



## Deep Learning Based Multi-Modal Fusion for Fast MR Reconstruction

**Lei Xiang,**

Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

**Yong Chen,**

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

**Weitang Chang,**

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

**Yiqiang Zhan,**

Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

**Weili Lin,**

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

**Qian Wang,**

Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China. wang.qian@sjtu.edu.cn

**Dinggang Shen, and**

Department of Radiology and BRIC, UNC-Chapel Hill, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic of Korea dgshen@med.unc.edu

**Fellow, IEEE**

### Abstract

T1-weighted image (T1WI) and T2-weighted image (T2WI) are the two routinely acquired magnetic resonance (MR) modalities that can provide complementary information for clinical and research usages. However, the relatively long acquisition time makes the acquired image vulnerable to motion artifacts. To speed up the imaging process, various algorithms have been proposed to reconstruct high-quality images from under-sampled  $k$ -space data. However, most of the existing algorithms only rely on mono-modality acquisition for the image reconstruction. In this paper, we propose to combine complementary MR acquisitions (i.e., T1WI and under-sampled T2WI particularly) to reconstruct the high-quality image (i.e., corresponding to the fully-sampled T2WI). To the best of our knowledge, this is the first work to fuse multi-modal MR acquisitions through deep learning to speed up the reconstruction of a certain target image. Specifically, we

present a novel deep learning approach, namely Dense-Unet, to accomplish the reconstruction task. The proposed Dense-Unet requires fewer parameters and less computation, while achieving promising performance. Our results have shown that Dense-Unet can reconstruct a 3D T2WI volume in less than 10 seconds with an under-sampling rate of 8 for the  $k$ -space and negligible aliasing artifacts or signal-noise-ratio (SNR) loss. Experiments also demonstrate excellent transferring capability of Dense-Unet when applied to the datasets acquired by different MR scanners. The above results imply great potential of our method in many clinical scenarios.

## Keywords

Deep learning; Dense Block; Multi-Model Fusion; Fast MR Reconstruction

## I. INTRODUCTION

Magnetic resonance (MR) imaging is widely applied for numerous clinical applications, as it can provide non-invasive, reproducible and high-quality measurements of both anatomical and functional information that are essentially important for disease diagnosis and treatment guidance. MR is better suited to reveal contrasts in different soft tissues than many other medical imaging modalities. It also avoids exposing patients to harmful ionizing radiation, thus bringing in great safety. However, MR acquisition usually requires sampling of the full  $k$ -space for encoding the spatial-frequency information. This leads to a relatively slow acquisition process if no acceleration is adopted. During the slow acquisition, patient movement or physiological motion, e.g., respiration, cardiac motion, and blood flow, can cause significant artifacts in the MR images [1]. Long scan time also increases the healthcare cost to the patient and limits the availability of MR scanners [2].

Attempts to accelerate MR acquisition could date back to the late 1970s, even before MR was widely applied for clinical purpose. The  $k$ -space data that encodes the spatial-frequency information are commonly acquired line-by-line. The acquisition time for a given sequence thus depends on the number of the sampled lines in the  $k$ -space. Many methods focus on the reduction of the  $k$ -space sampling rate, i.e., by under-sampling the  $k$ -space. These approaches rely on the intrinsic redundancy in the  $k$ -space, where individually sampled points do not arise from distinct spatial locations [3]. There are two major acceleration techniques, i.e., 1) parallel imaging (PI) [4], [5] and 2) compressed sensing (CS) [2], [3], [6], [7]. PI sequences under-sample the  $k$ -space and reconstruct the skipped data points using the coil sensitivity profiles across multiple-channels. However, an in-plane acceleration rate higher than 2 may cause artifacts and substantially reduce signal-to-noise ratio (SNR) in clinical practice [8]. On the other hand, CS aims at reconstructing an MR image from a small number of sparsely selected data points in the  $k$ -space. The success of CS in MR reconstruction originates from two facts: (a) the medical image is naturally compressible by sparse coding in an appropriately transformed domain [6]; (b) the sequence acquires data points in the transformed spatial-frequency domain, instead of the image domain.

Recent deep learning techniques can also be applied to accelerate MR acquisition. Deep convolutional neural networks (CNNs) have been widely applied in various intelligent tasks,

including image classification [9], [10], [11], object detection [12], [13], image segmentation [14], [15], [16] and image translation [17]. CNNs also have the capability of addressing inverse problems such as image super-resolution (SR) [17], denoising [18], depth map prediction [19], and medical image synthesis [20], [21], [22], [23]. Particularly in medical image reconstruction, CNNs have shown promising capability for handling incompletely sampled  $k$ -space data [24], [25], [26], [27]. Wang et al. [28] proposed to train a CNN to model the mapping between the MR images obtained from zero-filled under-sampled and fully-sampled  $k$ -space data, respectively. The deep learning based reconstruction result can be used either as an initialization or a regularization in the classical CS reconstruction process. Lee et al. [27] further introduced a deep multi-scale residual learning algorithm to reconstruct the under-sampled MR data by formulating CS as a residual regression problem. Sun et al. [24] proposed a deep architecture (ADMM-Net) which was defined over a dataflow graph. Its parameters, e.g., image transforms, shrinkage functions, etc., are trained in the end-to-end way. Schlemper et al. [26] developed a deep cascade of CNNs to reconstruct the aggressive Cartesian under-sampled MR image. When the frames of the sequences were reconstructed jointly, they demonstrated to learn the spatiotemporal correlation efficiently by leveraging the convolution and data sharing layers together. Quan et al. [29] proposed a novel fully-residual convolutional autoencoder and generative adversarial network (RefineGAN) with consistency loss for fast and accurate CS-MR reconstruction. In addition, they leveraged a chained network to further improve the quality of the reconstructed image.

More recently, several studies have been performed to integrate advanced deep neural network architectures, strategies, and loss designs to address the MRI reconstruction problems [30], [31], [32], [33], [34]. Schlemper et al. [30] proposed a novel cascaded CNN-based CS technique and a stochastic variation for Diffusion Tensor Cardiac Magnetic Resonance (DT-CMR) reconstruction. Yang et al. [31] designed a novel conditional generative adversarial network based model (DAGAN) to reconstruct CS-MRI and proposed a refinement learning method to further improve the reconstruction performance. Seitzer et al. [32] combined adversarial loss, perceptual loss and mean squared error (MSE) loss together to promote the visual quality of CS-MRI reconstruction. Qin et al. [33] developed a novel convolutional recurrent neural network (CRNN) architecture to simultaneously exploit the dependency of the temporal sequences as well as the iterative nature of traditional optimization algorithms for high-quality cardiac MR image reconstruction.

Although deep learning methods have shown promising results for reducing reconstruction time while maintaining superior image quality, most methods reported in the literature focus on reconstructing high-quality MR image (i.e., corresponding to the fully-sampled  $k$ -space data) by just using the under-sampled data of the same modality. Most of them have not explored the highly coupled relationship between different MR sequences to accelerate image reconstruction, i.e., through multi-modal fusion.

In clinical routines, T1WI and T2WI images are two basic MR sequences for assessing anatomical structures and pathologies, respectively. The two modalities are closely related with each other. T1WI is useful for identifying fatty tissue, characterizing focal liver lesions and in general for obtaining morphological information, as well as for post-contrast imaging.

T2WI is useful for detecting edema and inflammation, revealing brain white-matter lesions and assessing zonal anatomy in the prostate and uterus, etc. Examples of the co-registered T1WI and T2WI data from the same subject are shown in Fig. 1. Two white-matter lesions are highlighted by the red circle and the green box, respectively. The two sequences of T1WI and T2WI in general provide complementary information to reveal the anatomical details of the patient, although they have very different image appearance and contrast for corresponding anatomies. For example, the white matter in T1WI is of high intensity, while it appears dark in T2WI; whereas the gray matter in T1WI is of low intensity, while it appears bright in T2WI.

The acquisition of T2WI is usually slower than T1WI due to the relatively longer repetition time (TR) and echo time (TE) of T2WI. To this end, one may consider under-sampling the  $k$ -space for faster T2WI, which, however, may reduce the imaging quality at the same time. In Fig. 1, the example of the 1/8 under-sampled T2WI (i.e., by reducing the sampling rate to 1/8 of the case of fully sampling the  $k$ -space) is provided. Although the lesions can still be roughly observed, the boundaries are not clear and many details are missing. In addition, the quality of the non-lesion areas, as indicated by the blue arrows, is also reduced in the 1/8 under-sampled T2 image. Therefore, we argue that it is impractical to speed up the acquisition of T2WI by simply under-sampling the  $k$ -space with such a high reduction factor.

