n}\right)\left(m_2^n(\mathbf{x})-\overline{m_2^n}\right)$$

$$\mathrm{MSE}(n) = \frac{1}{|\mathcal{X}|}\sum_{\mathbf{x}\in\mathcal{X}}\left(m_c^n(\mathbf{x})-m_s^n(\mathbf{x})\right)^2 \tag{15}$$

$$\mathrm{PSNR}(n) = 10\log_{10}\left(\frac{\max_{\mathbf{x}\in\mathcal{X}}\left(m_c^n(\mathbf{x})\right)^2}{\mathrm{MSE}(n)}\right) \tag{16}$$

$$\mathrm{CORR}(n) = \frac{c_{m_c^n m_s^n}}{\sqrt{c_{m_c^n m_c^n}c_{m_s^n m_s^n}}} \tag{17}$$

$$\text{SSIM}(n) = \frac{(2\overline{m_c^n}\ \overline{m_s^n} + C_1)(2c_{m_c^n m_s^n} + C_2)}{\left(\left(\overline{m_c^n}\right)^2 + \left(\overline{m_s^n}\right)^2 + C_1\right)\left(c_{m_c^n m_c^n} + c_{m_s^n m_s^n} + C_2\right)} \qquad (18)$$

with $m_c^n(\mathbf{x})$ and $m_s^n(\mathbf{x})$ the computed and acquired images, respectively, for the nth subject of *Brain MRI*, $1 \leq n \leq 8$, evaluated at point $\mathbf{x}$; voxels take on values within domain $\mathcal{X}$. Unless otherwise stated, this domain will consist of the brain area. These four metrics have also been used with parametric maps for performance assessment on synthetic data.

3.4. Experiments

Network verification with synthetic images as inputs has been accomplished by visual assessment as well as with the NSE map defined in Eq. (5). In addition, the parameters defined in Eqs. (15)–(18) have also been employed.

As for the network validation with real images, all the parameters defined in the previous section have been employed, and assessment has been carried out both directly on the network outputs (i.e., on the parametric maps) as well as on the synthesized weighted images. For the former, we have employed the silver standard maps from the five subjects involved in the *Quantitative Brain MRI* protocol. For the latter, and as indicated in Fig. 1b), we analytically synthesized the same weighted images acquired in the *Brain MRI* protocol with the same sequence parameters as those described in Table 2. The equations used for each sequence are Eqs. (1)–(4) for the synthesis of T1w, PDw/T2w, T2*w, and T2w FLAIR, respectively. In addition, we synthesized additional weighted images with the same sequences as in the *Brain MRI* protocol, but varying the sequence parameters (i.e. TE, TR, TI). These sequences are SE (Eq. (2)) with TE in the range of 20 to 100 ms and TR of 120 and 4000 ms, and IR-SE (Eq. (4)) with three different combinations of TE, TR, and TI. We do not have the corresponding acquired weighted images as ground-truth due to scan time restrictions, but we pursue to investigate the versatility of our approach to synthesize any weighted image with coherent contrast.

We have also tested how the network deals with non skull-stripped images, a fact that is indicated to be an issue in [17]. To this end, non skull-stripped T1w and T2w images were input to the CNN. In this case, normalization was done by dividing each of them by the skull-stripped images average intensity in accordance with the synthetic dataset generation and network training. From the parametric maps with skull computed by the CNN, we then analytically synthesized the same weighted images acquired in the *Brain MRI* protocol.

Finally, we propose a network refinement by performing additional training with a small number of real weighted images and their corresponding silver standard parametric maps obtained with the *Quantitative Brain MRI* protocol. We have carried out a cross validation procedure; specifically, we tested with $5 - t$ subjects, where $2 \leq t \leq 4$, and the remaining t subjects have been divided into training and early-stopping monitoring datasets; cross validation stems from the fact that we have $\binom{5}{t}$ combinations of testing datasets for each t; each combination will be hereinafter referred to as a split. Note that the case $t = 4$ corresponds to a leave-one-out scheme. Within this scheme, we have carried out two experiments: (i) the CNN previously trained with the synthetic dataset is fined tuned with the parametric maps and (ii) the CNN is trained from scratch making use exclusively of the parametric maps of the *Quantitative Brain MRI* protocol (i.e., no synthetic data are shown to the CNN). Maps from experiment (i) will be referred to as $\text{MAP}_{c\text{-(i)}}$ while maps from experiment (ii) will be denoted by $\text{MAP}_{c\text{-(ii)}}$.
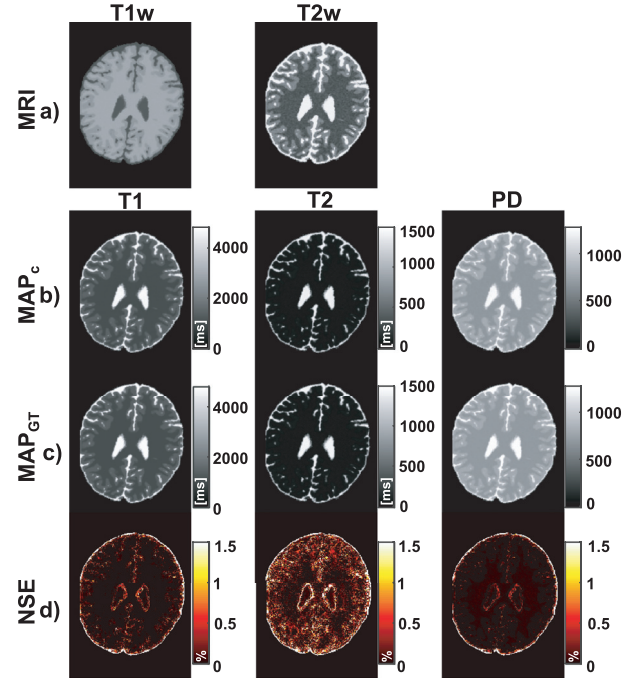

Fig. 2. A representative axial slice of the T1, T2, and PD maps computed from a test brain volume of the synthetic dataset. a) T1w and T2w images input to the network. b–d) Computed, ground-truth, and normalized squared error (NSE) maps from the same slice for the T1, T2, and PD parameter maps, respectively. The T1 and T2 values are given in miliseconds (ms). Main differences between the computed and the ground-truth maps appear in the boundary of the brain, although the NSE is predominantly below 1% on the three computed T1, T2, and PD maps.

3.5. Statistical analysis

The parameters defined in Eqs. (15)–(18) when applicable, are shown as averages (and sample standard deviation) along the 14 synthetic volumes used for testing or the 8 subjects used for system validation; these parameters are calculated within a 3D domain of the 14 central slices (where largest brain areas are found).

As for parameters defined in Eqs. (8), (9), (12)–(14) we have measured the Pearson correlation coefficient and the intra-class correlation coefficient ICC(2,1) [33]. As for the former, we run a correlation test based on the Fisher transformation to test the hypothesis that the correlation coefficient is less than or equal to a predefined value; a $p - value = P < 0.05$ was considered significant so as to reject the hypothesis. We have also analyzed Eqs. (12)–(14) using linear regression. Additionally, for the rectangular ROI drawn in the synthesized and acquired weighted images we have computed the Pearson correlation coefficient and performed an F-test for linear regression. Finally, we carry out a Bland-Altman plot analysis of a representative slice per subject where pixel values were normalized so that a value of "1.0" represented the signal strength of WM for each particular weighted image as in [25].

