



Ultra-Fast T2-Weighted MR Reconstruction Using Complementary T1-Weighted Information

Lei Xiang¹, Yong Chen², Weitang Chang², Yiqiang Zhan¹,
Weili Lin², Qian Wang^{1(✉)}, and Dinggang Shen^{2(✉)}

¹ Institute for Medical Imaging Technology, School of Biomedical Engineering,
Shanghai Jiao Tong University, Shanghai, China
wang.qian@sjtu.edu.cn

² Department of Radiology and BRIC, University of North Carolina
at Chapel Hill, Chapel Hill, NC, USA
dgshen@med.unc.edu

Abstract. T1-weighted image (T1WI) and T2-weighted image (T2WI) are the two routinely acquired Magnetic Resonance Imaging (MRI) protocols that provide complementary information for diagnosis. However, the total acquisition time of ~ 10 min yields the image quality vulnerable to artifacts such as motion. To speed up MRI process, various algorithms have been proposed to reconstruct high quality images from under-sampled k-space data. These algorithms only employ the information of an individual protocol (e.g., T2WI). In this paper, we propose to combine complementary MRI protocols (i.e., T1WI and under-sampled T2WI particularly) to reconstruct the high-quality image (i.e., fully-sampled T2WI). To the best of our knowledge, this is the first work to utilize data from different MRI protocols to speed up the reconstruction of a target sequence. Specifically, we present a novel deep learning approach, namely Dense-Unet, to accomplish the reconstruction task. The Dense-Unet requires fewer parameters and less computation, but achieves better performance. Our results have shown that Dense-Unet can reconstruct a 3D T2WI volume in less than 10 s, i.e., with the acceleration rate as high as 8 or more but with negligible aliasing artefacts and signal-noise-ratio (SNR) loss.

1 Introduction

Magnetic Resonance Imaging (MRI) is widely applied for numerous clinical applications, as it can provide non-invasive, reproducible and quantitative measurements of both anatomical and functional information that are essential for disease diagnosis and treatment planning. MRI can better measure different soft tissue contrasts than many other medical imaging modalities, e.g., Computed Tomography (CT). It also avoids exposing patients to harmful ionizing radiation, implying higher safety. However, MRI acquisitions need to sample full k-space for orthogonal encoding of the spatial information if no acceleration scheme is employed, therefore limiting acquisition speed. The k-space data that encode spatial-frequency information are commonly acquired line-by-

The original version of this chapter was revised: an Acknowledgements section has been added. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-00928-1_100

line with a fixed interval (repetition time). During MR acquisition, patient movement or physiological motion, e.g., respiration, cardiac motion, and blood flow, could result in significant artefacts in the MR images. Long scan time also increases the healthcare cost and limits the availability of MR scanner for patients.

In clinical routines, T1WI and T2WI images are the two basic MR sequences for assessing anatomical structure and pathology, respectively. T1WI is useful for assessing the cerebral cortex, 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 white matter lesions and assessing zonal anatomy in the prostate and uterus. These two sequences provide complementary information to each other and help characterize the abnormalities of the patients. Typical scanning time for T1WI and T2WI is ~ 10 min, in which T2WI takes the majority due to its longer repetition time (TR) and echo time (TE). In clinical practice, further acceleration in MR acquisition is desired to (1) scan more patients and (2) reduce motion artefacts. Since image acquisition time for a given sequence depends on the number of sampled lines in the k-space, many methods focus on the reduction of the k-space sampling rate, i.e., under-sampling the k-space. These approaches capitalize the inherent redundancy in MRI, where individually sampled points in the k-space do not arise from distinct spatial locations.

Recently, deep learning techniques have been applied to accelerate MRI acquisition. Wang et al. [1] proposed to train a CNN model to learn the mapping between the MR image obtained from zero-filled under-sampled and fully-sampled k-space data, and then use the reconstruction result either as an initialization or a regularization term in the classical CS MRI process. Lee et al. [2] further introduced a deep multi-scale residual learning algorithm to reconstruct the under-sampled MR image by formulating the CS MRI problem as a residual regression problem. Schlemper et al. [3] developed a deep cascade of CNNs to reconstruct the aggressive Cartesian under-sampled MR image. When the frames of the sequences are reconstructed jointly, they demonstrated to learn the spatiotemporal correlation efficiently by leveraging convolution and data sharing layers together.

