



Deep Learning Super-Resolution Enables Rapid Simultaneous Morphological and Quantitative Magnetic Resonance Imaging

Akshay Chaudhari¹(✉), Zhongnan Fang², Jin Hyung Lee³, Garry Gold¹, and Brian Hargreaves¹

¹ Department of Radiology, Stanford University, Stanford, CA, USA
{akshaysc,gold,bah}@stanford.edu,

² LVIS Corporation, Palo Alto, CA, USA
zhongnanf@gmail.com

³ Department of Neurology, Stanford University, Stanford, CA, USA
ljinhy@stanford.edu

Abstract. Obtaining magnetic resonance images (MRI) with high resolution and generating quantitative image-based biomarkers for assessing tissue biochemistry is crucial in clinical and research applications. However, acquiring quantitative biomarkers requires high signal-to-noise ratio (SNR), which is at odds with high-resolution in MRI, especially in a single rapid sequence. In this paper, we demonstrate how super-resolution (SR) can be utilized to maintain adequate SNR for accurate quantification of the T_2 relaxation time biomarker, while simultaneously generating high-resolution images. We compare the efficacy of resolution enhancement using metrics such as peak SNR and structural similarity. We assess accuracy of cartilage T_2 relaxation times by comparing against a standard reference method. Our evaluation suggests that SR can successfully maintain high-resolution and generate accurate biomarkers for accelerating MRI scans and enhancing the value of clinical and research MRI.

Keywords: Super-resolution · Quantitative MRI · T_2 relaxation

1 Introduction

Magnetic resonance imaging (MRI) is an excellent non-invasive diagnostic tool to accurately assess pathologies in several anatomies. However, MRI is fundamentally constrained in optimizing for either high-resolution, high signal-to-noise ratio (SNR), or low scan durations. Enhancing one of the three outcomes necessarily degrades one or both of the others. Additionally, unlike other imaging modalities, MR images are qualitative in nature and do not directly correlate to the underlying tissue physiology. While quantitative MRI may help in assessing

tissue biochemistry and longitudinal changes, biomarker accuracy is extremely sensitive to image SNR. Consequently, it is challenging to develop a single MRI method to produce high-resolution morphological images with high quantitative biomarker accuracy in a reasonable scan time, which is tolerable for patients and which ultimately limits cost of the procedure.

1.1 Background

The double-echo in steady-state (DESS) pulse sequence can generate high-resolution images with diagnostic contrast as well as the quantitative biomarker of T_2 relaxation time, in only five-minutes of scan time [1]. The T_2 relaxation time has shown to be sensitive to collagen matrix organization and tissue hydration levels, and is useful for assessing degradation of tissues such as cartilage, menisci, tendons, and ligaments [2]. DESS intrinsically produces two images with independent contrasts. The first echo of DESS (S_1) has a T_1/T_2 weighting while the second echo of DESS (S_2) has a high T_2 weighting.

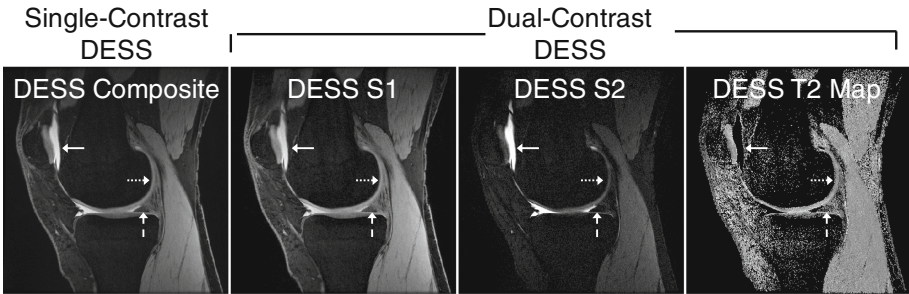


Fig. 1. Compared to the single-contrast DESS, dual-contrast DESS provides additional morphological information and automatic quantitative T_2 relaxation time maps. The separate DESS contrasts (S_1 and S_2) and T_2 maps are useful in assessing the cartilage (dashed arrow), the menisci (dotted arrow), and inflammation (solid arrow). The T_2 maps are not affected by noisy fat-suppression of bony signal.

In previous applications of DESS, the S_1 and S_2 scans are combined during the reconstruction process to produce an output with a singular contrast (herein referred as *single-contrast DESS*) [3]. However, separating the two echoes can provide considerable diagnostic utility since both echoes are sensitive to varying pathologies. Additionally, the two independent-contrast images (herein referred as *dual-contrast DESS*) can be used to analytically determine the tissue T_2 relaxation time, which is a promising biomarker for tissue degradation and OA progression [2,4]. Example images comparing the output of single-contrast DESS and dual-contrast DESS are shown in Fig. 1. Dual-contrast DESS has shown to be useful in diagnostic musculoskeletal imaging of knee as well as in research studies for evaluating OA progression [1,5].