4. Results

4.1. Network testing with the synthetic dataset

Figure 2 shows a representative axial slice of the T1, T2, and PD maps computed from one of the test brain volumes of the synthetic dataset together with their corresponding NSE maps. Main differences between the computed and ground-truth maps appear in the boundary of the brain and in the tissue interfaces to a lower extent. Nevertheless, the NSE is predominantly below 1% on the three computed T1, T2, and PD maps. Further, the mean evalua-
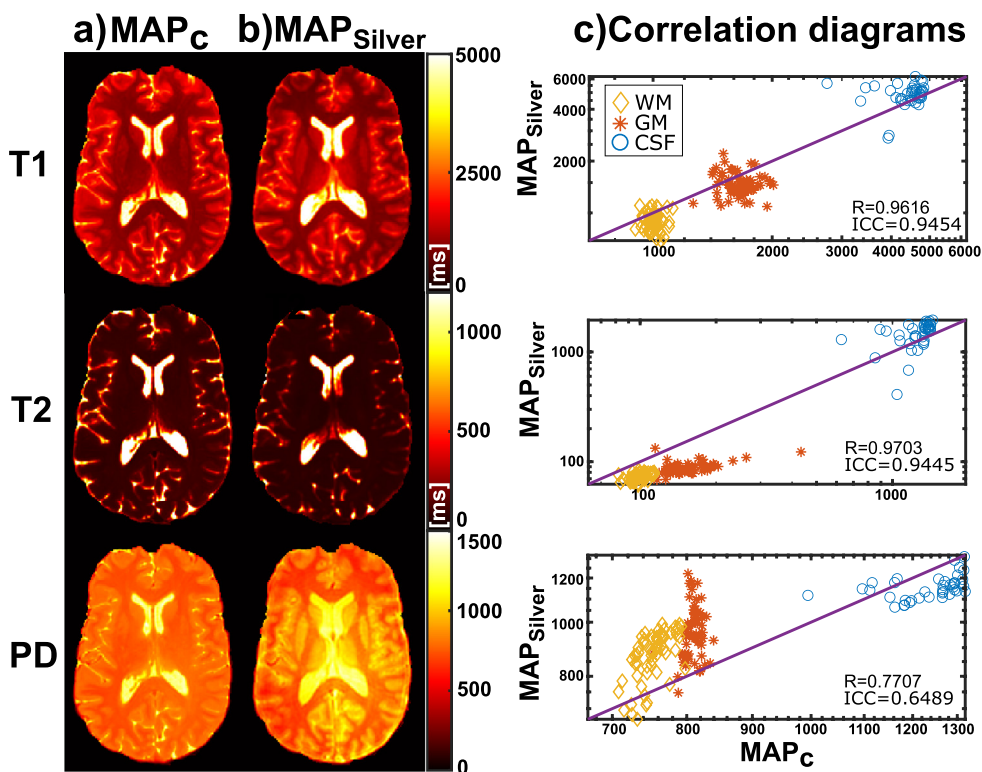
Fig. 3. A representative axial slice of T1, T2, and PD maps computed from a subject of the *Quantitative Brain MRI* protocol. a) Computed T1, T2, and PD parametric maps. b) Their corresponding silver standard relaxometry maps. c) Correlation of parameter $\mu_i^k(n)_L$ (Eq. 8) between the computed and the silver standard relaxometry maps for the five healthy subjects. T1 and T2 values are given in miliseconds (ms). The markers indicate the mean values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). Diagonal lines represent the identity.

Table 3

Metrics (mean ± std) used to evaluate the performance of the CNN to compute each set of T1, T2, and PD maps from each pair of T1w and T2w images of the test brain volumes of the synthetic dataset. These metrics are the mean squared error (MSE), structural similarity error index (SSIM), peak signal-to-noise ratio (PSNR), and correlation coefficient (CORR). The metrics were calculated between the computed parametric maps and the ground-truth T1, T2, and PD maps. Note that for the calculation of the metrics the background voxels were not considered.

	T1 map	T2 map	PD map
MSE	0.0072	0.0013	0.0004
	(0.0044)	(0.0010)	(0.0002)
SSIM	0.9932	0.9933	0.9912
	(0.0016)	(0.0044)	(0.0033)
PSNR	36.1274	33.8001	37.2614
	(2.2800)	(2.5282)	(1.5868)
CORR	0.9983	0.9975	0.9990
	(0.0007)	(0.0007)	(0.0004)

tion metrics obtained in the synthetic data testing of all 14 test brain volumes show good agreement between the computed and the ground-truth maps as can be seen in Table 3. The SSIM is always above 0.99 and the MSE below 1%.

4.2. Validation with actual MR acquisitions

Table 8 shows the parameters defined in Eqs. (6) and (7) for the T1, T2, and PD parametric maps obtained from all the ROIs within a tissue along all the subjects of the *Brain MRI* protocol. The values obtained in this work are mostly within the range of the values previously reported in the literature for a 3T scanner. As for the particular case of the PD maps, the GM/WM ratio is close to the ratio reported in the literature (1.22 vs. 1.10, respectively). In addi-

tion, Fig. 3 shows a representative axial slice of the T1, T2, and PD maps computed from a subject of the *Quantitative Brain MRI* protocol and their corresponding silver standard relaxometry maps. The computed parametric maps are visually realistic and capture most of the structural information without computational errors. Note that no outliers appear in the CSF of the T1 map. The correlation diagrams include the values of parameter $\mu_i^k(n)_L$ (Eq. (8)) for the five subjects of the *Quantitative Brain MRI* protocol. There is high correlation between the computed and the silver standard relaxometry maps, namely 0.9616, 0.9703, and 0.7707 for the T1, T2, and PD, respectively; the first two values are statistically higher than 0.90 ($P < 0.01$). Similarly, respective ICC values are 0.9454, 0.9445, and 0.6489.

Figure 4 shows a representative axial slice of weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding acquired images for a subject of the *Brain MRI* protocol. Overall, the synthesized and acquired weighted images are visually similar regarding both structural information and contrasts between tissues. The image modalities used to train the network present higher similarity than the others, being the T1w the most similar and the T2w FLAIR the least similar but yet with visual resemblance. The boundary of the CSF on the cortical area is hyperintense on the synthesized T2w FLAIR, which is presumably caused by partial volume effects. Supplementary Material Figure S1 shows an extended version of Fig. 4 with a representative axial slice for each of the eight subjects of the *Brain MRI* protocol.

Figure 5 shows a scatter plot between the synthesized and the acquired weighted images (we show the pairs of values $\mu_i^k(n)$ defined in Eq. (9) for the acquired and the synthesized images) for the eight subjects of *Brain MRI*. There is high correlation between the pairs of weighted images, namely 0.9979, 0.9952, 0.9912, 0.9820, and 0.9602 for the T1w, T2w, PDw, T2*w, and T2w
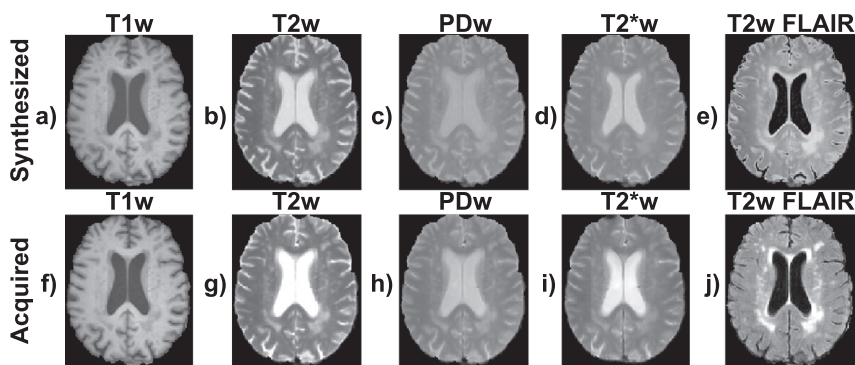
Fig. 4. A representative axial slice of the weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding acquired images. a–e) The synthesized T1w, T2w, PDw, T2*w, and T2w FLAIR images. f–j) Their corresponding acquired weighted images.
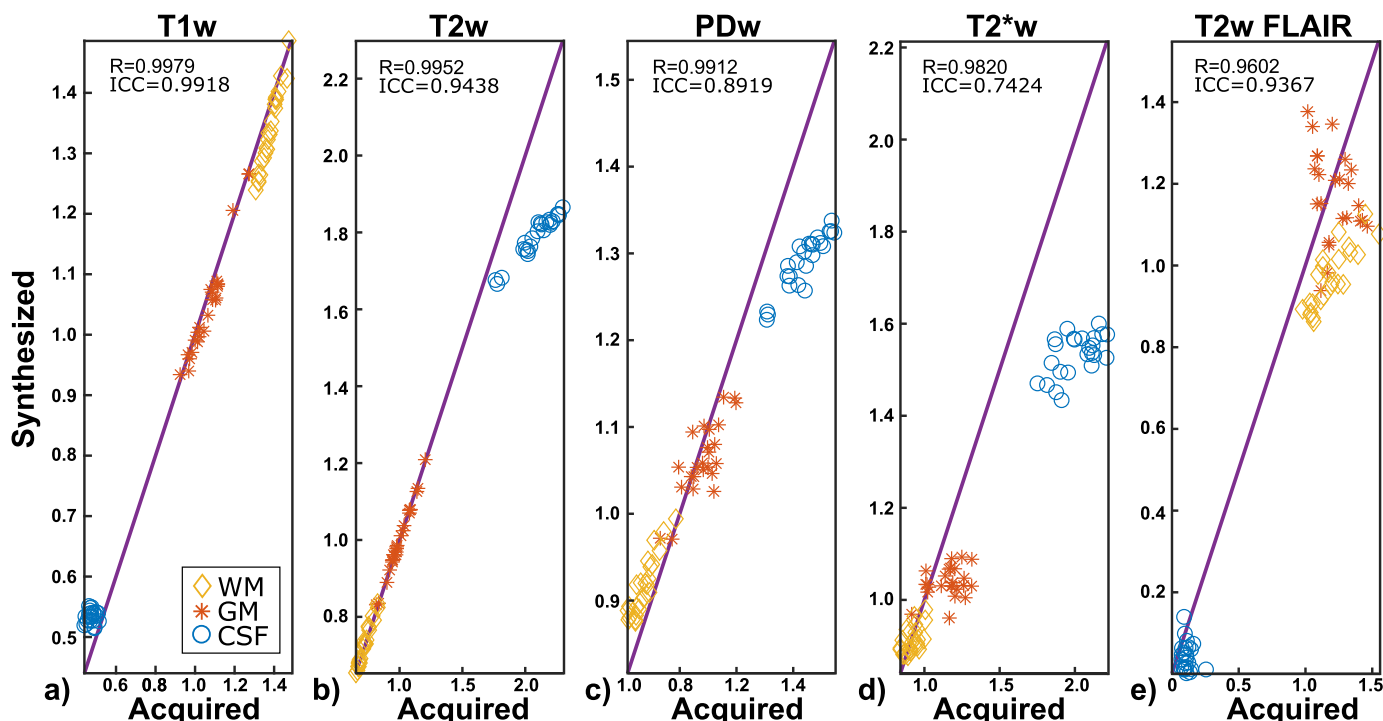

Fig. 5. Correlation of parameter $\mu_i^k(n)$ (Eq. 9) between the synthesized and the acquired weighted images. a) T1w, b) T2w, c) PDw, d) T2*w, and e) T2w FLAIR. The markers indicate the mean values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). The diagonal lines represent the identity.