Although T1WI and T2WI have different contrasts, these two protocols contain highly related information. White matter in T1WI is of high signal intensity, while it becomes dark in T2WI. Similarly, grey matter is of low signal intensity in T1WI, while it appears white in T2WI. Two lesion examples are labeled in the red circle and green box in Fig. 1. These two lesions are both clear in T2WI, while only one lesion in green box can be noted in the T1WI. In 1/8 under-sampled T2WI, the lesions can still be found, but the boundary details are mostly missing. The reduced image quality in 1/8 T2WI is also prevalent in normal areas as indicated by the blue arrows. These observations inspire us to design a deep learning approach that achieves ultra-fast T2WI reconstruction by combining T1WI with highly under-sampled T2WI.

Several studies [4, 5] attempt to reconstruct T2WI from T1WI. However, to the best of our knowledge, this is the *first work* that intends to reconstruct T2WI using both T1WI and under-sampled T2WI. Our method can leverage the complex relation between T1WI and T2WI and utilize the unique information in under-sampled T2WI to reconstruct the fully-sampled T2WI. We adapt the deep fully convolutional neural network, i.e., the Unet architecture [6] that consists of a contracting (or encoding) path and a symmetric expanding (decoding), to leverage the context information from multi-

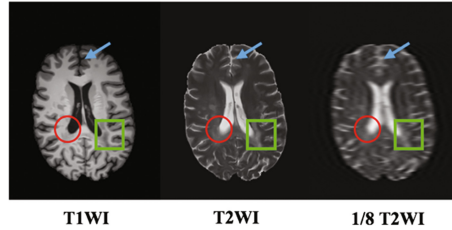


Fig. 1. Examples of the pairs of T1WI, T2WI and 1/8 under-sampled T2WI data from the same patient. Multiple sclerosis lesions are marked by circles and boxes in the figure.

scale feature maps. In this way, we can combine T1WI and T2WI through the network. We particularly select multiple corresponding 2D slices from T1WI and under-sampled T2WI, and concatenate them as multi-channel input to the network. We further develop the Dense-Unet architecture by introducing dense blocks, which significantly reduces the parameters of the network while boosts the reconstruction quality. Experimental results suggest that the acceleration rate can be as high as 8 or more with negligible aliasing artefacts and signal-noise-ratio (SNR) loss.

2 Method

Overview. The framework for accelerating the T2WI reconstruction with T1WI and under-sampled T2WI is shown in Fig. 2(a). First, the original T2WI is retrospectively under-sampled and only the center part of the k-space with the ratio R , e.g., $R = 1/2, 1/4, 1/8, \dots$, is utilized to reconstruct the under-sampled T2WI. The under-sampled T2WI is then concatenated with the fully-sampled T1WI, and fed together into the Dense-Unet architecture. To be specific, we implement the input as two groups of consecutive axial slices (N from fully-sampled T1WI and N corresponding ones from under-sampled T2WI). These $2N$ slices are jointly considered as part of the feature maps of the first convolutional layer in the network. The output of the network consists of N consecutive axial slices that correspond to the input. In testing, we synthesize every N consecutive axial slices for the reconstructed T2WI, and combine the outputs into the final 3D volume of T2WI by simple averaging. We regard this joint synthesis of the N consecutive slices as a quasi-3D mapping.

