Deep Learning for Fast and Spatially-Constrained Tissue Quantification from Highly-Accelerated Data in Magnetic Resonance Fingerprinting – *Supplementary Materials*

Zhenghan Fang, Yong Chen, Mingxia Liu, Lei Xiang, Qian Zhang, Qian Wang, Weili Lin, and Dinggang Shen*, *Fellow, IEEE*

I. Effects of Important Components in Our Method (cont.)

In the main text, we have analyzed the influence of three major components in our proposed spatially-constrained tissue quantification (SCQ) method. Here, we further study the influence of another two components in SCQ via experiments.

A. Influence of Network Structures

In this group of experiments, we study the influence of the network structures in our proposed feature extraction (FE) module and spatially-constrained quantification (SQ) module.

Specifically, for the FE module, we compare fully-connected neural networks (FNNs) with different numbers of fully-connected (FC) layers: 2 (**2FC**), 4 (**4FC**, i.e., SCQ), and 6 (**6FC**). The output dimensions of all FC layers are set as 46. The results are summarized in Table SI. As shown in Table SI, for T1 quantification, the number of FC layers does not have any influence on the accuracy, whereas for T2, 6FC achieves a slightly higher accuracy than 2FC and 4FC.

For the SQ module, we compare U-Nets [1] with different numbers of down- and up-sampling (DUS) layers: 0 (**0DUS-U-Net**), 1 (**1DUS-U-Net**), 2 (**2DUS-U-Net**), 3 (**3DUS-U-Net**, i.e., SCQ), and 4 (**4DUS-U-Net**). The network structures of 0, 1, 2, and 4DUS-U-Nets are shown in Fig. S1 (a)-(d), respectively, while the network structure of 3DUS-U-Net is shown in Fig. 3 in the main text. The results are summarized in Table SII.

As shown in Table SII, the quantification error generally decreases with the increase of number of DUS layers. This is likely because the increase of DUS layers enlarges the receptive field of U-Net, which allows more spatial context information to be exploited.

TABLE SI
Quantification errors of T1 and T2 (mean ± standard deviation, unit: %) yielded by our SCQ method using different network structures in the FE module. 2FC, 4FC, and 6FC: FNNs with 2, 4, and 6 FC layers, respectively. Here, the acceleration rate (ar) is set as 4.

	2FC	4FC (SCQ)	6FC
T1	2.1 ± 0.3	2.1 ± 0.2	2.1 ± 0.3
T2	5.8 ± 0.6	5.8 ± 0.7	5.6 ± 0.7

TABLE SII
Quantification errors of T1 and T2 (mean ± standard deviation, unit: %) yielded by our SCQ method using different network structures in the SQ module. 0, 1, 2, 3, and 4DUS-U-Net: U-Nets with 0, 1, 2, 3, and 4 down- and up-sampling layers, respectively. Here, $ar = 4$.

	0DUS-U-Net	1DUS-U-Net	2DUS-U-Net	3DUS-U-Net (SCQ)	4DUS-U-Net
T1	2.3 ± 0.4	2.3 ± 0.4	2.0 ± 0.3	2.1 ± 0.3	2.0 ± 0.3
T2	6.0 ± 0.8	5.9 ± 0.7	5.8 ± 0.7	5.8 ± 0.7	5.7 ± 0.8