FLAIR, respectively; all of these values are statistically higher than 0.90 ($P < 0.001$). Similarly, respective ICC values are 0.9918, 0.9438, 0.8919, 0.7424, and 0.9367.

Figure 6 shows similar scatter plots of the contrast, CNR, and SNR samples between the synthesized and acquired weighted images (as defined in Eqs. (12)–(14)). There is high correlation with values between 0.9907 and 0.9241 for the contrast, 0.9807 and 0.8739 for the CNR, and 0.9845 and 0.9082 for the SNR for all the weighted images, as shown in Table 7; most of these values are statistically larger than 0.90 ($P < 0.05$), except in the case of the T2*w and T2w FLAIR for the CNR and SNR, and the PDw only for the SNR, which are statistically greater than 0.84. The ICC values for the three same samples are between 0.9730 and 0.9421 for the T1w, between 0.9579 and 0.8954 for the T2w, and between 0.9120 and 0.8607 for the T2w FLAIR. In contrast, the ICC values are lower for the PDw and T2*w, as shown in Table 7.

Linear regression showed that the SNR of the synthesized weighted images is generally better with an improvement that reaches 47.74% (CI: [41.93%; 53.54%]). CNR is fairly similar for

the T1w (-5.31%, CI: $[-6.33\%; -4.29\%]$), the T2*w (-0.12%, CI: $[-2.75\%; 2.52\%]$), and the T2w FLAIR (5.36%, CI: [1.67%; 9.05%]), although it is slightly worse for the T2w (-17.46%, CI: $[-18.94\%; -15.99\%]$) and the PDw (-26.45%, CI: $[-27.79\%; -25.10\%]$). Finally, contrast only improves in the T2w FLAIR (22.77%, CI: [20.57%; 24.98%]). See details in Table 4.

Figure 7 shows a scatter plot between the pixel values of the rectangular ROIs drawn on the synthesized and the acquired weighted images of a representative subject of *Brain MRI*. There is high correlation between the pairs of weighted images, namely 0.9911, 0.9684, 0.8898, 0.8477, and 0.6403 for the T1w, T2w, PDw, T2*w, and T2w FLAIR, respectively; all of these values are statistically significant ($P << 0.0001$) in the F-test for linear regression. For the sake of completeness, the correlation values for each of the eight subjects of the *Brain MRI* protocol are given in Supplementary Material Table S1.

Additionally, the high values of the mean SSIM, SNR, and CORR and low values of the MSE obtained in the subjects of *Brain MRI* show good agreement between synthesized and acquired weighted
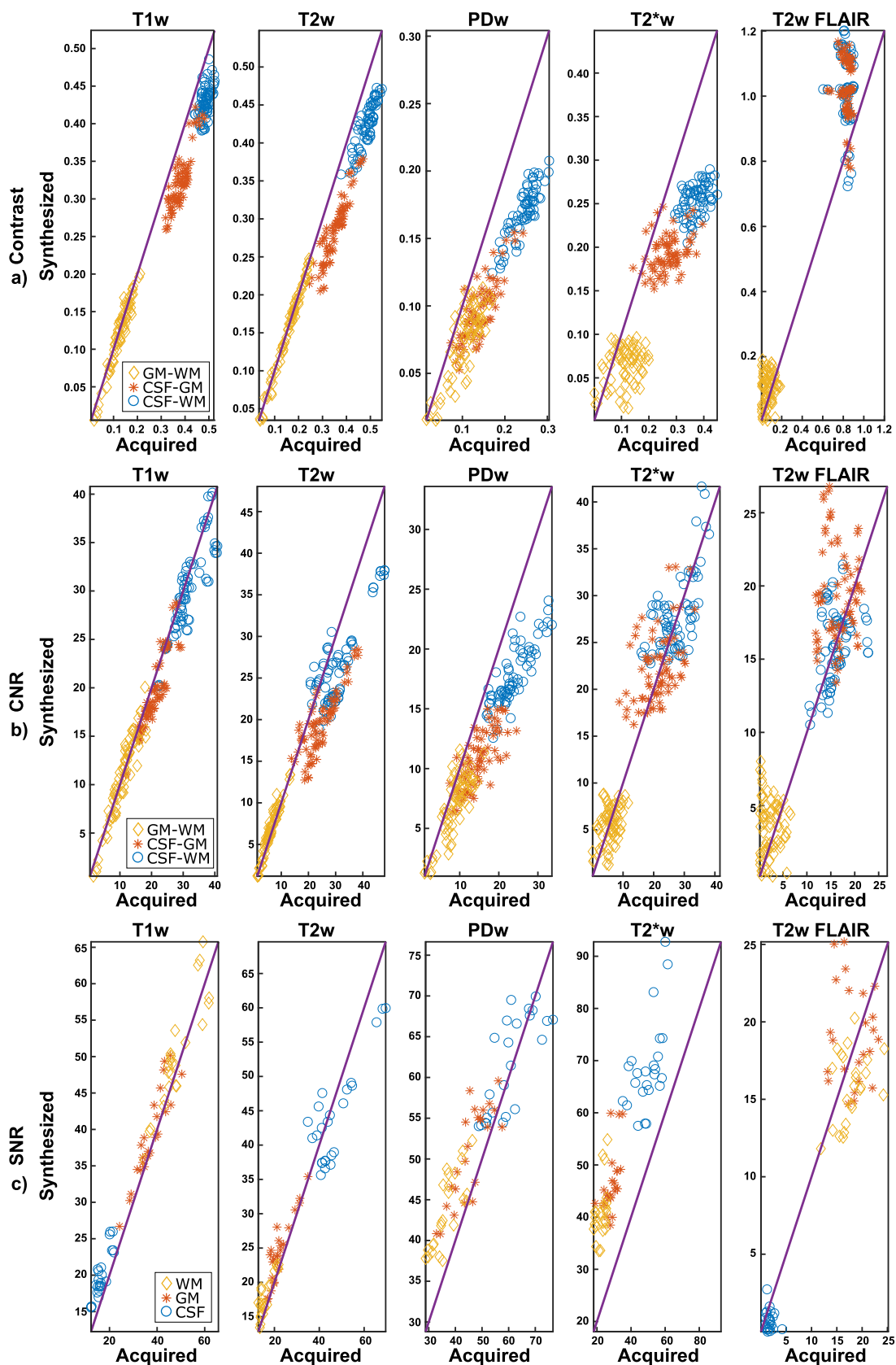
Fig. 6. Correlation of the contrast, the contrast-to-noise ratio (CNR), and the signal-to-noise ratio (SNR) between the synthesized and the acquired weighted images. a) Contrast, b) CNR, and c) SNR of the T1w, T2w, PDw, T2*w, and T2w FLAIR images. For the Contrast (a) and the CNR (b) the markers indicate the contrast/CNR values between each combination of the GM ROIs with the WM ROIs (yellow diamonds), each combination of CSF ROIs with the GM ROIs (red stars), and each combination of CSF ROIs with the WM ROIs (blue circles). For the SNR (c) the markers indicate the mean SNR values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). The diagonal lines represent the identity.

Table 4

Values defined in Eqs. (6) and (7) (the latter, within braces) for each tissue in each of the computed parametric maps (i.e. T1, T2, and PD maps). Comparison with the values previously reported in the literature for a 3T scanner. Note that WM is the white matter, GM is the grey matter, and CSF is the cerebrospinal fluid.