1.2 Motivation

While promising, the dual-contrast DESS is limited in acquiring slices with 1.5 mm section-thickness to maintain adequate SNR for T_2 measurements of the cartilage and menisci. Compared to an in-plane resolution of 0.4×0.4 mm, such a high-section thickness precludes multi-planar reformations, which are essential for evaluating thin knee tissues in arbitrary planes, due to excessive image blurring. An ideal acquisition would provide sub-millimeter section thickness without biasing T_2 measurements. Advances in convolutional neural networks (CNNs) and 3D super-resolution (SR) methods may enable acquisition of slices with a thickness of 1.5 mm followed by retrospectively achieving sub-millimeter resolution, while maintaining SNR for T_2 measurements [6]. However, unlike the single-contrast DESS that has hundreds of datasets publicly available, the dual-contrast DESS is a newer sequence with very limited amounts of high-resolution data available, which makes it challenging to create a SR CNN from scratch. In such scenarios, transfer learning methods may be helpful in overcoming the limitations of a paucity of high-resolution ground-truth dual-contrast DESS training data. Specifically, it may be possible to train a SR CNN initially using single-contrast DESS datasets and subsequently adapt the network to enhance dual-contrast DESS images using limited training data.

Consequently, this study aimed to answer: **1.** Can transfer learning enhance through-plane MRI resolution for the clinically-relevant dual-contrast DESS sequence and **2.** Can transfer learning enable accurate quantitative imaging of the T_2 relaxation time by overcoming SNR limitations commonly faced in high-resolution imaging? The overall goal of this study was to evaluate whether there can be an efficient methodology to create a SR CNN for dual-contrast DESS to produce high-resolution morphological and quantitative images.

2 Related Work

Sparse-coding SR (ScSR) is a state-of-the-art non-deep-learning method that has been used for 2D MRI SR [7]. CNN-based 3D SR MRI has previously shown to transform MRI images with a high section-thickness (low slice-direction resolution) into images with lower section-thickness (high slice-direction resolution) [8]. However, this initial training was performed on single-contrast DESS sequence that does not produce quantitative biomarkers. These scans were originally acquired with a section thickness of 0.7 mm and retrospectively downsampled by a factor of 2x to a section thickness of 1.4 mm to exactly duplicate a faster, lower-resolution acquisition. The SR network was then utilized to evaluate whether the original 0.7 mm scans could be recovered from the 1.4 mm slices. We build upon these results and to extend SR to MRI sequences that can simultaneously produce multiple diagnostic contrasts and quantitative biomarkers.

3 Methods

3.1 Imaging Methodology

We utilized a CNN termed Magnetic Resonance Super-Resolution (MRSR) to extend the SR capabilities of the network initially trained for single-contrast DESS scans. The dual-contrast DESS datasets used in this study were acquired with a slice thickness of 0.7mm (imaging parameters: $TE_1/TE_2/TR = 7/39/23$ ms, matrix size = 416×416 , field of view = 160mm, flip angle = 20° , scan time = 5min, phase encoding parallel imaging = 2x, slices = 160). A slice thicknesses of 0.7mm was maintained for the single-contrast and dual-contrast DESS scans.

A pre-trained network for performing SR with a slice downsampling factor of 2x for the single-contrast DESS sequence was utilized to simultaneously enhance both images from the dual-contrast DESS. This pre-training was performed on image patches with input and output sizes of $32 \times 32 \times 32$ using convolutional filters of size $3 \times 3 \times 3$ and a feature map length of 64. This SR CNN network transforms an input low-resolution image into a residual image through a series of 20 convolutions and rectified linear unit (ReLU) activations [8]. An approximate high-resolution image is generated through the sum of the low-resolution input and the resultant residual using the L2-norm between the approximate and true high-resolution images as the loss function.

3.2 Transfer Learning Training for Dual-Contrast DESS

Since dual-contrast DESS contains an extra image contrast, the initial single-contrast DESS weights for the first convolution layer were duplicated to account for the dual-echoes. Similarly, the final layer output weights were modified to output two echo images instead of one, as shown in Fig. 2. In such a manner, the single-contrast DESS MRSR architecture was modified and subsequently fine-tuned to simultaneously enhance dual-contrast DESS images.