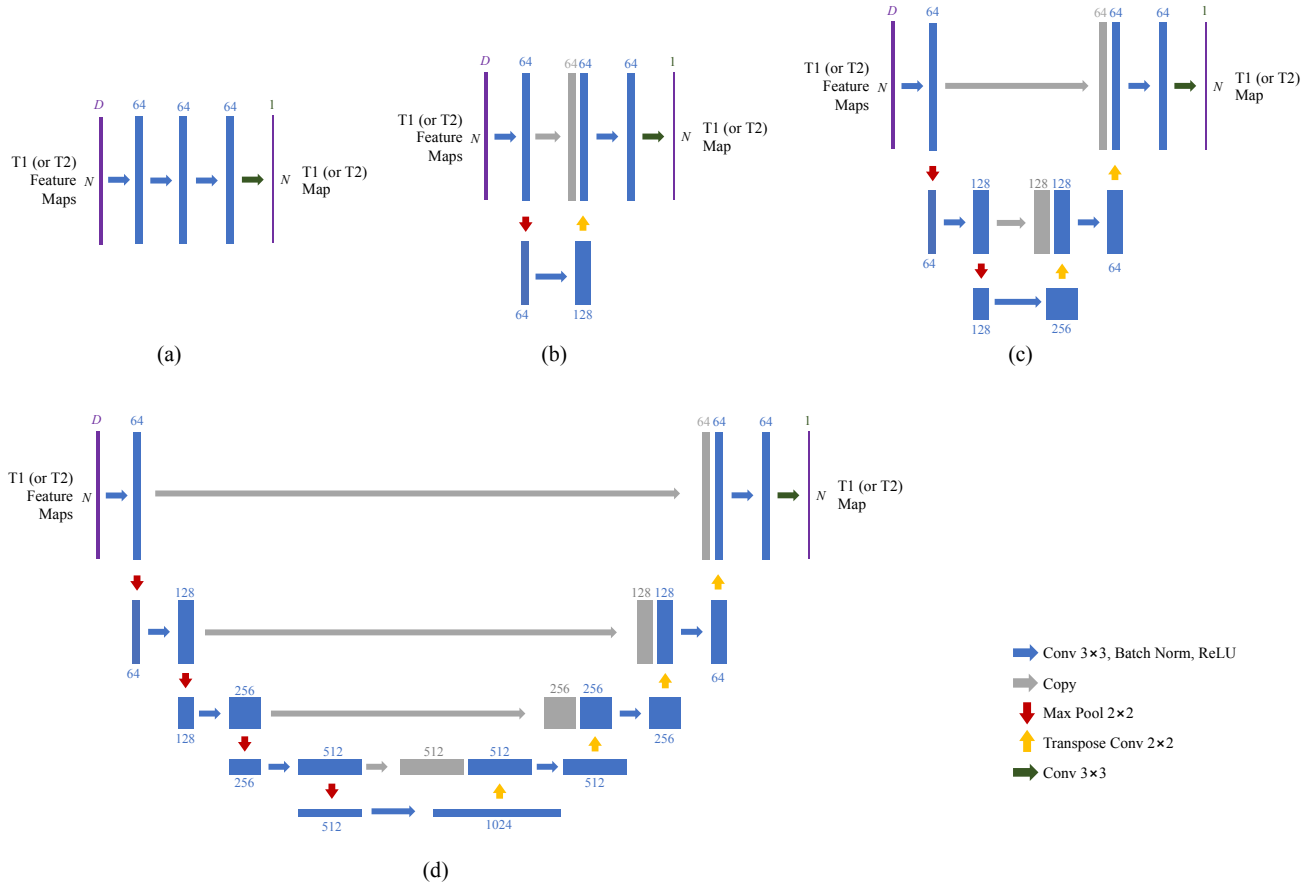

Fig. S1. The network structures used for the spatially-constrained quantification (SQ) module. (a)-(d): 0, 1, 2, and 4DUS-U-Nets, i.e., U-Nets with 0, 1, 2, and 4 down-sampling layers, respectively.

B. Influence of Data Normalization

In this group of experiments, we study the influence of the data normalization step in our data preprocessing method (see Section II. A in the main text). Specifically, we compare our SCQ method (i.e., with data normalization) and its variant without data normalization (named **SCQ_N**). We also report the results of the FE-only model (the variant of our method without the SQ module, introduced in Section III. E. 2 in the main text) when using normalized signals (FE-only) and original signals (named **FE-only_N**, i.e., without data normalization) as input. These results are reported in Table SIII.