In this paper, we propose a deep learning solution to fuse T1WI and a highly under-sampled T2WI to reconstruct a high-quality T2WI. Our method can (1) leverage the highly coupled relation between T1WI and T2WI and (2) utilize the unique cues in the under-sampled T2WI to reconstruct the high-quality image corresponding to the fully-sampled T2WI. Particularly, we argue that the appearance of different tissues is highly related, though diverse, in T1WI and T2WI especially in non-lesion areas. In this way, a nonlinear mapping modeled by deep learning can bridge the two different image modalities. Meanwhile, even though a lesion might not be easily observable in T1WI, the cues in under-sampled T2WI allow us to reconstruct it at high quality after integrating the information from the fully-sampled T1WI.

To this end, we adapt the Unet architecture [14] to fuse the two modalities for the reconstruction of T2WI. First, we concatenate the corresponding T1WI and under-sampled T2WI as the input. Next, in the proposed Dense-Unet that consists of a contracting (or encoding) path and a symmetric expanding (decoding) path, we introduce the dense block to significantly boost the reconstruction quality of the target fully-sampled T2WI. The dense blocks result in fewer network parameters, making our computation much easier. Our experimental results suggest that we could accelerate the imaging process by under-sampling the  $k$ -space at the rate of 8 for T2WI, with negligible aliasing artifacts and SNR loss in the reconstructed T2 images. Moreover, we find that the deep network trained from the data of a Siemens Verio 3T scanner can be well applied to the data acquired by a Philips Ingenia 3T scanner, implying a high transferring and generalization capability of our method.

Although there are several studies attempting to estimate T2WI from T1WI in the literature, our work attains fast T2WI reconstruction by fusing multi-modal T1WI with under-sampled T2WI. A similar work can be found in [34], which also tries to improve the quality of down-sampled images with the help of high resolution images of different contrast. For just using T1WI to predict T2WI, Alkan et al. [35] proposed a nine-layer 2D CNN using T1WI combined with tissue mask to generate T2WI. Vemulapalli et al. [36] generated T1WI from T2WI by unsupervised cross-modal synthesis. However, our work differs from these works in terms of the following contributions.

- We have demonstrated that our proposed method could achieve 8x acceleration rate while still preserves high reconstruction quality. It only takes 9.5s to finish the reconstruction of a 3D T2 image volume by fusing T1WI and under-sampled T2WI acquisitions.
- We propose a novel Dense-Unet architecture with the dense blocks in both encoding and decoding paths. The dense blocks dramatically reduce the number of the network parameters to 1/3, while still achieving superior reconstruction quality of T2WI.
- We demonstrate the possibility of transferring the networks trained with different datasets. In particular, we train with the data from a Siemens Verio 3T and test with data from a Philips Ingenia 3T scanner. The quality of the reconstruction is satisfactory, which demonstrates the good transferring and generalization capability of our proposed method.

A preliminary version of this work has been presented at a conference [37]. Herein, we (i) demonstrate transferability of proposed method across different scanners, (ii) evaluate and further analyze the reconstruction performance of using different under-sampling masks, (iii) provide more detailed comparison of our proposed method with base Unet both qualitatively and quantitatively, and (iv) include additional discussions that are not in the conference publication.

## II. METHODS

We provide details of our proposed Dense-Unet in this section, which is capable of reconstructing high-quality T2WI by fusing T1WI and under-sampled T2WI acquisitions. The architecture of our proposed Dense-Unet is illustrated in Fig. 2.

### A. Objective Function

We can denote the under-sampled T2  $k$ -space data as

$$f_{T2} = MFy_{T2} \quad (1)$$

where  $M$  is the mask to under-sample the  $k$ -space,  $F$  represents the full Fourier encoding matrix complying with  $F^H F = I$  ( $F^H$  for the Hermitian transpose operation), and  $y_{T2}$  is the T2 image corresponding to fully-sampled  $k$ -space data (fully-sampled T2 image). Therefore,

$F_{yT2}$  denotes the fully sampled  $k$ -space data. With the under-sampled  $k$ -space data  $f_{T2}$ , we can apply zero-filling to the  $k$ -space and get the under-sampled T2 image by

$$x_{T2} = F^H f_{T2} \quad (2)$$

Our goal in this work is to reconstruct the fully-sampled T2 image ( $f_{T2}$ ) when only the under-sampled  $x_{T2}$  or  $f_{T2}$  is available. Specifically, we adopt a deep CNN architecture to accomplish this task. Given pair data of  $x_{T2}$  and  $y_{T2}$ , we train the network to minimize the following loss function in a supervised way:

$$\underbrace{\arg \min}_{\theta} \left\{ \frac{1}{2N} \sum_{n=1}^N \|C(x_{T2,n}; \theta) - y_{T2,n}\|_2^2 \right\} \quad (3)$$

In the above,  $C(\cdot; \cdot)$  is the desired mapping function with the network parameters  $\theta = \{(W_1, B_1), \dots, (W_L, B_L)\}$ .  $N$  is the total number of the paired images for training, and  $L$  is the maximal layer depth of the network.

T1WI provides critical guidance to the reconstruction of T2WI in this work, which has been motivated early and will be verified in subsequent experiments. To this end, we need to fuse the T1WI information to better reconstruct T2WI. The corresponding T1WI is thus concatenated with the under-sampled T2WI, and then input to the network. Accordingly, the mapping function of the network needs to accommodate the multi-modal input by following

$$\underbrace{\arg \min}_{\theta} \left\{ \frac{1}{2N} \sum_{n=1}^N \|C([x_{T2,n}, y_{T1,n}]; \theta) - y_{T2,n}\|_2^2 \right\} \quad (4)$$

where  $y_{T1,n}$  is the fully-sampled T1 image

## B. The Dense-Unet Architecture

Our proposed deep neural network can handle multi-modal input, which is concatenated from fully-sampled T1WI and under-sampled T2WI. Specifically, we concatenate  $m$  consecutive under-sampled T2 slices and  $m$  consecutive T1 slices in the same positions. The concatenated  $2m$  slices are input to the network, which outputs corresponding T2 slices. In our implementation, we set  $m = 3$  for synthesizing every 3 axial slices in each training and testing task. The performance is mostly stable by increasing  $m$  according to our experiments.

The Dense-Unet, whose detailed architecture is shown in Fig. 2(a), contains a contracting path and an expanding path. The feature map sizes decrease along the contracting path by pooling, and then increase in the expanding path by deconvolution. There are four basic components in the Dense-Unet architecture, i.e., pre-feature extraction layer, dense block, transition layer and reconstruction layer. The four components are marked with different indices in the figure. An example of the dense block is shown in Fig. 2(b).

**Pre-feature extraction layer:** Dense-Unet first extracts feature maps from the concatenated under-sampled T2WI and T1WI using a convolutional layer. These feature maps are forwarded to the latter dense blocks for further mapping. Denoting the concatenated input as  $(x_{T2}, y_{T1})$ , we can compute the output of the first layer as

$$F_1 = \sigma(W_1 * [x_{T2}, y_{T1}] + B_1) \quad (5)$$

where  $W_1$  and  $B_1$  represent the kernels associated with the first convolutional layer, and  $*$  denotes the convolution operator.

**Dense block:** Dense connectivity has been proposed in [10] to improve the feature flow across layers. We adopt the strategy in our model so that we can effectively increase the depth of the whole network without being trapped in optimization. Moreover, the dense block requires substantially fewer parameters and less computation, which makes the model efficient to train. Fig. 2(b) illustrates the layout of the dense block. Consequently, the  $l$ -th layer receives the feature maps of all preceding layers,  $[z_0, \dots, z_{l-1}]$ , as the input

$$z_l = H_l([z_0, \dots, z_{l-1}]) \quad (6)$$

where  $[z_0, \dots, z_{l-1}]$  refers to the concatenation of the feature maps coming from layers  $0, \dots, l-1$ .  $H_l(*)$  is defined as a composite function of three consecutive operations: batch normalization (BN) [38], followed by a rectified linear unit (ReLU) [39], and a  $3 \times 3$  convolution (Conv). The hyper-parameters for the dense block are the growth rate (GR) and the number of the convolutional layers (NC). Fig. 2(b) gives an example of the dense block with GR=16 and NC=5.

**Transition layer:** We refer to the layer behind the dense block as the transition layer. In the contracting path, it consists of BN,  $1 \times 1$  Conv, and  $2 \times 2$  average pooling. On the expanding path, it consists of BN and deconvolution (filter number: 64, size:  $4 \times 4$ , stride: 2). The transition layer is introduced to fix the number of the feature maps to 64. The hyper-parameters in the preceding dense block may alter the number of the feature maps. With the transition layer, we can flexibly adopt the dense blocks in both the contracting and the expanding paths.

**Reconstruction layer:** The proposed Dense-Unet ends with the reconstruction layer that yields the fully-sampled T2WI from the feature maps generated by the last dense block. The reconstruction can be attained by a single convolutional layer by following

$$y_{\tilde{T}2} = W_R * F_D + B_R \quad (7)$$

Here,  $F_D$  is the feature maps outputted by the last dense block, and  $y_{\tilde{T}2}$  is the T2WI estimation by the reconstruction layer. Note that there is no activation function employed in



the reconstruction layer. We use the mean squared error (MSE) between the estimation of fully-sampled T2 image and the ground-truth fully-sampled T2 image as the loss function, which supervises the training of the whole network.