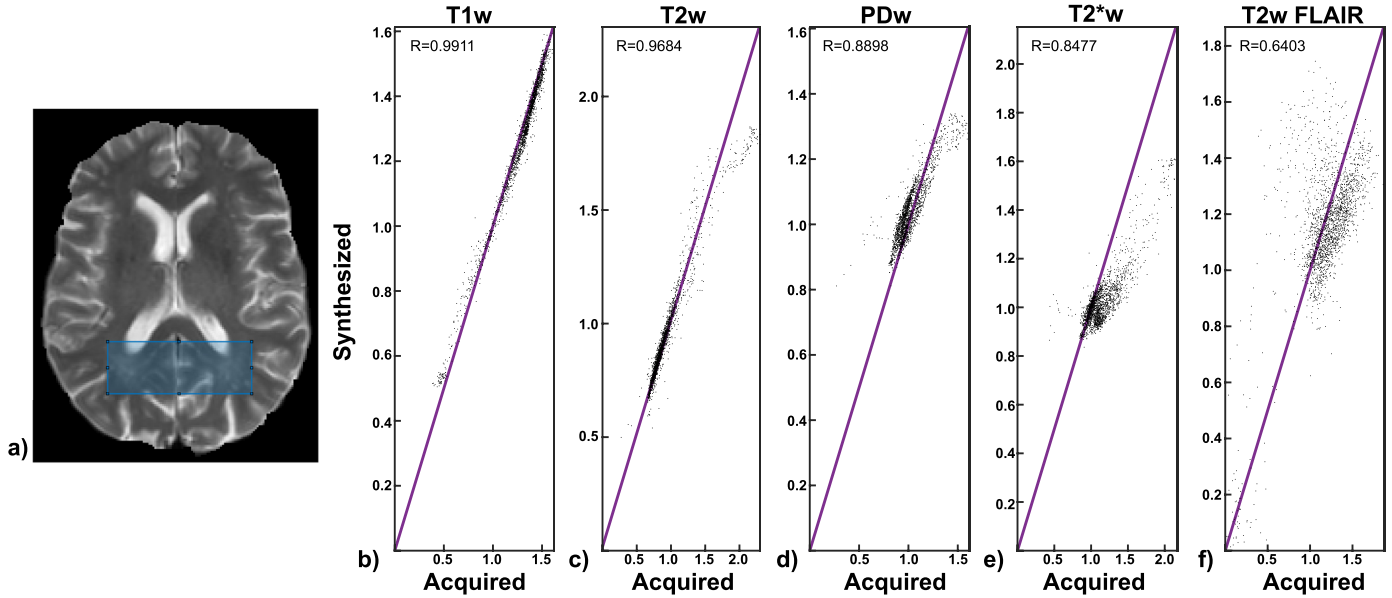|	T1 (s)		T2 (s)		PD	
	This work	Literature	This work	Literature	This work	Literature
WM	0.9741	0.7370–1.1000	0.0890	0.0560–0.0840	0.7222	0.6330
	(0.0585)	[Ref. [36,37]]	(0.0061)	[Ref. [38,39]]	(0.0142)	[Ref. [40]]
GM	1.4474	1.3310–1.8200	0.1257	0.0710–0.1320	0.7988	0.7720
	(0.1361)	[Ref. [39,41]]	(0.0160)	[Ref. [38,39]]	(0.0135)	[Ref. [40]]
CSF	4.6785	3.7000–6.8730	1.3705	0.5000–1.8700	1.2601	
	(0.1060)	[Ref. [14,42]]	(0.0335)	[Ref. [43,44]]	(0.0380)	

Fig. 7. Correlation between pixel values from the rectangular ROI of the synthesized and the acquired weighted images for a representative subject of the *Brain MRI* protocol. a) Rectangular region of interest chosen for linear regression, b) T1w, c) T2w, d) PDw, e) T2*w, and f) T2w FLAIR. The markers indicate the value of each pixel in the rectangular ROI (approximate 2025 values). The diagonal lines represent the identity. $P << 0.0001$ for all image modalities for the correlation test carried out (see correlation values tested ín main text).

Table 5

Correlation coefficient (R) and intraclass correlation coefficient (ICC) of the contrast, the contrast-to-noise ratio (CNR), and the signal-to-noise ratio (SNR) between the synthesized and the acquired weighted images (see Fig. 6). The bold correlation values indicate that they are statistically significant superior to a correlation value of 0.9. The * indicates $P < 0.05$ and ** $P < 0.001$.

		T1w	T2w	PDw	T2*w	T2w FLAIR
Contrast	R	**0.9907****	**0.9855****	**0.9591****	**0.9241***	**0.9689****
	ICC	0.9421	0.9155	0.6521	0.6530	0.8961
CNR	R	**0.9807****	**0.9658****	**0.9453****	0.9193	0.8739
	ICC	0.9730	0.8954	0.7425	0.9187	0.8607
SNR	R	**0.9845****	**0.9734****	0.9280	0.9082	0.9145
	ICC	0.9712	0.9579	0.8300	0.4634	0.9120

images as shown in Table 5. Specifically, the SSIM achieves values above 0.96 for the T1w and the T2w, and of 0.91, 0.78, and 0.56 for the PDw, T2*w, and T2w FLAIR, respectively. The MSE is below 1% for the T1w, T2w, and PDw, and below 9% for the T2*w and the T2w FLAIR. Similarly to Fig. 4, the image modalities used to train the network show higher SSIM, SNR, and CORR and lower MSE than the others.

Figure 8 represents Bland-Altman plots including data from a representative slice of all subjects of *Brain MRI*. It compares synthesized and acquired pixel values for T1w, T2w, PDw, T2*w, and T2w

FLAIR. The absolute mean difference for the each image modality is 0.0041, 0.0021, 0.0089, 0.0794 and 0.0249, respectively.

Figure 9 displays a representative axial slice of additional weighted images synthesized with the same sequences as in *Brain MRI*, but varying the sequence parameters. This proves the versatility of the proposed approach to synthesize any weighted images. The images obtained are realistic and with coherent contrasts.

Finally, Fig. 10 shows a representative axial slice of the non skull-stripped weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding non skull-stripped acquired images for a subject of *Brain MRI*. Similarly to Fig. 4, both images are visually apparent regarding both structural information and contrast between tissues. Nevertheless, the inhomogeneities in the skull interfaces might cause a mismatch between the synthesized and the acquired images.

4.3. Fine tuning: refining the network with actual parametric maps

Figure 11 shows a representative slice of both MAP$_{c-(ii)}$ and MAP$_{c-(i)}$ (columns a) and b), respectively) with MAP as T1, T2, and PD (first, second, and third rows, respectively). For MAP$_{c-(i)}$ the figure also shows the correlation diagrams that include the values of the ROIs for all subjects of the *Quantitative Brain MRI* protocol tested with a leave-one-out scheme (i.e., $t = 4$). It can be seen that the fine tuning procedure improves the accuracy of the
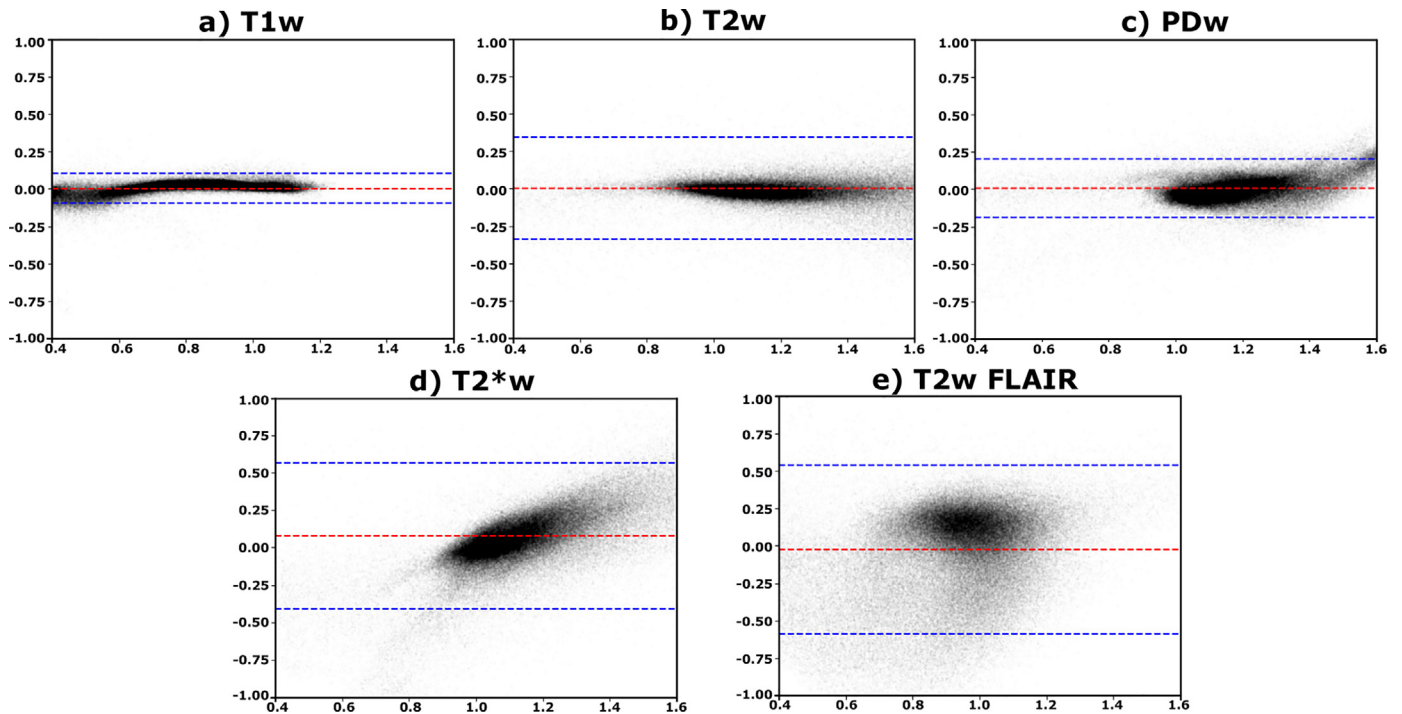
Fig. 8. Bland-Altman plots used to compare synthesized and acquired weighted images as in [25]. Each plot combines results from a representative axial slice of all subjects of the *Brain MRI* protocol. The corresponding image modalities are: a–e) T1w, T2w, PDw, T2*w, and T2w FLAIR, respectively. Red dashed lines represents the bias and blue dashed lines the 95% confidence interval.