As shown in Table SIII, the quantification errors of FE-only_N are higher than those of FE-only, indicating that data normalization is helpful for quantification accuracy when SQ module is removed and hence spatial context information is not exploited. Besides, the quantification errors of SCQ_N and SCQ are similar, indicating that data normalization is less needed when spatial context information is exploited.

TABLE SIII
QUANTIFICATION ERRORS OF T1 AND T2 (MEAN ± STANDARD DEVIATION, UNIT: %) YIELDED BY FE-ONLY_N (FE-ONLY WITHOUT DATA NORMALIZATION), FE-ONLY (A VARIANT OF SCQ WITHOUT SPATIALLY-CONSTRAINED QUANTIFICATION MODULE), SCQ_N (SCQ WITHOUT DATA NORMALIZATION), AND SCQ (OUR SPATIALLY-CONSTRAINED TISSUE QUANTIFICATION METHOD). HERE, $ar = 4$.

	FE-only_N	FE-only	SCQ_N	SCQ
T1	2.91 ± 0.69	2.41 ± 0.49	2.04 ± 0.25	2.08 ± 0.23
T2	7.33 ± 0.77	7.02 ± 0.78	5.83 ± 0.72	5.81 ± 0.66

C. Influence of Feature Number, and Comparison between FNN and SVD (cont.)

In this section, we study the visual results of our SCQ method and its variant (SCQ_S) which uses PCA, instead of FNN, for feature extraction in FE module. The images of tissue property maps estimated with 10 or 46 extracted features (i.e., $D =$ 10 or 46) are shown in Fig. S2. As shown in Fig. S2, using only 10 features yields reasonably good results, but with a larger