### III. EXPERIMENTAL RESULTS

First, we introduce the datasets used in the experiments and present the experimental settings (Sections 3.1). After that, we investigate the necessity of fusing the multi-modal T1 information for the reconstruction of T2WI (Section 3.2). Next, we explore the influence of different hyper-parameter settings of Dense-Unet upon the reconstruction performance (Section 3.3). Then, we compare Dense-Unet with the existing Unet and other state-of-the-art methods to demonstrate the effectiveness of our proposed method (Section 3.4). After that, we explore the effect of using different under-sampling masks on the performance of reconstruction (Section 3.5). Finally, we demonstrate that the Dense-Unet trained with a certain dataset can be transferred and applied to reconstruct another dataset acquired from a different MR scanner (Section 3.6).

#### A. Data and Experimental Setting

We utilize the imaging data from MICCAI Multiple Sclerosis (MS) Segmentation Challenge 2016 [40] for the demonstration of our proposed Dense-Unet. There are two datasets, each containing 5 subjects of paired T1WI and T2WI. **Dataset 1** comes from a Philips Ingenia 3T scanner. The voxel size is  $0.7 \times 0.74 \times 0.74 \text{ mm}^3$ . **Dataset 2** comes from a Siemens Verio 3T scanner. The voxel size is  $1.1 \times 0.5 \times 0.5 \text{ mm}^3$ . Multiple pre-processing steps are applied on all subjects in both datasets, including: 1) denoising with the non-local means algorithm [41]; 2) rigid registration [42]; 3) brain extraction using the volBrain platform [43] from T1WI and then applied to T2WI with sinc interpolation; 4) bias correction using the N4 algorithm [44]; 5) intensity normalization to the range [0,1] by dividing the maximal intensity value. Note that Steps 1–4 above are processed by the challenge organizers. The final size of each image is cropped to  $336 \times 336 \times 261 \text{ mm}^3$  in our experiments.

To prepare the data of under-sampled T2WI, we adopt a center mask (c.f. Fig. 7) to under-sample the  $k$ -space. Note that the center part of the  $k$ -space (and also covered by the center mask) accounts for a major contrast source of the anatomical structures in the reconstructed T2 image. The outer area of the  $k$ -space, on the other hand, provides high-frequency spatial information. In our proposed method, we fuse T1WI and under-sampled T2WI as the input to the network. We expect that all spatial information in the fully-sampled T1 image can help estimate anatomical structures in T2WI, where central  $k$ -space in T2WI will provide essential information for image contrast. Particularly, we follow Eq. (1) to simulate the under-sampled T2 image given each fully-sampled T2WI acquisition. Note that there are alternative ways to design the mask, which will be discussed later.

We then use PyTorch [45] to implement the proposed Dense-Unet. In the training phase, we use consecutive 2D slices ( $m = 3$ ) to train the deep network. We extract 2D slices from the 3D volumes of fully-sampled T1WI and under-sampled T2WI as the input, while the corresponding slices from the fully-sampled T2WI are used as ground truth. In this way, each training subject can contribute 200 training samples in axial slices, while the slices



containing background only are excluded from training. Data augmentation of horizontal flipping is also applied. We adopt Adam optimization [46] with a momentum of 0.9 and perform 100 epochs in the training stage. The batch size is set to 4 and the initial learning rate is set to 0.0001, which is divided by 10 after 50 epochs. We use zero-padding during every convolutional layer to ensure that the size of the output is the same as the size of the input. All the experiments are conducted in a desktop with an Intel Core i7 4.00GHz CPU, 32 GB RAM, and an NVIDIA GeForce GTX Titan X GPU.

The leave-one-out cross-validation strategy is employed in the experiment. To quantitatively evaluate the reconstruction performance, we use the standard metrics of mean absolute error (MAE) and peak signal-noise ratio (PSNR)

$$MAE = \frac{|y_{T2} - y_{\tilde{T}2}|}{V} \quad (8)$$

$$PSNR = 10 \ln \left( \frac{VD^2}{\|y_{T2} - y_{\tilde{T}2}\|_2^2} \right) \quad (9)$$

Here,  $V$  is the number of voxels in the image,  $y_{T2}$  is the ground-truth T2 image,  $y_{\tilde{T}2}$  is the reconstructed T2WI, and  $D$  is the intensity range of image  $y_{T2}$ . The PSNR and MAE measures encode the difference between the reconstructed T2WI and the ground truth. In general, the higher PSNR and lower MAE indicate better perceptive quality of the reconstructed T2WI.

## B. Necessity of Fusing T1WI

To demonstrate the effectiveness of integrating T1WI data for the reconstruction of T2WI, we compare the performances of several different cases performed on **Dataset 1** and report the average PSNRs/MAEs in Table I. The under-sampling rate of the  $k$ -space is set to 8 here. The cases under comparisons include (1) using under-sampled T2WI to reconstruct fully-sampled T2WI (Reconstructed T2 with 1/8 T2) and (2) using the combination of T1WI and under-sampled T2WI to reconstruct fully-sampled T2WI (Reconstructed T2 with T1 and 1/8 T2). For better comparison, in Table I we provide the PSNR/MAE scores between the fully-sampled T2WI and under-sampled T2WI (1/8 T2) before deep learning based reconstruction. Meanwhile, one may also consider synthesizing T2WI by using T2WI input only (Reconstructed T2 with T1). Note that the same network as in Fig. 2 was used when dealing with a single image input for fairness (i.e., Reconstructed T2 with 1/8 T2 and Reconstructed T2 with T1).

As shown in Table I, better performance is observed with Reconstructed T2 with T1 and 1/8 T2 as compared to Reconstructed T2 with T1 and Reconstructed T2 with 1/8 T2. In particular, the average PSNR score for Reconstructed T2 with T1 and 1/8 T2 is 36.9dB,

comparing to 30.6dB for Reconstructed T2 with T1 and 33.9dB for Reconstructed T2 with 1/8 T2. On the other hand, minimum improvement in image quality is observed with Reconstructed T2 with 1/8 T2 (PSNR: 33.9dB) as compared to 1/8 T2 (PSNR: 32.5dB), suggesting difficulty in enhancing the quality of T2WI with deep learning under such a high reduction factor. However, with additional information from T1WI, the image quality (Reconstructed T2 with T1 and 1/8 T2, PSNR: 36.9dB) is largely improved, which demonstrates the importance to fuse multi-modal inputs for the reconstruction of highly undersampled T2WI. The MAE scores are mostly consistent with the results measured by PSNR.

Representative images for all four reconstruction cases are shown in Fig. 3 for visual inspection. Compared to other methods, our proposed approach with Reconstructed T2 with T1 and 1/8 T2 provides the best image quality and preserves both image contrast and detailed tissue boundaries with respect to the ground-truth T2WI. Although the performance of Reconstructed T2 with T1 in PSNR/MAE is poor (c.f. Table I), major tissue and anatomies are correctly synthesized, e.g., white matter and gray matter in the red circle. However, the lesion in the green box is largely missing if using only T1 input for cross-modal synthesis, as reflected by the large differences in the error map with respect to the ground truth. This indicates the difficulty of reconstructing T2WI based on T1WI alone by image synthesis, as the unique information in the  $k$ -space center of T2WI (though under-sampled and low-quality) is essential for preserving image contrast.

On the contrary, referring to Reconstructed T2 with 1/8 T2, we find that the white-matter lesion (green box) is clearly observable, while its interface with the surrounding tissues becomes blurry. Moreover, in the red circle, the contrast appears worse compared to the ground truth. Even though deep learning is adopted here to enhance image quality, we argue that the missing data points in the masked-out  $k$ -space make it difficult to fully reconstruct T2WI from under-sampled  $k$ -space data. To this end, we need to take advantages from complementary T1WI and under-sampled T2WI. While T1WI provides detailed anatomical information that both modalities share, the unique cues in under-sampled T2WI are essential for the proposed Dense-Unet. In general, Reconstructed T2 with T1 and 1/8 T2 yields the most satisfactory reconstruction result with high perceptive quality.

### C. Parameter Setting in Dense-Unet

There are two hyper-parameters in the dense block, i.e., the growth rate (GR) and the number of convolutional layers (NC). In this experiment, we explore the effect of different hyper-parameter settings on the reconstruction quality of T2WI by using Dataset 1. Concerning the physical limit of the GPU memory, we have tested five pairs of parameters. The results as reported in Table II are also associated with three different rates to under-sample the  $k$ -space. In this way, we can evaluate the influence of the parameters regarding the acceleration factor in the reconstruction of fully-sampled T2WI.