Table 6

Percentage of variation of the linear regression coefficient [95% confidence interval (CI)] in comparison to the identity (i.e., linear regression coefficient of one) for the contrast, the contrast-to-noise ratio (CNR), and the signal-to-noise ratio (SNR) samples. For the linear regression computation, the x-axis is considered as the samples values of the acquired image and the y-axis the samples values of the synthesized image as shown in Fig. 6. Positive values indicate an improvement of the corresponding samples.

	T1w	T2w	PDw	T2*w	T2w FLAIR
Contrast	−12.69% [−13.36%; −12.01%]	−14.11% [−14.88%; −13.33%]	−31.22% [−32.23%; −30.21%]	−32.89% [−34.37%; −31.34%]	22.77% [20.57%; 24.98%]
CNR	−5.31% [−6.33%; −4.29%]	−17.46% [−18.94%; −15.99%]	−26.45% [−27.79%; −25.10%]	−0.12% [−2.75%; 2.52%]	5.36% [1.67%; 9.05%]
SNR	4.76% [2.89%; 6.62%]	−3.10% [−5.92%; −0.28%]	6.86% [4.22%; 9.51%]	47.74% [41.93%; 53.54%]	−5.76% [−11.08%; −0.44%]

Table 7

Metrics (mean ± std) used to evaluate the capability to synthesize weighted images from a set of T1, T2, and PD maps computed by the CNN. These metrics are the mean squared error (MSE), structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and correlation coefficient (CORR). The metrics were calculated between both the synthesized and the acquired weighted images. Note that for the calculation of the metrics the background voxels were not considered.

	T1w	T2w	PDw	T2*w	T2w FLAIR
MSE	0.0058 (0.0009)	0.0095 (0.0020)	0.0061 (0.0010)	0.0392 (0.0058)	0.0815 (0.0081)
SSIM	0.9651 (0.0051)	0.9620 (0.0039)	0.9194 (0.0078)	0.7823 (0.0222)	0.5693 (0.0190)
PSNR	30.6338 (1.5330)	26.2621 (0.7607)	25.3160 (0.7972)	18.9098 (0.5905)	19.6598 (1.9280)
CORR	0.9910 (0.0015)	0.9858 (0.0023)	0.9886 (0.0017)	0.9438 (0.0076)	0.8726 (0.0093)

Table 8

Mean correlation coefficient (R) and intraclass correlation coefficient (ICC) of parameter $\mu_i^k(n)_L$ (Eq. (8)) between the computed and the silver standard relaxometry maps for three values of the number of training subjects t in the cross validation of the fine tuning. Values reported have been computed for each test subjects of each split, and then, mean values were computed along all splits and subjects.

		T1	T2	PD
R	$t = 4^*$	0.9784	0.9682	0.8912
	$t = 3^\dagger$	0.9733	0.9648	0.8857
	$t = 2^\dagger$	0.9722	0.9645	0.8752
ICC	$t = 4^*$	0.9517	0.9607	0.7935
	$t = 3^\dagger$	0.9324	0.9496	0.7930
	$t = 2^\dagger$	0.9275	0.9475	0.7536

* five splits, † 10 splits.

computed parametric maps in terms of ICC (compare the values shown in Fig. 3), whereas without the previous synthetic training the results worsen noticeably and the maps blur. Furthermore, Table 6 shows the mean correlation coefficient and ICC of parameter $\mu_i^k(n)_L$ (Eq. (8)) for the different configurations of t. Both correlation and ICC have been computed for each test subject of each split, and then, mean values were computed along all splits and subjects. Results show that, as expected, both parameters increase with the number of training subjects.
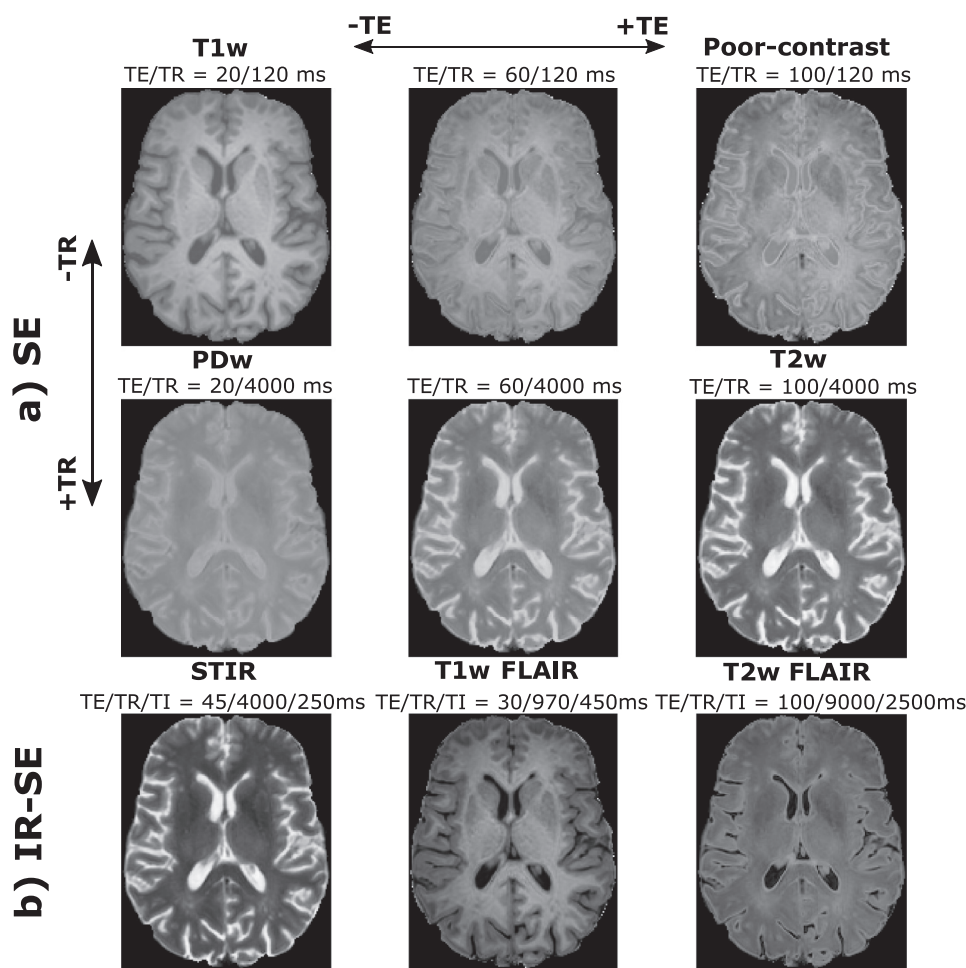
Fig. 9. A representative axial slice of other weighted images synthesized varying the sequence parameters from a set of the T1, T2, and PD maps computed by the CNN. a) Weighted images synthesized for a spin echo (SE) sequence with different TE and TR corresponding to T1w, T2w and PDw image modalities. b) Weighted images synthesized for an inversion recovery spin echo (IR-SE) sequence with different TE, TR, and TI corresponding to short-TI inversion recovery (STIR), T1w FLAIR, and T2w FLAIR image modalities. Note that the unlabeled images correspond to sequence parameter combinations which lead to weighted images with undefined contrast.