All data processing steps for the single-contrast DESS and MRSR networks were kept unchanged. This included data normalization between 0 and 1, simulation of thicker slices with a 48th-order anti-aliasing filter, a mini-batch size of 50, and a learning rate of 0.0001. All input patches had a size of $32 \times 32 \times 32 \times 2$ with a stride of 16 in the first three directions. Thus, an input image of dimensions $416 \times 416 \times 160$ was divided into 5625 patches. The MRSR network was trained for 10 epochs using 4 NVIDIA Titan 1080Ti graphical processing units.

30 dual-contrast DESS 3D datasets were used for training and 10 for validation. All datasets were collected from patients referred for a clinical MRI following institutional review board approval and informed consent, for ensuring unbiased representation of healthy and pathologic tissues.

Two unique datasets, described below, were tested using the MRSR transfer learning network because it is not currently possible to acquire a single high-resolution dataset that also has high-SNR for accurate quantitative imaging of

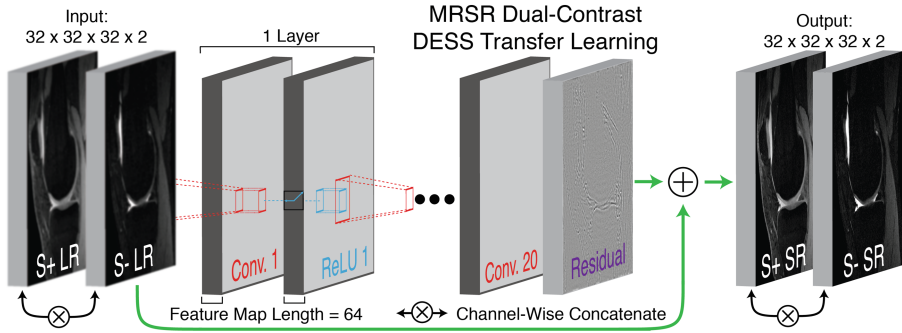


Fig. 2. The schematic of the Magnetic Resonance Super-Resolution (MRSR) network demonstrates how the low-resolution (LR) dual-contrast DESS images are simultaneously transformed into the super-resolution (SR) images.

the T_2 relaxation time. The goal of this two-fold testing was to acquire separate reference high-resolution and high-SNR scans. The dual-contrast DESS could therefore have intermediate SNR for accurate T_2 measurements and the intermediate resolution of the acquisition could be enhanced using MRSR.

Image Quality: Test Cohort 1. This dataset had identical scan parameters to the training dataset. Following the simulation of 2x thicker slices, image quality enhancements were evaluated by comparing the structural similarity (SSIM), peak SNR (pSNR), and root mean square error (RMSE) between the ground truth high-resolution and MRSR images, along with tricubic interpolated (TCI), Fourier interpolated (FI), and sparse coding super-resolution (ScSR) images.

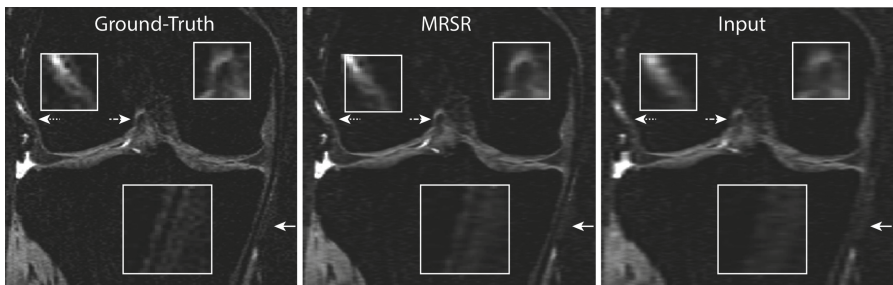


Fig. 3. MRSR coronal reformatted images demonstrate better resolution in the slice-direction (left-right) than the input TCI images, compared to the ground-truth.

T_2 Accuracy: Test Cohort 2. The second dataset had thicker slices (1.6 mm) to maintain a higher SNR for accurate T_2 quantification, since T_2 has a high sensitivity to noise [1]. Accuracy of the T_2 maps was evaluated by comparing the T_2 values in two combined adjacent slices in the medial femoral cartilage of the

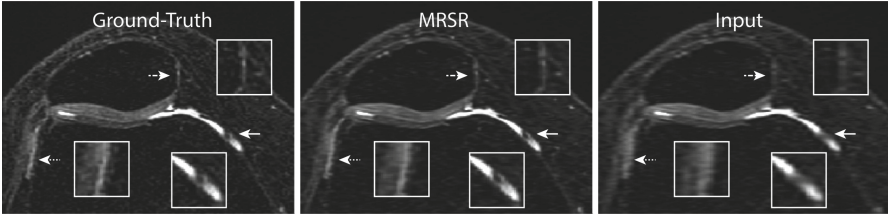


Fig. 4. Example axial reformatted MRSR images, depict finer image details considerably better than the input TCI image compared to the ground-truth.