As shown in Table II, Dense-Unet with GR = 16 and NC = 5 yields the best reconstruction performance, regardless of the under-sampling rates (i.e., 1/4, 1/8, and 1/16). Reduced image quality (decreased PSNR and increased MAE) is observed with higher under-sampling rates, which is expected. Fig.4 presents examples for visual inspection, which are derived from the

optimal hyper-parameters and corresponding to different under-sampling rates. Similar image quality and error maps are noticed for Reconstructed T2 with T1 and 1/8 T2 and Reconstructed T2 with T1 and 1/4 T2. However, substantial errors are observed with the under-sampling rate of 1/16, especially as indicated in the green box of white-matter lesion. To this end, we adopt the recommended under-sampling rate as 1/8 in this work, and argue that our method can reach 8x acceleration factor for T2WI.

#### D. Comparison with Unet

In this section, we conduct comprehensive comparisons between our proposed Dense-Unet and the standard Unet, in order to demonstrate the advantage of the dense block. For Unet, it also has two down-sampling layers and two up-sampling layers. Similar to the original Unet, the numbers of convolution feature map in each stage are 64, 128 and 256, respectively. And it has 2 convolution layers in each stage. First, we compare the training/testing loss convergence of Unet and Dense-Unet, respectively (Fig. 5). In the training stage, the blue curve for Dense-Unet converges faster than the green curve for Unet, and arrives at lower loss in the final iteration. In the testing stage, Dense-Unet also produces better performance than Unet, which is reflected by the lower loss calculated from test samples. Note that there are ~2.2 million parameters in U-Net, but only ~0.4 million in Dense-Unet.

Quantitative evaluation in PSNR and MAE for both Unet and Dense-Unet is summarized in Table III. For three different under-sampling rates, Dense-Unet consistently outperforms Unet, suggesting the contribution of the proposed dense block in the multi-modal reconstruction task. Fig.6 gives a visual comparison of the reconstruction results by Unet and Dense-Unet, respectively. The under-sample ratio studied in Fig.6 is 1/8 and all three views are presented. We can see that with T1WI and under-sampled T2WI as input, both Unet and Dense-Unet can reconstruct the details of T2WI, including the lesion. However, the Dense-Unet is superior in that it provides a clearer boundary and less error especially in the ventricular area.

#### E. Effect of under-sampling masks for T2WI

A center mask is used to under-sample the  $k$ -space of T2WI. In this section, we evaluate the choices of different under-sampling masks and compare the reconstructed T2 images. With a constant under-sample ratio of 1/8, we particularly test three different masks, including Center Mask, Equidistance Mask and Gaussian Mask. Center Mask only extracts information from the center  $k$ -space. Equidistance Mask extracts half of the  $k$ -space lines (1/16 of the whole  $k$ -space) from the center, and the rest half lines are uniformly sampled at both sides. For Gaussian Mask, 3/32 of  $k$ -space lines are included in the center and the rest are sampled based on the Gaussian distribution. All three masks are shown in Fig.7. Note that more information is always extracted from the center of  $k$ -space for all three masks, as the data in the center is critical to determine the image contrast in T2WI. Results of using different masks are then summarized in Table IV. We observe that Center Mask provides the best performance out of the three cases. This further suggests that, when the same amount of  $k$ -space data is utilized, sampling at the  $k$ -space center of T2WI will provide the best results with our proposed method.

## F. Transferring across Scanners

In this section, we explore the robustness of our proposed method on different datasets that were acquired from different scanner makers. We investigate the way to transfer the network across the datasets of different scanners. Specifically, as we have two datasets in this work, we use Dataset 2 (from a Siemens Verio 3T scanner) to train the network and then apply it to Dataset 1 for testing (from a Philips Ingenia 3T scanner). The average PSNRs and MAEs with respect to different under-sampling rates are reported in Table V. For comparison, the average PSNRs and MAEs that are tested on Dataset 1 using the leave-one-out strategy are also listed in Table V. We refer Dense-Unet (D1-D1) to the case when our proposed method is trained on Dataset 1 and tested on Dataset 1 too. Dense-Unet (D2-D1) then indicates the case that the network trained from Dataset 2 is transferred to Dataset 1.

From the table, we observe that the reconstruction performances are mostly comparable between Dense-Unet (D1-D1) and Dense-Unet (D2-D1). For example, with T1 and 1/8 under-sampled T2, the transferred network reaches a PSNR of 36.7dB, which is just slightly lower than the case trained and tested with the same dataset (36.9dB). Similar observations can be found for different under-sampling rates, indicating high robustness of our method when it is transferred across different datasets.

## IV. DISCUSSION

In this study, we have demonstrated that, with the complementary information from T1WI, the under-sampled T2WI can be reconstructed with superior image quality compared to the images reconstructed using only under-sampled T2WI or only fully-sampled T1WI. Note that the techniques of image enhancement (i.e., corresponding to Reconstructed T2 with 1/8 T2) and cross-modal synthesis (Reconstructed T2 with T1) have drawn a lot of attention in medical image analysis recently. However, it is still necessary to fuse multi-modal input for a better reconstruction as reflected by our experimental results (Reconstructed T2 with T1 and 1/8 T2). As in Fig.3, we have demonstrated that under-sampled T2 is critical to provide cues of white-matter lesion, which is barely observable in T1WI.

While the importance of T2WI for diagnosis (e.g., toward the lesion as in Fig.3) is evident, one may argue that for healthy subjects, the reconstructed T2WI from T1WI alone is enough without the need of under-sampled T2WI [35]. To this end, we present an example slice without any lesion in Fig.8. The first row is the reconstruction results using our proposed Dense-Unet methods, and the second row contains error maps generated by comparison with the ground-truth T2. For visual inspection, one may notice that Reconstructed T2 with T1 has a higher systematic error compared to Reconstructed T2 with T1 and 1/8 T2, especially in the ventricles highlighted by the green box. Meanwhile, Reconstructed T2 with 1/8 T2 has no such bias, though it suffers from blurry boundary in the image. Therefore, we argue that the cross-modal synthesis task (from T1 to T2) is challenging. Whereas, it becomes much easier to reconstruct high-quality T2WI with multi-modal input.

MR images reconstructed from  $k$ -space contain complex-value data. The phase of MR images often provides valuable information as the magnitude image, particularly when performing image reconstruction from under-sampled dataset. In this study, all our

experiments were performed using data from the 2016 MICCAI Multiple Sclerosis (MS) Segmentation Challenge. The data provided by the challenge only contains magnitude information. Therefore, we are unable to evaluate the reconstruction results with the phase information. In the future, we will collect real data containing both the magnitude and phase to further evaluate the proposed reconstruction method. It is straightforward to extend the current Dense-Unet architecture to handle complex-value data by including the phase data as a separate channel for the input network. In our case, we will have 12 channels in total for the input, i.e., 6 for magnitude channels (3 consecutive slices for T1WI and 3 for under-sampled T2WI) and 6 for phase channels (also 3 for T1WI and 3 for under-sampled T2WI). Similarly, the network can output the 6 channels corresponding to the magnitude and phase parts. This setting could combine the information from magnitude and phase parts simultaneously for the reconstruction, and further improvement in the reconstruction performance is expected.

The proposed framework that uses T1WI to help the reconstruction of T2WI can be easily applied to other MRI sequence reconstruction, such as Fluid Attenuated Inversion Recovery (Flair) image and Diffusion Weighted Imaging (DWI). In the future, we will explore the potential of using more modalities to reconstruct the target modality. Moreover, currently we use one fully sampled modality and an under-sampled modality to implement the reconstruction; we could extend the work by using two under-sampled modalities (i.e., one slightly under-sampled and one highly sampled) to complete the reconstruction of highly sampled image. Following this extension, we could improve our framework by outputting two reconstruction results corresponding to the two under-sampled inputs, resulting in a faster scanning process for multi-sequence acquisitions in practical clinical applications. These are all based on the observation that multi-modalities have high inherent corresponding relationship with each other, but they also have their own unique information.

## V. CONCLUSION

In this paper, we propose a novel Dense-Unet model to reconstruct the T2WI from the T1WI and under-sampled T2WI. Our approach of using T1WI makes the reconstruction of T2WI from 1/8 under-sample ratio in  $k$ -space possible and leads to the process being sped up by a factor of 8. The dense block, which requires substantially fewer parameters and less computation, is integrated within the Unet architecture in our work. This proposed Dense-Unet could converge faster than that of the baseline model Unet and attain a lower loss point. This enables our model to further improve the quality of the reconstructed T2WI. Comprehensive experiments showed the superior performance of our method, including the greater perceptive quality and the faster running speed. Moreover, the transferability across different datasets from different scanners further shows the superiority of our proposed method. The trained model can be directly applied to different scanners without any pre-training or refinement process. This work can greatly improve the acquisition efficiency and image quality in clinical settings. While all the experiments in this work are performed on simulated MR images, we will investigate the performance of our framework on real data in the future, with both magnitude and phase values to improve the reconstruction results gained from under-sampled images.