5. Discussion

In this work, we have presented a novel joint synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images from only a pair of input weighted images. The pair of input images are a T1w and a T2w acquired with clinical routine sequences. The parametric maps are obtained by training the CNN with a synthetic dataset; hence, we overcome the lack of a public and sufficiently large database of conventional images that should be accompanied by their corresponding parametric maps. Our synthetic training dataset departs from 120 instances of BrainWeb maps, in which we add intensity variability, by means of random noise, as well as spatial variability, by registering these maps to different anatomies from the PPMI database. We show the feasibility of this solution by computing accurate and realistic parametric maps from both synthetic and actual MR brain acquisitions; the computed maps are then used to synthesize different weighted images, so our end-to-end synthetic MRI solution is not limited to a number of predefined weighted images that have entered the learning process, but is capable of generalizing to any image modality that can be synthesized out of the parametric maps. Hence, our solution fulfils the three conditions needed to become a synthetic MRI method. To the best of our knowledge this is the first synthetic MRI method that is based on conventional routine sequences and can be trained on the basis of synthetic data.

We have shown that synthesized weighted images from five clinical routine sequences achieve high similarity metrics, with SSIM usually above 0.90 and low error with MSE always below 9%. The correlation analysis shown in the scatter plots of Fig. 5 provide values above 0.95 for all modalities. Similarly, for the scatter plots of contrast, CNR, and SNR (Fig. 6), both correlation and ICC also obtain high values, as shown in Table 7. Note that the agreement when the ICC values are above 0.75 is considered good while when the values are above 0.90 is considered excellent [33]; our results indicate that we lie in these ranges for at least one parameter for each synthesized modality. In addition, spatial resolutions of training and test images do not need to exactly match. Our testing images have resolution of $0.94 \times 1.25 \times 5$ mm while the PPMI datase resolution is of $1 \times 1 \times 1.2$ mm; despite the in-plane resolution does not differ much, slice thicknesses are clearly different and no partial volumes effects in the through-plane direction are obvious in our solution.

The synthetic MRI approach proposed may have important implications in neuroimaging due to the utility of the parametric maps for tissue characterization and the possibility of synthesizing any weighted image. Specifically, the obtained T1, T2, and PD values of the three tissues (WM, GM, and CSF) present a good correspondence with the values reported in the literature, as shown in Table 8 and with the silver standard relaxometry parametric maps with correlation values above 0.95 for the T1 and T2 maps. The output quality increases noticeably when the network is fine tuned
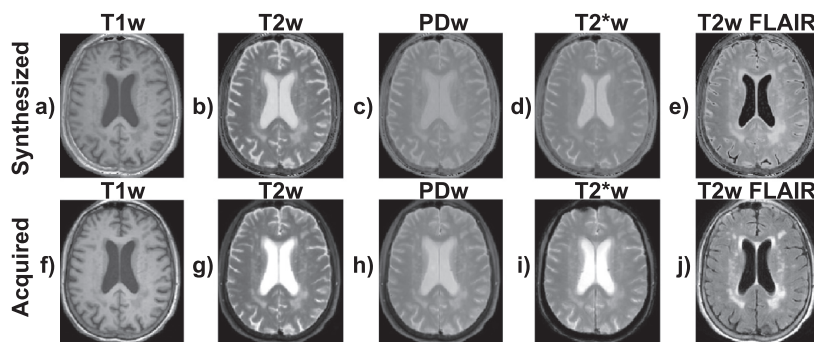
Fig. 10. A representative axial slice of the non skull-stripped weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding non skull-stripped acquired images. a–e) The synthesized T1w, T2w, PDw, T2*w, and T2w FLAIR images. f–j) Their corresponding acquired weighted images.
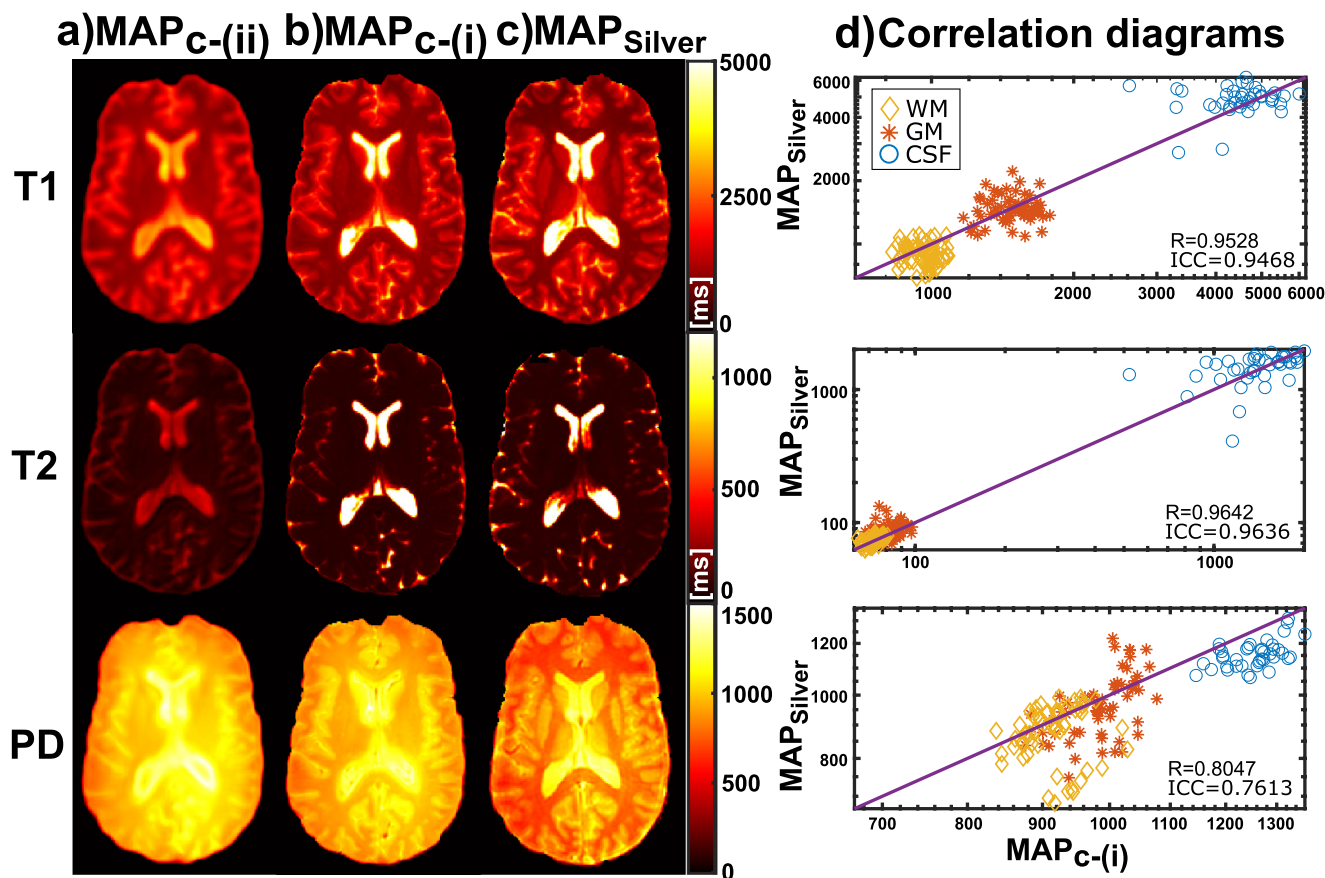

Fig. 11. A representative axial slice of T1, T2, and PD maps computed from a subject of the *Quantitative Brain MRI* protocol in a leave-one-out scheme. a) T1, T2, and PD parametric maps computed by the network trained from scratch with actual parametric maps. b) Corresponding maps computed by the fine tuned network with previous synthetic training. c) Their corresponding silver standard relaxometry maps. d) Correlation of parameter $\mu_t^k(n)_L$ (Eq. (8)) between MAP$_{c-(i)}$ and MAP$_{Silver}$ for the five healthy subjects. Diagonal lines represent the identity.

with a small number of silver standard maps. This provides a way to obtain parametric maps with increased accuracy, at the cost of employing a (small) number of silver standard maps for additional training. Note, however, that training with synthetic data is a key step, since training from scratch with this small amount of silver standard maps by no means suffices. Moreover, the proposed approach avoids the need of lengthy relaxometry sequences; the total scan time of the full-brain acquisition described in this manuscript (T1w and T2w acquisitions) is less than 8 min versus the 18 min scan time of an inversion recovery golden standard acquisition only for T1 mapping [34], and the 17 min scan time of the DESPOT algo-

rithm for T1 and T2 mapping [10]. The computed parametric maps are therefore less prone to motion artifacts. Interestingly, the proposed approach is not based on specific and complex sequences as MRF [15], IR-TrueFISP [21], QRAPMASTER [22], and MPME [25] or private protocols as the SyMRI IMAGE software [35]. Also, the feasibility of synthesizing weighted images and/or retrospectively optimizing sequence parameters can further reduce scan time. Thus, a radiologist could have the parametric maps together with various conventional weighted images based on the same widespread short scan protocol. In addition, we provide the possibility of creating databases of perfectly registered weighted images accompa-

nied with their corresponding parametric maps; these databases can be used to train machine learning algorithms for different purposes, perform data augmentation or improve the performance of registration or segmentation algorithms. The field of *radiomics* also seems a natural target for our methodology.