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

$$F_1 = \sigma(W_1 * [x_{T2}, y_{T1}] + B_1) \quad (1)$$

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

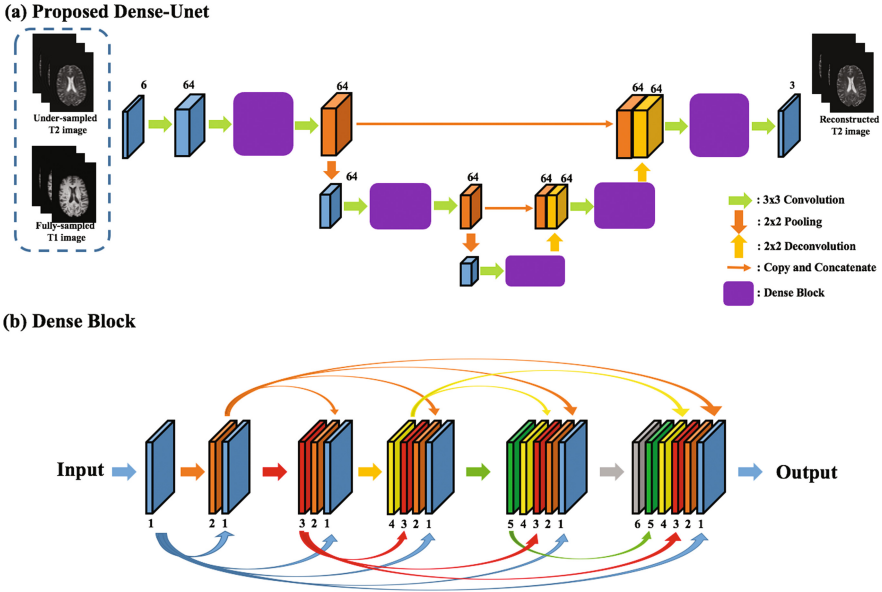


Fig. 2. Illustration of (a) the framework for T2WI reconstruction with T1WI and under-sampled T2WI and (b) the detailed configuration of dense block.

Dense Block. Dense connectivity has been proposed in [7] to further improve the information flow between layers. We adopt this new convolutional network architecture in our model so that we can increase the depth of the whole network to dozens of layers with feasible optimization. Moreover, the dense block requires substantially fewer parameters and less computation, which makes the model more efficient and costs less memory. Figure 2(b) illustrates the layout of the dense block. Consequently, the m^{th} layer receives the feature maps of all preceding layers as the input:

$$z_l = H_l([z_0, \dots, z_{l-1}]) \tag{2}$$

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

Transition Layers. We refer the convolution (or deconvolution) layers following the dense block as the transition layers, which include batch normalization, convolution and pooling (if in the contracting path), or just batch normalization and deconvolution (if in the expanding path). The transition layer in the contracting path used in our experiments consists of a BN layer and a 1×1 Conv layer followed by a 2×2 average pooling layer. In our experiment, the feature map number is always set to 64. On the

expanding path, the dense block is followed by batch normalization and deconvolution layer consisting of 64 filters of size 3×3 .

Reconstruction Layer. The proposed Dense-Unet model ends with the reconstruction layer that reconstructs fully-sampled T2WI from the feature maps outputted by the last dense block. The reconstruction can be attained by a single convolutional layer as

$$\tilde{y}_{T2} = \mathbf{W}_R * \mathbf{F}_D + \mathbf{B}_R. \quad (3)$$

Here, \mathbf{F}_D is the feature maps outputted by the last dense block, and \tilde{y}_{T2} is the T2WI estimated by the reconstruction layer. Note that there is no activation function employed in the reconstruction layer. We use mean squared error (MSE) as loss function.

3 Experimental Result

3.1 Dataset

We utilized the dataset from the MICCAI Multiple Sclerosis (MS) segmentation challenge 2016 [8] to demonstrate the capability of proposed Dense-Unet. We selected 5 subjects with paired T1WI and T2WI. These subjects were scanned by the same Philips Ingenia 3T scanner. The voxel size is $0.7 \times 0.744 \times 0.744 \text{ mm}^3$. Multiple pre-processing steps are applied, including: (1) denoising with the NL-means algorithm of each image; (2) rigid registration of each image; (3) brain extraction using the volBrain platform from T1WI and then being applied on the other modalities with sinc interpolation; (4) bias correction using the N4 algorithm; (5) intensity normalization to range (0,1) by dividing the maximum intensity. The final matrix size of the images is $336 \times 336 \times 261$. We used consecutive 2D slices to train our model. In this way, each subject can provide hundreds of samples, which is sufficient for the training of the network.

3.2 Experimental Setting

In this paper, PyTorch was used to implement the Dense-Unet architecture. In the training phase, we extracted 2D whole slices from the 3D image, and each 3D image could contribute ~ 200 samples. We excluded the slices without any brain tissue. The leave-one-out cross-validation strategy was employed for evaluation. Also, data augmentation of left-right flipping was applied. Therefore, we prepared enough samples for training our deep model. We adopted Adam optimization with momentum of 0.9 and performed 100 epochs in training stage. The batch size was set to 4 and the initial learning rate was set to 0.0001, which was divided by 10 after 50 epochs. We used zero-padding during every convolution layer to make sure that the size of the output is the same as the input. To quantitatively evaluate the reconstruction performance, we used the standard metric of peak signal-noise ratio (PSNR) and normalized mean absolute error (MAE).