## Acknowledgment

This paper was supported in part by the National Natural Science Foundation of China under Grants 81471733 and 61471390, the program of China Scholarship Council, the National Key R&D Program of China under Grant 2017YFC0107602, the Science and Technology Commission of Shanghai Municipality under Grant 16410722400, 16511101103, and 17411953300.

## REFERENCES

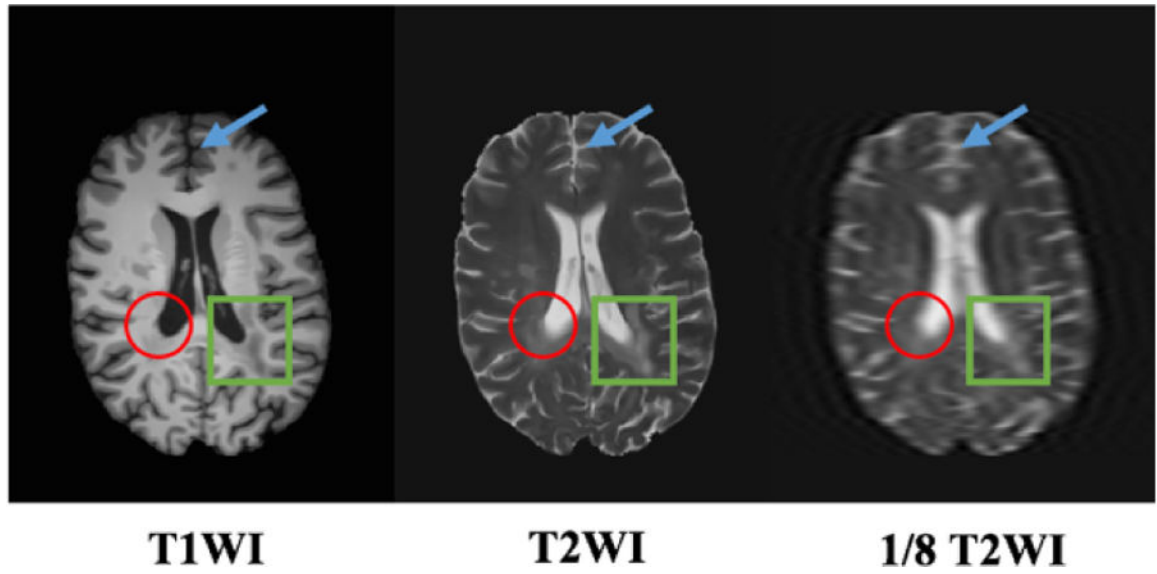
- [1]. Krupa K and Bekiesin'ska-Figatowska M, "Artifacts in magnetic resonance imaging," *Polish journal of radiology*, vol. 80, p. 93, 2015. [PubMed: 25745524]
- [2]. Jaspán ON, Fleysher R, and Lipton ML, "Compressed sensing mri: a review of the clinical literature," *The British journal of radiology*, vol. 88, no. 1056, p. 20150487, 2015. [PubMed: 26402216]
- [3]. Lustig M, Santos JM, Lee J-H, Donoho DL, and Pauly JM, "Application of compressed sensing for rapid mr imaging," *SPARS*, (Rennes, France), 2005.
- [4]. Pruessmann KP, Weiger M, Scheidegger MB, and Boesiger P, "Sense: sensitivity encoding for fast mri," *Magnetic resonance in medicine*, vol. 42, no. 5, pp. 952–962, 1999. [PubMed: 10542355]
- [5]. Sodickson DK and Manning WJ, "Simultaneous acquisition of spatial harmonics (smash): fast imaging with radiofrequency coil arrays," *Magnetic resonance in medicine*, vol. 38, no. 4, pp. 591–603, 1997. [PubMed: 9324327]
- [6]. Lustig M, Donoho DL, Santos JM, and Pauly JM, "Compressed sensing mri," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 72–82, 2008.
- [7]. Gamper U, Boesiger P, and Kozerke S, "Compressed sensing in dynamic mri," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 59, no. 2, pp. 365–373, 2008.
- [8]. Hutchinson M and Raff U, "Fast mri data acquisition using multiple detectors," *Magnetic resonance in Medicine*, vol. 6, no. 1, pp. 87–91, 1988. [PubMed: 3352509]
- [9]. Krizhevsky A, Sutskever I, and Hinton GE, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [10]. Huang G, Liu Z, Van Der Maaten L, and Weinberger KQ, "Densely connected convolutional networks," in *CVPR*, vol. 1, no. 2, 2017, p. 3.
- [11]. He K, Zhang X, Ren S, and Sun J, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [12]. Girshick R, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [13]. Ren S, He K, Girshick R, and Sun J, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.
- [14]. Ronneberger O, Fischer P, and Brox T, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [15]. Badrinarayanan V, Kendall A, and Cipolla R, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *arXiv preprint arXiv:1511.00561*, 2015.
- [16]. Zhou S, Nie D, Adeli E, Gao Y, Wang L, Yin J, and Shen D, "Fine-grained segmentation using hierarchical dilated neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 488–496.
- [17]. Dong C, Loy CC, He K, and Tang X, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2016. [PubMed: 26761735]
- [18]. Zhang K, Zuo W, Chen Y, Meng D, and Zhang L, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017. [PubMed: 28166495]



- [19]. Eigen D, Puhrsch C, and Fergus R, "Depth map prediction from a single image using a multi-scale deep network," in Advances in neural information processing systems, 2014, pp. 2366–2374.
- [20]. Xiang L, Qiao Y, Nie D, An L, Lin W, Wang Q, and Shen D, "Deep auto-context convolutional neural networks for standard-dose pet image estimation from low-dose pet/mri," Neurocomputing, vol. 267, pp. 406–416, 2017. [PubMed: 29217875]
- [21]. Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q, and Shen D, "Medical image synthesis with context-aware generative adversarial networks," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2017, pp. 417–425.
- [22]. Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, Wang Q, and Shen D, "Medical image synthesis with deep convolutional adversarial networks," IEEE Transactions on Biomedical Engineering, 2018.
- [23]. Xiang L, Wang Q, Nie D, Zhang L, Jin X, Qiao Y, and Shen D, "Deep embedding convolutional neural network for synthesizing ct image from t1-weighted mr image," Medical image analysis, vol. 47, pp. 31–44, 2018. [PubMed: 29674235]
- [24]. Sun J, Li H, Xu Z et al., "Deep admm-net for compressive sensing mri," in Advances in Neural Information Processing Systems, 2016, pp. 10–18.
- [25]. Yu S, Dong H, Yang G, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Firmin D et al., "Deep de-aliasing for fast compressive sensing mri," arXiv preprint arXiv:1705.07137, 2017.
- [26]. Schlemper J, Caballero J, Hajnal JV, Price AN, and Rueckert D, "A deep cascade of convolutional neural networks for dynamic mr image reconstruction," IEEE transactions on Medical Imaging, vol. 37, no. 2, pp. 491–503, 2018. [PubMed: 29035212]
- [27]. Lee D, Yoo J, and Ye JC, "Deep residual learning for compressed sensing mri," in Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on IEEE, 2017, pp. 15–18.
- [28]. Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D, and Liang D, "Accelerating magnetic resonance imaging via deep learning," in Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on IEEE, 2016, pp. 514–517.
- [29]. Quan TM, Nguyen-Duc T, and Jeong W-K, "Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss," IEEE transactions on medical imaging, vol. 37, no. 6, pp. 1488–1497, 2018. [PubMed: 29870376]
- [30]. Schlemper J, Yang G, Ferreira P, Scott A, McGill L-A, Khalique Z, Gorodezky M, Roehl M, Keegan J, Pennell D et al., "Stochastic deep compressive sensing for the reconstruction of diffusion tensor cardiac mri," arXiv preprint arXiv:1805.12064, 2018.
- [31]. Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge J, Keegan Y, Guo et al., "Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction," IEEE transactions on medical imaging, vol. 37, no. 6, pp. 1310–1321, 2018. [PubMed: 29870361]
- [32]. Seitzer M, Yang G, Schlemper J, Oktay O, Würfl T., Christlein V, Wong T, Mohiaddin R, Firmin D, Keegan J et al., "Adversarial and perceptual refinement for compressed sensing mri reconstruction," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2018, pp. 232–240.
- [33]. Qin C, Hajnal JV, Rueckert D, Schlemper J, Caballero J, and Price AN, "Convolutional recurrent neural networks for dynamic mr image reconstruction," IEEE transactions on medical imaging, 2018.
- [34]. Kim KH, Do W-J, and Park S-H, "Improving resolution of mr images with an adversarial network incorporating images with different contrast," Medical physics, 2018.
- [35]. Alkan C, Cocjin J, and Weitz A, "Magnetic resonance contrast prediction using deep learning"
- [36]. Vemulapalli R, Van Nguyen H, and Kevin Zhou S, "Unsupervised cross-modal synthesis of subject-specific scans," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 630–638.
- [37]. Xiang L, Chen Y, Chang W, Zhan Y, Lin W, Wang Q, and Shen D, "Ultra-fast t2-weighted mr reconstruction using complementary t1-weighted information," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2018, pp. 215–223.