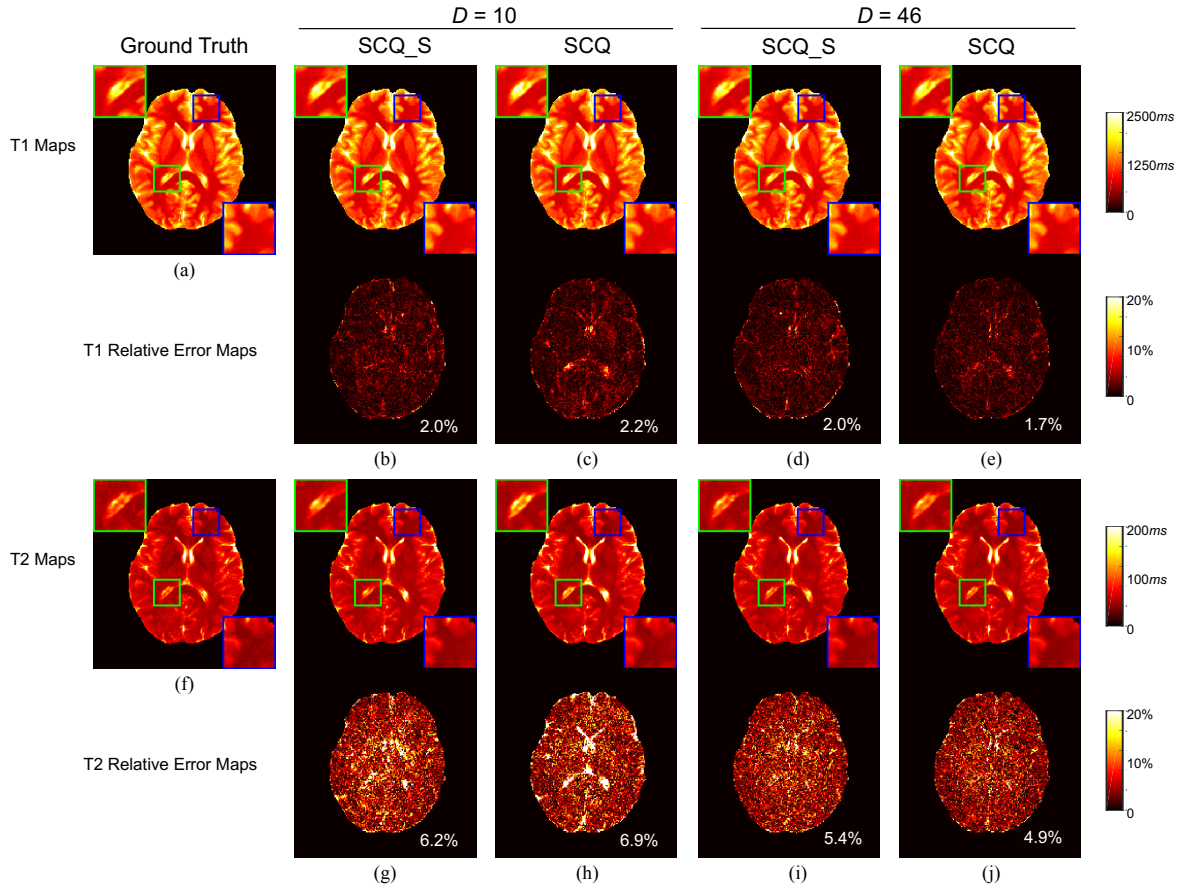

Fig. S2. Visual results of T1 and T2 quantification by our method (SCQ) and its variant (SCQ_S) which uses SVD (instead of FNN) for feature extraction in the FE module, with feature number $D = 10$ or 46. (a) Ground-truth T1 map. (b)-(e): Estimated T1 maps and associated relative error maps. (f) Ground-truth T2 map. (g)-(j): Estimated T2 maps and associated relative error maps. The overall error is labeled at the lower right corner of each error map. Here, $ar = 4$.

error in the CSF area, compared to the images obtained with 46 features. Moreover, the result from SCQ with 10 features has a larger error in the CSF area than that from SCQ_S, which causes the quantification error of SCQ to be higher than SCQ_S when the feature number is around 10 (as shown in Fig. 6 of the main text).

II. Spatial Resolution

As observed in T2 result with 8× acceleration in Fig. 4 of the main text, the spatial constraint imposed by our method can cause blurriness in the estimated tissue property map and degrade its spatial resolution when high acceleration rate is applied. In this section, we further study the observed image blurriness through the frequency-domain representations of tissue property maps. We also examine the influence of two settings in the SQ module on spatial resolution.

A. Image Blurriness in Estimation Results

Since spatial smoothing will reduce the energy of high-frequency (HF) components in an image, we calculate the energy of HF components in the tissue property map estimated by our SCQ method and compare it to that of the ground-truth map. Specifically, the frequency components in the upper 25% of the spectrum are defined as high-frequency components and the proportion of the sum of their energy to the sum of energy of all frequency components is calculated. The brain slice shown in Fig. 4 is selected for a case study and the results are summarized in Table SIV. Table SIV clearly shows a decrease in HF energy in our results compared to the ground-truth maps, demonstrating the blurring effects caused by our proposed method. Moreover, the higher reduction in HF energy is observed when the acceleration rate increases, which is in accordance with the degree of blurriness observed in Fig. 4 of the main text. The highest HF energy reduction is noted in T2 quantification with 8× acceleration (ar = 8), with a reduction of ∼50% compared to the ground truth.

To more accurately evaluate the blurriness in our estimation results, we artificially blur the ground-truth tissue property maps with 2-D Gaussian kernels and calculate the resulting HF energy reduction in the blurred images with different Gaussian standard deviations (STD). The results are summarized in Table SV. Comparing Tables SIV and SV, we find that the blurriness caused by our SCQ method is comparable to that caused by Gaussian smoothing with a standard deviation between 0.35 and 0.45 pixels.

TABLE SIV
THE PROPORTION OF ENERGY OF HIGH-FREQUENCY COMPONENTS IN THE ENTIRE SPECTRUM (UNIT: ‰) IN THE GROUND-TRUTH (GT) TISSUE PROPERTY MAPS AND ESTIMATED MAPS BY OUR PROPOSED SCQ METHOD. ar: ACCELERATION RATE.