3.3 Contribution of Adding T1WI

To demonstrate the effectiveness of integrating T1WI data for the reconstruction of under-sampled T2WI, we compare the performance achieved (1) by using under-sampled T2WI as the only input and (2) by using the combination of T1WI and under-sampled T2WI. When dealing with the only input of under-sampled T2WI, we employed the same setting as in Fig. 2, but the input layer only includes the under-sampled T2WI. The under-sample ratio was 1/8 for this experiment. Averaged PSNRs and MAEs are listed in Table 1. ‘1/8 T2’ indicates the PSNR/MAE scores computed by comparing the input 1/8 under-sampled T2WI with the ground-truth T2WI directly. ‘Reconstructed T2 with T1 (or 1/8 T2)’ represents the reconstructed T2WI results using only T1WI (or only 1/8 T2WI) as input. ‘Reconstructed T2 with T1 and 1/8 T2’ represents the reconstructed T2WI using combined inputs of T1WI and 1/8 T2WI.

Table 1. The evaluation of the reconstructed T2WI using different input settings.

	1/8 T2	Reconstructed T2 with T1	Reconstructed T2 with 1/8 T2	Reconstructed T2 with T1 and 1/8 T2
PSNR	32.5	30.6	33.9	36.9
MAE	0.023	0.033	0.020	0.014

We can see that the results of ‘Reconstructed T2 with T1 and 1/8 T2’ are superior than both ‘Reconstructed T2 with T1’ and ‘Reconstructed T2 with 1/8 T2’. The image quality of ‘Reconstructed T2 with 1/8 T2’ (PSNR: 33.9 dB) improves just a little compared to ‘1/8 T2’ (PSNR: 32.5 dB), which implies that over under-sampled T2WI is difficult to be reconstructed. However, with T1WI added, the reconstructed results of ‘Reconstructed T2 with T1 and 1/8 T2’ (PSNR: 36.9 dB) demonstrate a clear improvement compared to that of ‘Reconstructed T2 with 1/8 T2’ (PSNR: 33.9 dB). This verifies the idea that different image contrasts have complementary information to each other in reconstruction. And, with the additional contrast information of T1WI, we can achieve better reconstruction of T2WI.

We also provide one example in Fig. 3 for visual observation, where our method yields satisfying reconstruction result regarding the ground-truth by keeping high contrast of tissue boundaries. For ‘Reconstructed T2 with T1’, it has predicted the common tissue structure correctly, e.g., white matter and gray matter in the red circle, but it misses the lesion part (green box). This is because there are no enough cues in T1WI to show the whole lesion information. When it comes to ‘Reconstructed T2 with 1/8 T2’, we can see that lesion part (green box) is preserved and reconstructed, but it has fuzzy boundary. Moreover, other contrast details (red circle) have become blurred because they are all recovered from only the under-sampled T2WI. Based on the experimental results above, we design the proposed model that takes advantages of both T1WI (for detailed and clear common structure) and under-sampled T2WI (for the unique information that only appears in the T2WI). So we combine the T1WI and 1/8 T2WI together as the input of our proposed Dense-Unet. Finally, ‘Reconstructed T2 with T1 and 1/8 T2’ gives the most satisfying reconstruction result with high perceptive

quality. Note that most of the current compressed sensing methods can only handle reconstruction of under-sample ratio around 1/4. Our proposed method pushes this limitation to a new level, i.e., for even under-sample 1/8 we can still reconstruct correct image details, which further accelerates the speed of scanning T2WI by twice faster while preserving image quality. Moreover, our proposed quasi-3D mapping preserves the consistency in the third view, as shown in both coronal and sagittal views.