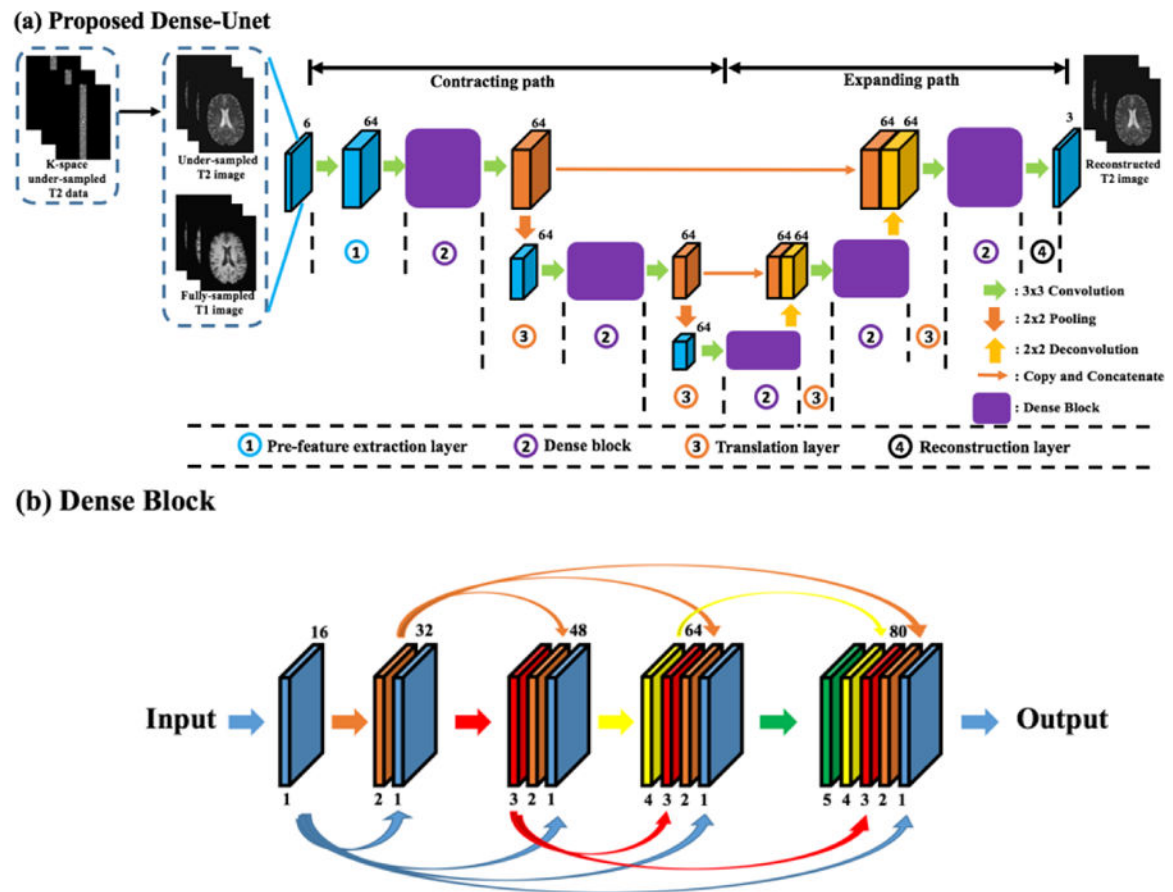
- [38]. Ioffe S and Szegedy C, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
- [39]. Glorot X, Bordes A, and Bengio Y, "Deep sparse rectifier neural networks," in Proceedings of the fourteenth international conference on artificial intelligence and statistics, 2011, pp. 315–323.
- [40]. Commowick O, Cervenansky F, and Ameli R, "Msseg challenge proceedings: Multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure," in MICCAI, 2016.
- [41]. Coupe P, Yger P, Prima S, Hellier P, Kervrann C, and Barillot C, "An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images," IEEE transactions on medical imaging, vol. 27, no. 4, pp. 425–441, 2008. [PubMed: 18390341]
- [42]. Commowick O, Wiest-Daessle N, and Prima S, "Block-matching strategies for rigid registration of multimodal medical images," in Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on IEEE, 2012, pp. 700–703.
- [43]. Manjón JV and Coupé P "volbrain: An online mri brain volumetry system," Frontiers in neuroinformatics, vol. 10, p. 30, 2016. [PubMed: 27512372]
- [44]. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, and Gee JC, "N4itk: improved n3 bias correction," IEEE transactions on medical imaging, vol. 29, no. 6, pp. 1310–1320, 2010. [PubMed: 20378467]
- [45]. Klein G, Kim Y, Deng Y, Senellart J, and Rush AM, "Opennmt: Open-source toolkit for neural machine translation," arXiv preprint arXiv:1701.02810, 2017.
- [46]. Kingma DP and Ba J, "Adam: A method for stochastic optimization," arXiv preprint arXiv: 1412.6980, 2014.



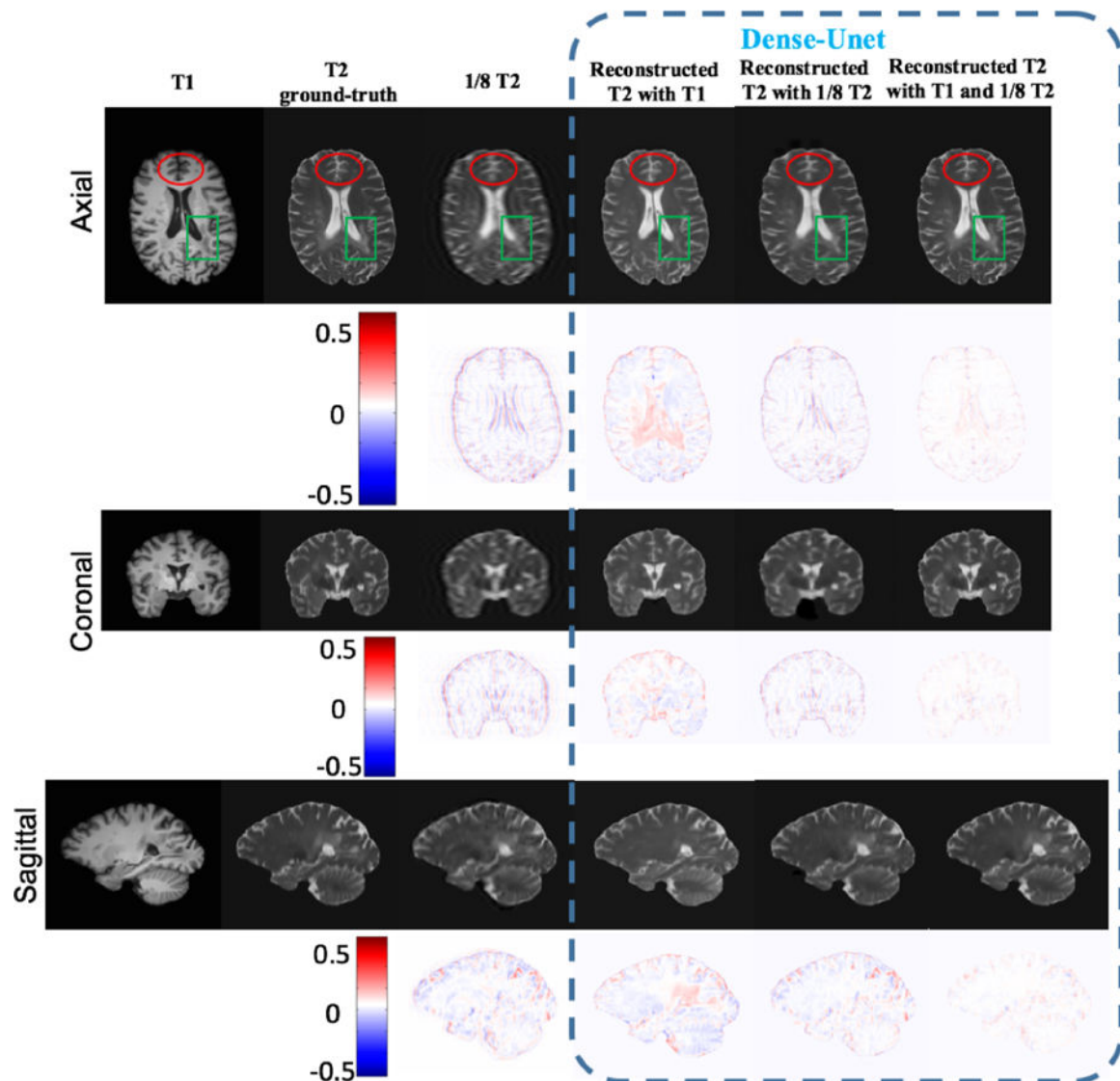
**Fig. 1.**

Examples of T1WI, T2WI and 1/8 under-sampled T2WI data from the same subject.