	GT	SCQ (ours)	
		$ar = 4$	$ar = 8$
T1	2.66	2.25	2.22
T2	7.81	5.39	3.81

TABLE SV
THE PROPORTION OF ENERGY OF HIGH-FREQUENCY COMPONENTS IN THE ENTIRE SPECTRUM (UNIT: ‰) IN THE GROUND-TRUTH TISSUE PROPERTY MAPS BLURRED BY 2-D GAUSSIAN KERNELS WITH STANDARD DEVIATION (STD) FROM 0.3 PIXELS TO 0.5 PIXELS.

	STD (pixels)				
	0.3	0.35	0.4	0.45	0.5
T1	2.54	2.17	1.56	0.97	0.54
T2	7.47	6.42	4.71	2.98	1.71

B. Influence of U-Net Structure

In this section, we study the influence of the number of DUS layers in U-Net on the spatial resolution of tissue property maps. Specifically, we use five DUS layer numbers: 0 (0DUS-U-Net), 1 (1DUS-U-Net), 2 (2DUS-U-Net), 3 (3DUS-U-Net), and 4 (4DUS-U-Net) and calculate the HF energy for estimation result from each layer number. The results calculated from the brain slice shown in Fig. 4 of the main text are summarized in Table SVI. As shown in Table SVI, the highest high-frequency energy (indicating the best spatial resolution) is achieved in the T1 map with 2 DUS layers (2DUS-U-Net) and the T2 map with 3 DUS layers (3DUS-U-Net).

C. Influence of Training Patch Size

In this section, we further study the influence of the patch size used in the end-to-end training on spatial resolution. Specifically, we train our model using patches with three different sizes, i.e., 32×32 (**P32**), 64×64 (**P64**), and 128×128 (**P128**), and calculate the HF energy for estimation result from each patch size. The results calculated from the brain slice shown in Fig. 4 of the main text are summarized in Table SVII. As shown in Table SVII, the use of the patch size of 128×128 pixels leads to the lowest high-frequency energy (indicating the worst spatial resolution) in both T1 and T2 maps, while the use of the patch size of 64×64 pixels leads to the highest high-frequency energy (indicating the best spatial resolution) in both T1 and T2 maps.

TABLE SVI
THE PROPORTION OF ENERGY OF HIGH-FREQUENCY COMPONENTS IN THE ENTIRE SPECTRUM (UNIT: ‰) IN THE GROUND-TRUTH (GT) TISSUE PROPERTY MAPS AND ESTIMATED MAPS BY OUR SCQ METHOD USING U-NETS WITH DIFFERENT NUMBERS OF DOWN- AND UP-SAMPLING (DUS) LAYERS. 0, 1, 2, 3, AND 4DUS-U-NET: U-NET WITH 0, 1, 2, 3, AND 4 DUS LAYERS, RESPECTIVELY.

	GT	SCQ (ours)				
		0DUS-U-Net	1DUS-U-Net	2DUS-U-Net	3DUS-U-Net	4DUS-U-Net
T1	2.66	2.27	2.34	2.42	2.25	2.19
T2	7.81	4.87	4.81	4.57	5.39	5.36

TABLE SVII
THE PROPORTION OF ENERGY OF HIGH-FREQUENCY COMPONENTS IN THE ENTIRE SPECTRUM (UNIT: ‰) IN THE GROUND-TRUTH (GT) TISSUE PROPERTY MAPS AND ESTIMATED MAPS BY OUR SCQ METHOD USING DIFFERENT TRAINING PATCH SIZES. P32, P64, AND P128: PATCH SIZE OF 32×32, 64×64, AND 128×128, RESPECTIVELY.

	GT	SCQ (ours)		
		P32	P64	P128
T1	2.66	2.23	2.25	2.21
T2	7.81	5.25	5.39	4.92

REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 9351, 2015, pp. 234–241.