We should stress that our method gives rise to different modalities, some of them unseen by the network throughout the training process, with comparable quality with recent medical image translation works; however, to the best of our knowledge, these works are limited to the specific modalities that enter the training and validation stages. Specifically, SSIM in our synthesized T1w, T2w, and PDw images is slightly higher than the values reported by Chartsias et al. [17] and Sohail et al. [19] for some of these image modalities, albeit the T2w FLAIR and sometimes the T2*w achieve lower quality. A more thorough comparison is not feasible since our actual acquired validation dataset is not large enough to train the state-of-the-art medical image translation methods. Also, note that in our approach no data from real acquisitions are used in the training stage thanks to the synthetic training, and only when fine tuning the network a very small database of actual maps is used; this is our main advantage with respect to the state-of-the-art. The non skull-stripped synthesized images, although visually realistic, achieve lower quality than their corresponding skull-stripped counterparts; however, this seems to be the case as well in [17]. The loss of quality is clearer in the neighbouring parts of the skull, and other tissues such as the eyes. We obtain a SSIM of 0.80 in both the T2w and the PDw images while in [17] the SSIM in the synthesis of the T2w from the PDw is of 0.86. Nevertheless, it is important to note that, as opposed to [17], in our work the skull has not entered the training process.

As for the comparison with synthetic MRI methods, in our work the synthesized weighted images show higher visual resemblance to the acquired images than in the other methods [21,25]. In addition, correlation coefficients calculated within the rectangular ROIs described in Section 2.1 are higher with our approach [21]. However, except for the T2w FLAIR modality, our synthesized weighted images present lower contrast values than the acquired weighted images as compared with [24], although we achieve similar or higher CNR and SNR figures. The Bland-Altman plots show better agreement than [25] in T1w and T2w, similar agreement in PDw, and only a slightly lower agreement in T2w FLAIR images. The T2*w modality is not synthetized by them. The loss of quality in T2w FLAIR images is a common issue in synthetic MRI [6] where the boundary of the CSF on the cortical area tends to be hyperintense presumably due to partial volume effects.

This work has several limitations. The method was evaluated in synthetic data, eight subjects —suspected of early Alzheimer disease—, and five healthy volunteers, so further validation in a larger cohort of both healthy volunteers and patients with other pathologies is still needed. Moreover, B0 and B1 inhomogeneties have not been taken into account in the synthetic dataset generation. Thus, including these inhomogeneties in the synthesis could be of interest. In addition, the equations used to synthesize the weighted images did not consider all the effects that occur in practice. For example, the T2*w image is synthesized from the T2 map instead of the T2* map which could be the cause of the worse metrics compared to the metrics of the T2w image. Also, the T2w FLAIR presents worse metrics than the other modalities, but the images obtained are comparable with those of the literature [24]. To address this, methods focused on improving T2w FLAIR images have been reported [45]. Additionally, the parametric maps of the synthetic dataset were generated with values corresponding to 3T scanners; hence our results do not directly carry over to other field strengths. The extension to high field scanners will presumably require to modify the synthetic training dataset and further postprocessing corrections because B1 and B0 field inhomogeneities are specially problematic at high fields.

Future work includes improvement and further tuning in the implemented CNN. In addition, the use of generative adversarial networks (GAN) architectures may be studied due to the recent works that achieve impressive results in medical image translation capturing high-frequency texture information [18,46]. On the other hand, the selection of the optimal input training sequences and/or sequence parameters could improve the computation of the parametric maps and, subsequently, the synthesis of the weighted images. A more realistic synthesis of weighted images with a detailed Bloch simulation [47,48] and other maps (e.g. T2*, B0, and B1 maps) could also enhance the synthesis quality of any MRI modality. Further, the simulation of motion in the synthetic dataset could enhance the network robustness against such artifacts. In addition, we could extend the proposed approach to other tissues, such as the heart or the liver, as long as synthetic anatomical volumes can be computed. To this end, for example, the extended Cardiac-Torso (XCAT) phantoms [49] could be employed. We also plan to address the oncology field, where parametric mapping may be a challenging task.

6. Conclusion

We propose a novel joint synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images which only needs two clinical routine weighted images as inputs (full-brain acquisition in less than 8 min of scan time). Based on a CNN, we are able to provide realistic parametric maps and weighted images when training the CNN with a synthetic dataset. The results in both synthetic data and actual MR acquisitions experiments demonstrate its feasibility for quantitative MRI in clinical viable times as well as its applicability for the synthesis of additional MR weighted image modalities.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2021.106371.

References

[1] H. Lu, L.M. Nagae-Poetscher, X. Golay, D. Lin, M. Pomper, P.C. Van Zijl, Routine clinical brain MRI sequences for use at 3.0 tesla, J. Magn. Reson. Imaging 22 (1) (2005) 13–22, doi:10.1002/jmri.20356.

[2] N. Weiskopf, J. Suckling, G. Williams, M.M. Correia, B. Inkster, R. Tait, C. Ooi, E.T. Bullmore, A. Lutti, Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation, Front. Neurosci. 7 (2013) 95, doi:10.3389/fnins.2013.00095.

[3] P. Conlon, M. Trimble, D. Rogers, C. Callicott, Magnetic resonance imaging in epilepsy: a controlled study, Epilepsy Res. 2 (1) (1988) 37–43, doi:10.1016/0920-1211(88)90008-3.

[4] H. Larsson, J. Frederiksen, J. Petersen, A. Nordenbo, I. Zeeberg, O. Henriksen, J. Olesen, Assessment of demyelination, edema, and gliosis by in vivo determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis, Magn. Reson. Med. 11 (3) (1989) 337–348, doi:10.1002/mrm.1910110308.

[5] T.E. Yankeelov, D.R. Pickens, R.R. Price, Quantitative MRI in Cancer, Taylor & Francis, 2011.

[6] A. Hagiwara, M. Warntjes, M. Hori, C. Andica, M. Nakazawa, K.K. Kumamaru, O. Abe, S. Aoki, SyMRI of the brain: rapid quantification of relaxation rates and proton density, with synthetic MRI, automatic brain segmentation, and myelin measurement, Invest. Radiol. 52 (10) (2017) 647–657, doi:10.1097/RLI.0000000000000365.

[7] G. Cooper, S. Hirsch, M. Scheel, A.U. Brandt, F. Paul, C. Finke, P. Boehm-Sturm, S. Hetzer, Quantitative multi-parameter mapping optimized for the clinical routine, Front. Neurosci. 14 (2020) 1290, doi:10.3389/fnins.2020.611194.

[8] S. Ji, D. Yang, J. Lee, S.H. Choi, H. Kim, K.M. Kang, Synthetic MRI: technologies and applications in neuroradiology, J. Magn. Reson. Imaging (2020), doi:10.1002/jmri.27440.

[9] S.A. Bobman, S.J. Riederer, J.N. Lee, S.A. Suddarth, H.Z. Wang, B.P. Drayer, J.R. MacFall, Cerebral magnetic resonance image synthesis., Am. J. Neuroradiol. 6 (2) (1985) 265–269.

[10] S.C. Deoni, T.M. Peters, B.K. Rutt, High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2, Magn. Reson. Med. 53 (1) (2005) 237–241, doi:10.1002/mrm.20314.

[11] A. Den Dekker, J. Sijbers, Data distributions in magnetic resonance images: a review, Physica Medica 30 (7) (2014) 725–741, doi:10.1016/j.ejmp.2014.05.002.

[12] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Björk, F. Vanhevel, P.M. Parizel, R.S.J. Estépar, J. Arnold, J. Sijbers, NOVIFAST: a fast algorithm for accurate and precise VFA MRI T1 mapping, IEEE Trans. Med. Imaging 37 (11) (2018) 2414–2427, doi:10.1109/TMI.2018.2833288.

[13] S. Barbieri, O.J. Gurney-Champion, R. Klaassen, H.C. Thoeny, Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI, Magn. Reson. Med. 83 (1) (2020) 312–321, doi:10.1002/mrm.27910.

[14] S. Clare, P. Jezzard, Rapid T1 mapping using multislice echo planar imaging, Magn. Reson. Med. 45 (4) (2001) 630–634, doi:10.1002/mrm.1085.

[15] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J.L. Sunshine, J.L. Duerk, M.A. Griswold, Magnetic resonance fingerprinting, Nature 495 (7440) (2013) 187–192, doi:10.1038/nature11971.

[16] B. Rieger, F. Zimmer, J. Zapp, S. Weingärtner, L.R. Schad, Magnetic resonance fingerprinting using echo-planar imaging: Joint quantification of T1 and relaxation times, Magn. Reson. Med. 78 (5) (2017) 1724–1733, doi:10.1002/mrm.26561.