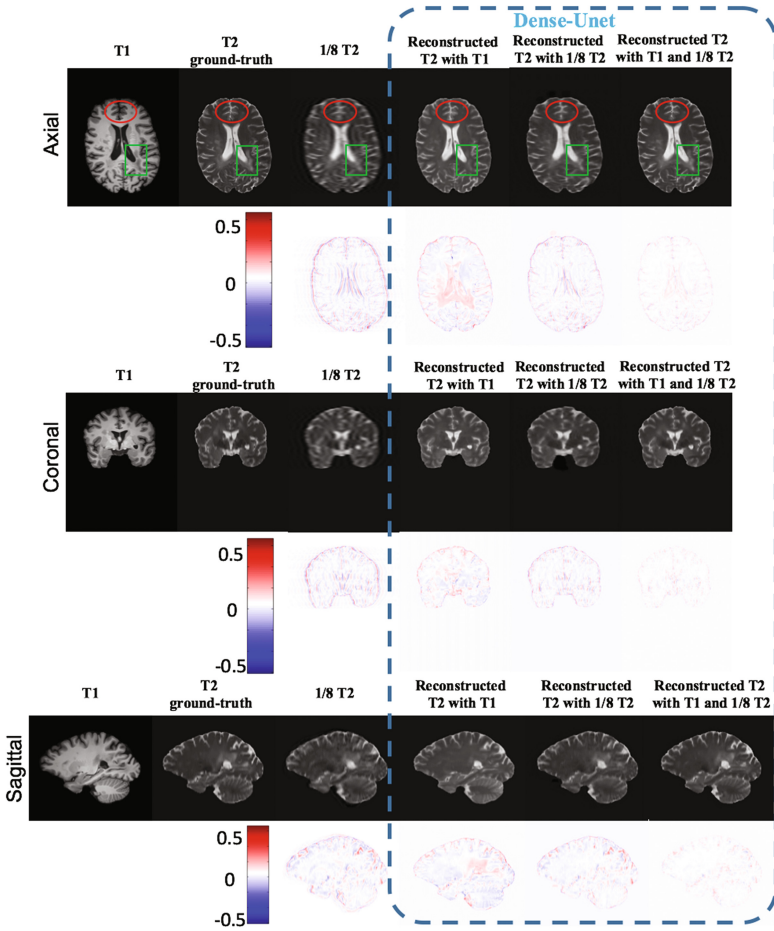


Fig. 3. Visual examples of using multi-inputs for T2WI reconstruction.

3.4 Comparison with Unet

In the literature, there are many studies that successfully reconstruct fully-sampled MR image from under-sampled images. But none of them attempted to solve this problem by utilizing information from another contrast, e.g., T1WI. This is the *first work* that utilizes T1WI and under-sampled T2WI to reconstruct the fully-sampled T2WI. In this

section, we mainly compare our method with the popular neural network architecture, Unet, which can handle two inputs naturally. Quantitative results of average PSNRs and MAEs are summarized in Table 2. We compare their performance in different under-sampled ratios. Dense-Unet clearly outperforms the Unet under all comparisons. Note that Unet has 9.5 M parameters, while Dense-Unet only has 3.2 M parameters. Thus, Dense-Unet consumes 3 times less storage than Unet. Also, the running-time for Dense-Unet is 9.5 s for a 3D volume ($336 \times 336 \times 260$), which is highly efficient.

Table 2. The reconstruction comparison of Unet and Dense-Unet.

	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	39.1	0.011	36.9	0.014	34.3	0.019

4 Conclusion

In this paper, we propose a novel Dense-Unet model to reconstruct the T2WI from the T1WI and under-sampled T2WI. The added T1WI makes the reconstruction of T2WI from 1/8 under-sample ratio in k-space possible, which leads to 8 times speed-up. The dense block, which requires substantially fewer parameters and less computation, is integrated within Unet architecture in our work. This enables our model to further improve the quality of the reconstructed T2WI. Comprehensive experiments showed superior performance of our method, including the perceptive quality and the running speed. This work thus can potentially improve the acquisition efficiency and image quality in clinical settings.

Acknowledgement. This work was supported in part by NIH grant EB006733.

References

1. Wang, S., Su, Z., Ying, L., Peng, S., Liang, F., Feng, D., Liang, D.: Accelerating magnetic resonance imaging via deep learning. In: IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 514–517. IEEE (2016)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
3. Schlemper, J., Caballero, J., Hajnal, J.V., Price, A., Rueckert, D.: A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. arXiv preprint [arXiv:1704.02422](https://arxiv.org/abs/1704.02422) (2017)
4. Alkan, C., Cocjin, J., Weitz, A.: Magnetic Resonance Contrast Prediction Using Deep Learning (2017)
5. Vemulapalli, R., Van Nguyen, H., Kevin Zhou, S.: Unsupervised cross-modal synthesis of subject-specific scans. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 630–638 (2015)

6. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, Alejandro F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
7. Huang, G., Liu, Z., Weinberger, K.Q., van der Maaten, L.: Densely connected convolutional networks. arXiv preprint [arXiv:1608.06993](https://arxiv.org/abs/1608.06993) (2016)
8. Commowick, O., Cervenansky, F., Ameli, R.: MSSEG challenge proceedings: multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure. In: MICCAI (2016)