Multiple sclerosis lesions are marked by circles and boxes, respectively

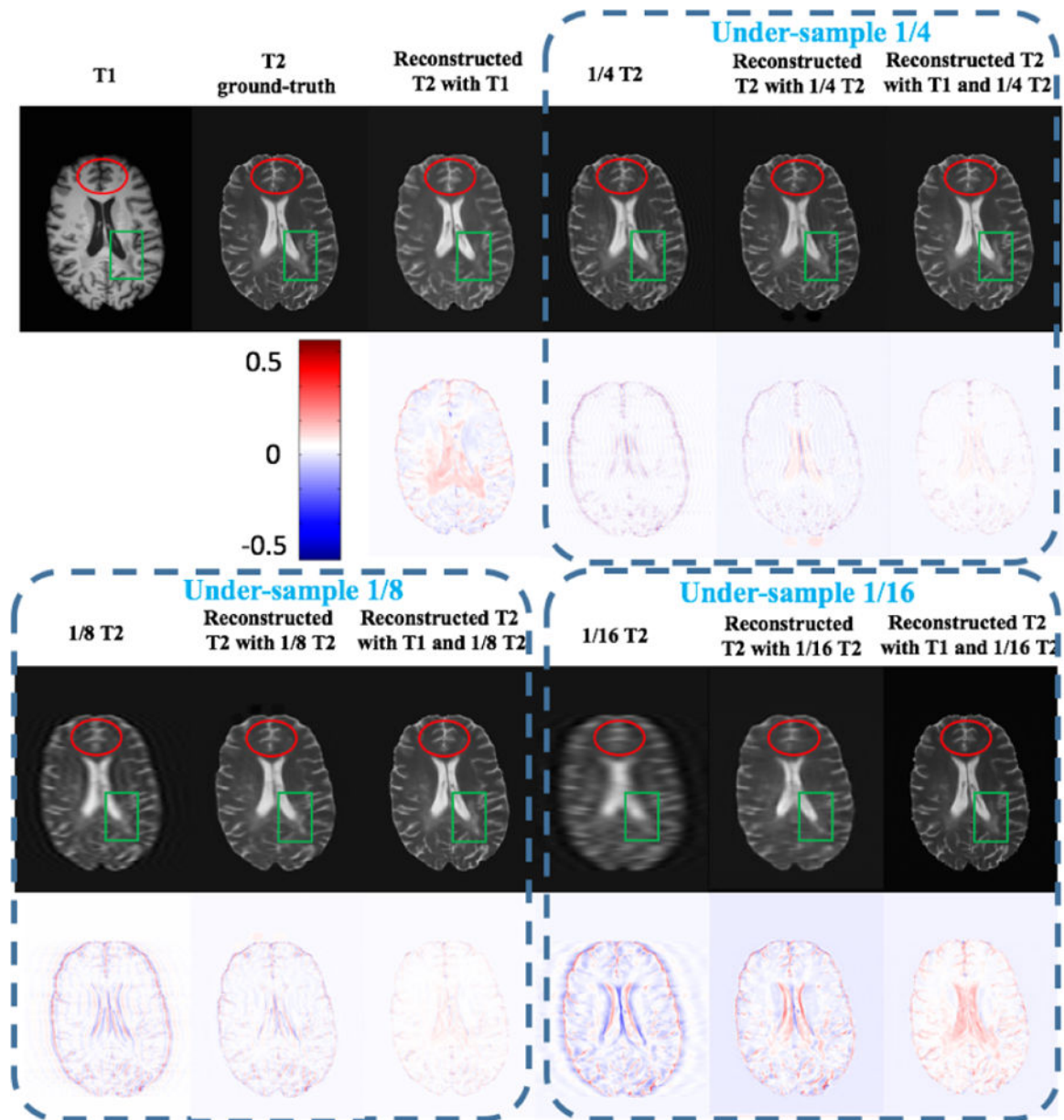
**Fig. 2.**

(a) Illustration of the proposed Dense-Unet for T2WI reconstruction with T1WI and under-sampled T2WI as concatenated input; (b) illustration of the detailed configuration of the dense block. Note that we implement the input in (a) as six consecutive axial slices (with three from fully-sampled T1WI and three from under-sampled T2WI). In (b), each dense block consists of five convolutional layers. The growth rate is set to 16, and the output of each block has 80 feature maps.

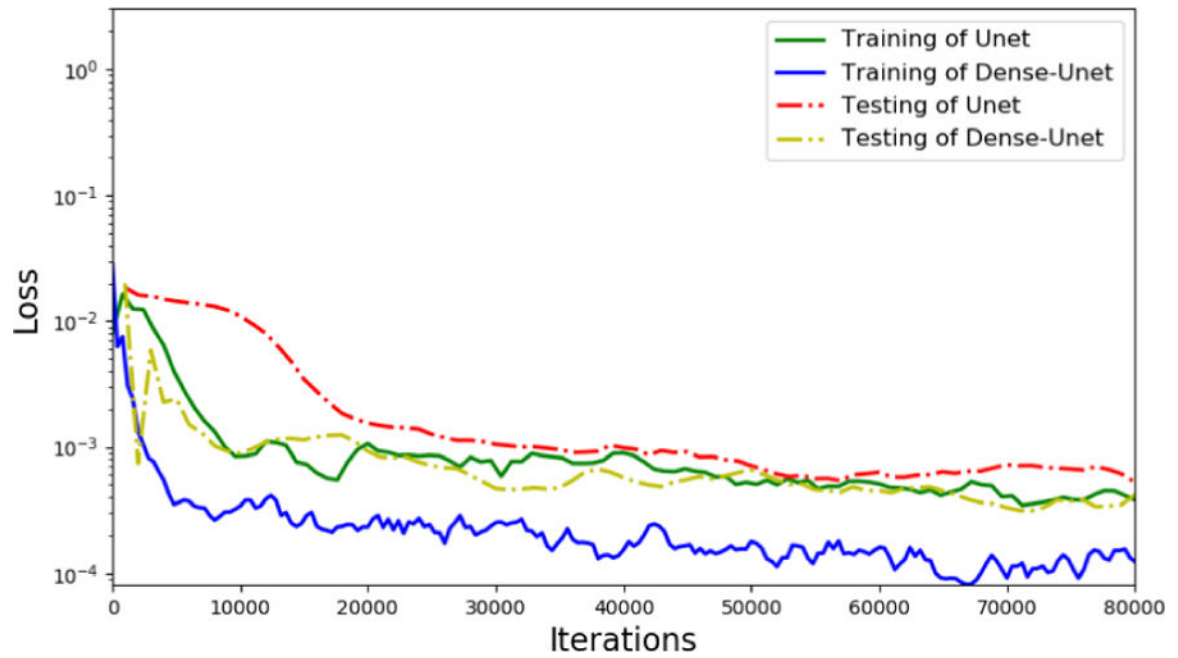


**Fig. 3.**

Visual inspection of using multi-modal input for T2WI reconstruction. There are three views in the figure. The first row of each view shows the input images, as well as the reconstruction results by different input settings. The second row provides the error map (encoded by the color bar) between the reconstructed T2 image and the ground truth. Red circles and green boxes are placed on the corresponding locations of individual images for better comparison of specific structures.

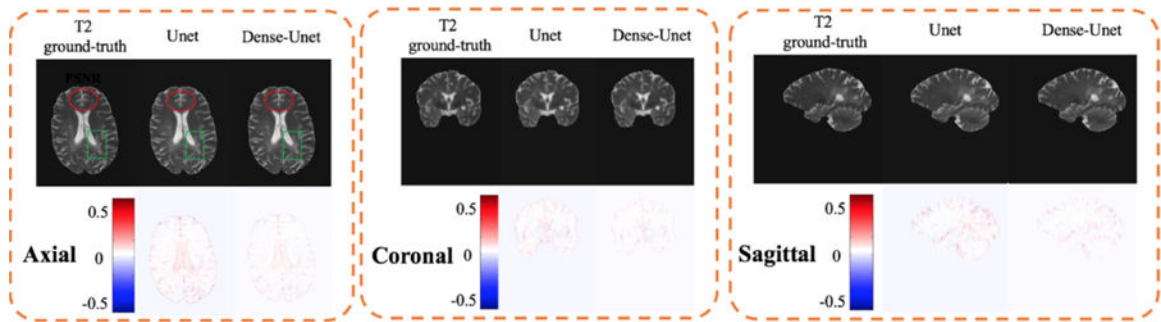


**Fig. 4.** Visualization of reconstructed T2WI with different under-sampling rates, including 1/4, 1/8 and 1/16. The reconstruction is obtained with Dense-Unet. For fair comparison, we also reconstruct the T2 images using fully sampled T1WI only as shown in the first row.