[17] A. Chartsias, T. Joyce, M.V. Giuffrida, S.A. Tsaftaris, Multimodal MR synthesis via modality-invariant latent representation, IEEE Trans. Med. Imaging 37 (3) (2017) 803–814, doi:10.1109/TMI.2017.2764326.

[18] S.U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, T. Çukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks, IEEE Trans. Med. Imaging 38 (10) (2019) 2375–2388, doi:10.1109/TMI.2019.2901750.

[19] M. Sohail, M.N. Riaz, J. Wu, C. Long, S. Li, Unpaired multi-contrast mr image synthesis using generative adversarial networks, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2019, pp. 22–31, doi:10.1007/978-3-030-32778-1_3.

[20] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1125–1134.

[21] V. Gulani, P. Schmitt, M.A. Griswold, A.G. Webb, P.M. Jakob, Towards a single-sequence neurologic magnetic resonance imaging examination: multiple-contrast images from an IR TrueFISP experiment, Invest. Radiol. 39 (12) (2004) 767–774, doi:10.1097/00004424-200412000-00008.

[22] J. Warntjes, O.D. Leinhard, J. West, P. Lundberg, Rapid magnetic resonance quantification on the brain: optimization for clinical usage, Magn. Reson. Med. 60 (2) (2008) 320–329, doi:10.1002/mrm.21635.

[23] L.N. Tanenbaum, A.J. Tsiouris, A.N. Johnson, T.P. Naidich, M.C. DeLano, E.R. Melhem, P. Quarterman, S. Parameswaran, A. Shankaranarayanan, M. Goyen, et al., Synthetic MRI for clinical neuroimaging: results of the magnetic resonance image compilation (magic) prospective, multicenter, multireader trial, Am. J. Neuroradiol. 38 (6) (2017) 1103–1110, doi:10.3174/ajnr.A5227.

[24] I. Blystad, J.B.M. Warntjes, O. Smedby, A.-M. Landtblom, P. Lundberg, E.-M. Larsson, Synthetic MRI of the brain in a clinical setting, Acta Radiol. 53 (10) (2012) 1158–1163, doi:10.1258/ar.2012.120195.

[25] C.C. Cheng, F. Preiswerk, B. Madore, Multi-pathway multi-echo acquisition and neural contrast translation to generate a variety of quantitative and qualitative image contrasts, Magn. Reson. Med. 83 (6) (2020) 2310–2321, doi:10.1002/mrm.28077.

[26] J. Bittoun, J. Taquin, M. Sauzade, A computer algorithm for the simulation of any nuclear magnetic resonance (NMR) imaging method, Magn. Reson. Imaging 2 (2) (1984) 113–120, doi:10.1016/0730-725X(84)90065-1.

[27] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, G.B. Pike, A.C. Evans, BrainWeb: online interface to a 3D MRI simulated brain database, NeuroImage 5 (1997) 425.

[28] J.Z. Bojorquez, S. Bricq, C. Acquitter, F. Brunotte, P.M. Walker, A. Lalande, What are normal relaxation times of tissues at 3 T? Magn. Reson. Imaging 35 (2017) 69–80, doi:10.1016/j.mri.2016.08.021.

[29] Y. Xiao, V. Fonov, M.M. Chakravarty, S. Beriault, F. Al Subaie, A. Sadikot, G.B. Pike, G. Bertrand, D.L. Collins, A dataset of multi-contrast population-averaged brain MRI atlases of a Parkinson's disease cohort, Data Brief 12 (2017) 370–379, doi:10.1016/j.dib.2017.04.013.

[30] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith, FSL, NeuroImage 62 (2) (2012) 782–790, doi:10.1016/j.neuroimage.2011.09.015.

[31] Ó. Peña-Nogales, T.M. Ellmore, R. de Luis-García, J. Suescun, M.C. Schiess, L. Giancardo, Longitudinal connectomes as a candidate progression marker for prodromal Parkinson's disease, Front. Neurosci. 12 (2019) 967, doi:10.3389/fnins.2018.00967.

[32] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980(2014).

[33] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropractic Med. 15 (2) (2016) 155–163, doi:10.1016/j.jcm.2016.02.012.

[34] G. Ramos-Llordén, J. Arnold, G. Van Steenkiste, B. Jeurissen, F. Vanhevel, J. Van Audekerke, M. Verhoye, J. Sijbers, A unified maximum likelihood framework for simultaneous motion and T1 estimation in quantitative MR T1 mapping, IEEE Trans. Med. Imaging 36 (2) (2016) 433–446, doi:10.1109/TMI.2016.2611653.

[35] SyntheticMR, SyMRI IMAGE, (https://syntheticmr.com/products/symri-image/), [Online; accessed 20-January-2021].

[36] D.C. Zhu, R.D. Penn, Full-brain T1 mapping through inversion recovery fast spin echo imaging with time-efficient slice ordering, Magn. Reson. Med. 54 (3) (2005) 725–731, doi:10.1002/mrm.20602.

[37] S.C. Deoni, High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI), J. Magn. Reson. Imaging 26 (4) (2007) 1106–1111, doi:10.1002/jmri.21130.

[38] N. Gelman, J.M. Gorell, P.B. Barker, R.M. Savage, E.M. Spickler, J.P. Windham, R.A. Knight, MR imaging of human brain at 3.0 T: preliminary report on transverse relaxation rates and relation to estimated iron content, Radiology 210 (3) (1999) 759–767, doi:10.1148/radiology.210.3.r99fe41759.

[39] J.P. Wansapura, S.K. Holland, R.S. Dunn, W.S. Ball Jr., NMR relaxation times in the human brain at 3.0 tesla, J. Magn. Reson. Imaging 9 (4) (1999) 531–538, doi:10.1002/(SICI)1522-2586(199904)9:4<531::AID-JMRI4>3.0.CO;2-L.

[40] A. Hagiwara, M. Hori, J. Cohen-Adad, M. Nakazawa, Y. Suzuki, A. Kasahara, M. Horita, T. Haruyama, C. Andica, T. Maekawa, et al., Linearity, bias, intrascanner repeatability, and interscanner reproducibility of quantitative multidynamic multiecho sequence for rapid simultaneous relaxometry at 3 T: a validation study with a standardized phantom and healthy controls, Invest. Radiol. 54 (1) (2019) 39–47, doi:10.1097/RLI.0000000000000510.

[41] G.J. Stanisz, E.E. Odrobina, J. Pun, M. Escaravage, S.J. Graham, M.J. Bronskill, R.M. Henkelman, T1, T2 relaxation and magnetization transfer in tissue at 3T, Magn. Reson. Med. 54 (3) (2005) 507–512, doi:10.1002/mrm.20605.

[42] G. Liberman, Y. Louzoun, D. Ben Bashat, T1 mapping using variable flip angle SPGR data with flip angle correction, J. Magn. Reson. Imaging 40 (1) (2014) 171–180, doi:10.1002/jmri.24373.

[43] S.C. Deoni, H.A. Ward, T.M. Peters, B.K. Rutt, Rapid T2 estimation with phase-cycled variable nutation steady-state free precession, Magn. Reson. Med. 52 (2) (2004) 435–439, doi:10.1002/mrm.20159.

[44] S.K. Piechnik, J. Evans, L. Bary, R.G. Wise, P. Jezzard, Functional changes in CSF volume estimated using measurement of water T2 relaxation, Magn. Reson. Med. 61 (3) (2009) 579–586, doi:10.1002/mrm.21897.

[45] A. Hagiwara, Y. Otsuka, M. Hori, Y. Tachibana, K. Yokoyama, S. Fujita, C. Andica, K. Kamagata, R. Irie, S. Koshino, et al., Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation, Am. J. Neuroradiol. 40 (2) (2019) 224–230, doi:10.3174/ajnr.A5927.

[46] K. Armanious, C. Jiang, M. Fischer, T. Küstner, T. Hepp, K. Nikolaou, S. Gatidis, B. Yang, MedGAN: medical image translation using GANs, Comput. Med. Imaging Graph. 79 (2020) 101684, doi:10.1016/j.compmedimag.2019.101684.

[47] T. Stöcker, K. Vahedipour, D. Pflugfelder, N.J. Shah, High-performance computing MRI simulations, Magn. Reson. Med. 64 (1) (2010) 186–193, doi:10.1002/mrm.22406.

[48] Z. Cao, S. Oh, C.T. Sica, J.M. McGarrity, T. Horan, W. Luo, C.M. Collins, Bloch-based MRI system simulator considering realistic electromagnetic fields for calculation of signal, noise, and specific absorption rate, Magn. Reson. Med. 72 (1) (2014) 237–247, doi:10.1002/mrm.24907.

[49] W.P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, B.M. Tsui, 4D XCAT phantom for multimodality imaging research, Med. Phys. 37 (9) (2010) 4902–4915, doi:10.1118/1.3480985.