**Fig. 5.**

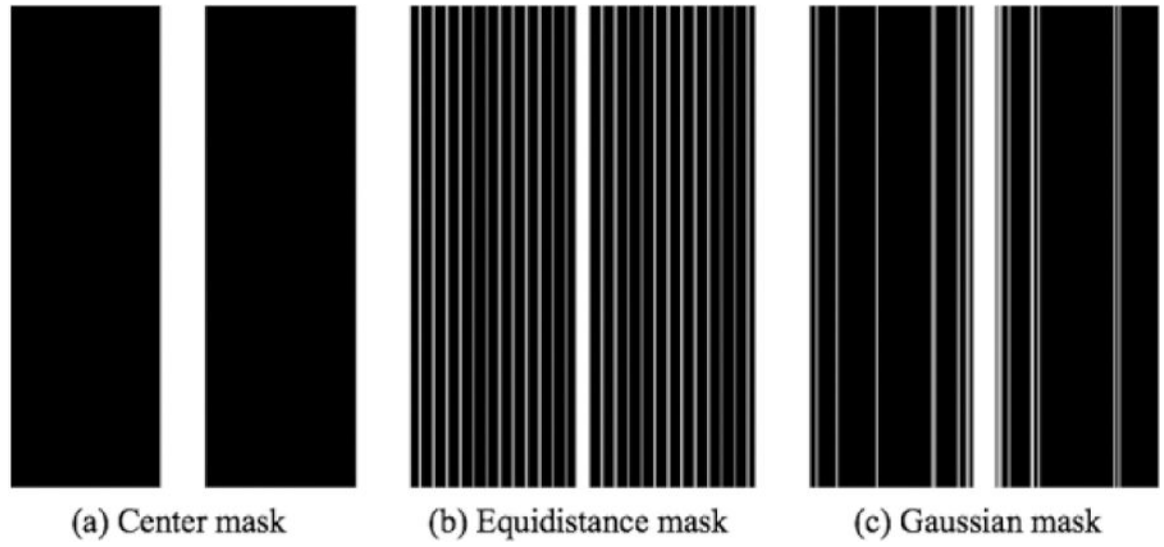
Comparison of training and testing loss convergence between Unet and Dense-Unet.



**Fig. 6.**

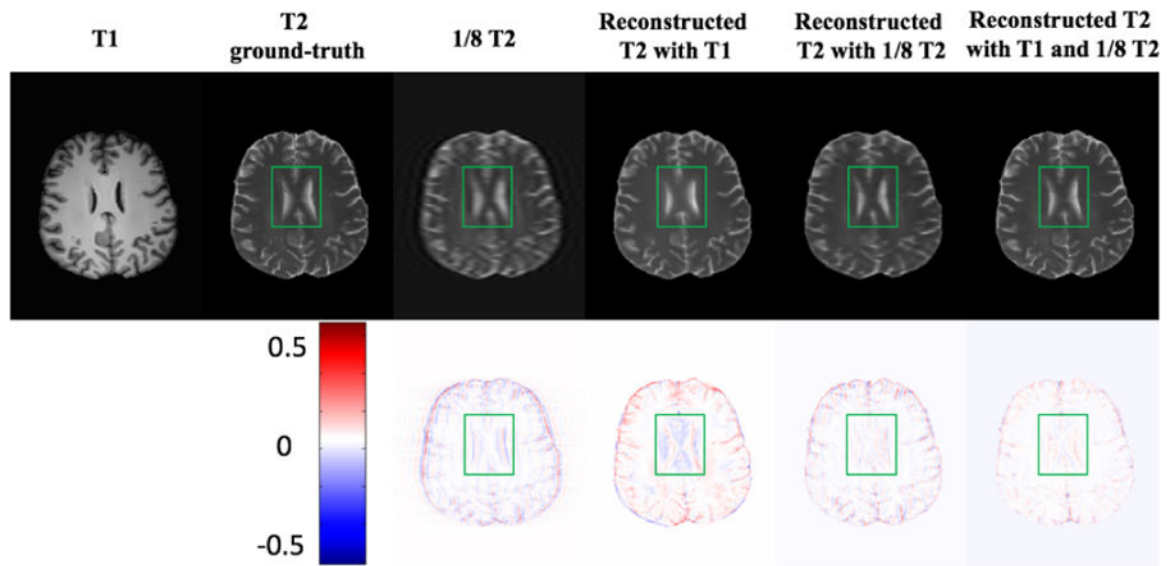
Comparisons between Unet and the proposed Dense-Unet using input data from T1WI and 1/8 under-sampled T2WI. Reconstructed images from all three views (axial, coronal and sagittal) are plotted. The first row for each view presents the reconstructed images and the second row shows the corresponding error map as compared to the ground truth.





**Fig. 7.**

Examples of different mask: (a) Center Mask, (b) Equidistance Mask, and (c) Gaussian Mask. The under-sample rate is  $1/8$  in  $k$ -space for all masks.



**Fig. 8.**

Visual comparison of various T2WI reconstruction results (no lesion). The first row represents reconstruction results by different inputs. The second row presents the corresponding error map as compared to the ground-truth T2WI.

**TABLE I**

EVALUATION OF THE RECONSTRUCTED T2WI USING DIFFERENT INPUT SETTINGS ON DATASET 1

	1/8 T2		Reconstructed T2 with T1		Reconstructed T2 with 1/8 T2		Reconstructed T2 with T1 and 1/8 T2	
	PSNR	MAE	PSNR	MAE	PSNR	MAE	PSNR	MAE
Subject 1	33.1	0.022	30.9	0.028	34.5	0.019	<b>37.6</b>	<b>0.013</b>
Subject 2	33.2	0.022	30.4	0.030	33.7	0.021	<b>37.0</b>	<b>0.014</b>
Subject 3	31.5	0.027	30.3	0.030	32.4	0.024	<b>36.5</b>	<b>0.015</b>
Subject 4	32.1	0.025	30.1	0.031	33.4	0.021	<b>36.4</b>	<b>0.015</b>
Subject 5	32.6	0.023	31.2	0.028	33.5	0.017	<b>37.0</b>	<b>0.014</b>
Average	32.5	0.024	30.6	0.030	33.9	0.020	<b>36.9</b>	<b>0.014</b>

**TABLE II**

RECONSTRUCTION PERFORMANCES BY USING DIFFERENT HYPER-PARAMETER SETTINGS AND UNDER-SAMPLING RATES IN DENSE-UNET.

Reconstructed T2 with different parameters	T1+1/4 T2		T1+1/8 T2		T1+1/16 T2	
	PSNR	MAE	PSNR	MAE	PSNR	MAE
GR=16, NC=4	37.7	0.013	34.1	0.020	33.5	0.021
GR=16, NC=5	<b>39.1</b>	<b>0.011</b>	<b>36.9</b>	<b>0.014</b>	<b>34.3</b>	<b>0.019</b>
GR=16, NC=6	38.9	0.012	34.7	0.019	32.4	0.024
GR=24, NC=4	38.5	0.012	36.5	0.015	34.1	0.020
GR=24, NC=5	38.8	0.012	36.3	0.016	34.0	0.020

**TABLE III**

COMPARISON OF PSNR AND MAE VALUES BETWEEN UNET AND DENSE-UNET AT DIFFERENT UNDER-SAMPLING RATIOS.

Reconstructed T2 with different models	T1+1/4 T2		T1+1/8 T2		T1+1/16 T2	
	PSNR	MAE	PSNR	MAE	PSNR	MAE
Unet	37.6	0.013	36.6	0.015	33.8	0.020
Dense-Unet	<b>39.1</b>	<b>0.011</b>	<b>36.9</b>	<b>0.014</b>	<b>34.3</b>	<b>0.019</b>

**TABLE IV**

COMPARISON OF PSNR AND MAE VALUES BETWEEN UNET AND DENSE-UNET AT DIFFERENT UNDER-SAMPLING RATIOS.

Center Mask		Equidistance Mask		Gaussian Mask	
PSNR	MAE	PSNR	MAE	PSNR	MAE
36.9	0.014	36.2	0.013	33.3	0.021

**TABLE V**

COMPARISON OF THE RECONSTRUCTION PERFORMANCE WHEN THE NETWORK TRAINED WITH DATASET 2 IS TRANSFERRED AND TESTED UPON DATASET 1.

Reconstructed T2 with different parametrs	T1+1/4 T2		T1+1/8 T2		T1+1/16 T2	
	PSNR	MAE	PSNR	MAE	PSNR	MAE
Dense-Unet (D1-D1)	39.1	0.011	36.9	0.014	34.3	0.019
Dense-Unet (D2-D1)	38.5	0.012	36.7	0.014	34.0	0.020