MRSR, TCI, FI, and ScSR outputs to the ground-truth thick-slice sequences. Segmentation was performed by a reader with 5 years of experience in knee MRI segmentation. T_2 relaxation time differences, coefficients of variation ($CV\%$), and concordance correlation coefficients (CCC) assessed T_2 variations between the methods, compared to the ground truth.

Mann-Whitney U-Tests assessed variations between morphological enhancement metrics as well as T_2 variations for all enhancement methods.

4 Results

Each epoch training duration was approximately 3h for the total of 170,000 training patches. The SSIM, pSNR, and RMSE values between the MRSR, TCI, FI, and ScSR images to the ground-truth are shown in Table 1, where MRSR was significantly superior compared to TCI, FI, and ScSR. Comparisons for T_2 values computed with all methods are shown in Table 2. MRSR had the best image quality metrics, as well as the closest matches for the T_2 values. Despite being compared on a pixel-wise basis, which can have a high sensitivity to noise, the MRSR T_2 values had the lowest inter-method CV of 3% and an excellent CCC of 0.93. There were no statistically significant variations for T_2 for any method compared to the ground truth, likely due to a limited sample size.

Example coronal and axial images of the resolution enhancement are shown in Figs. 3 and 4. The medial collateral ligament (solid arrow, approximately 1 mm thick) is completely blurred out in the input image (Fig. 3), but can be delineated well with MRSR. Similarly, the ligament bundles (dashed arrow) and the synovium (dotted arrow) appeared blurrier in the input image than the MRSR. Figure 4 shows that signal irregularities in medial synovium (solid arrow) delineated better using MRSR than in the input image. The lateral synovial membrane (dotted arrow) also appears thickened in the blurred input image but not in the ground-truth or MRSR, which may incorrectly lead to a diagnosis of synovitis. The patellar cartilage (dashed arrow) appears blurred with diffuse signal heterogeneity in the input image, which may lead to an incorrect cartilage lesion diagnosis. Example T_2 map comparisons (shown in Fig. 5) show minimal differences between the ground-truth and MRSR images, and that the per-pixel difference map has no organized structure, suggesting minimal systematic bias.

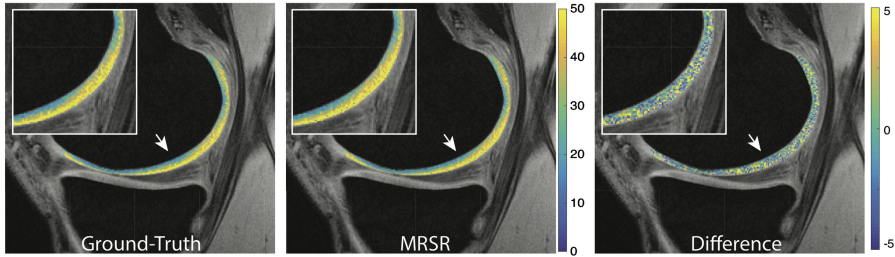


Fig. 5. MRSR T_2 relaxation time maps appear similar and provide a similar spatial distribution of T_2 values compared to the ground-truth. The difference map has no discernible structure, suggesting minimal systematic bias. (Note the different color scale). (Color figure online)

Table 1. Quantitative image quality metrics for both DESS echoes comparing the ground-truth to MRSR, TCI, FI, and ScSR images for test cohort 1. *indicates a significant difference ($p < 0.05$) compared to MRSR. †indicates that all displayed values are multiplied by 10^3 .

Metric	Image	MRSR	TCI	FI	ScSR
SSIM	S_1	0.98 ± 0.01	$0.95 \pm 0.02^*$	$0.92 \pm 0.02^*$	$0.97 \pm 0.01^*$
	S_2	0.98 ± 0.01	$0.96 \pm 0.02^*$	$0.95 \pm 0.02^*$	0.97 ± 0.01
pSNR	S_1	37.7 ± 1.5	$32.5 \pm 3.6^*$	$32.4 \pm 2.8^*$	36.6 ± 1.1
	S_2	38.7 ± 2.0	33.6 ± 4.2	$33.6 \pm 3.5^*$	37.5 ± 1.6
RMSE†	S_1	0.18 ± 0.06	$0.72 \pm 0.56^*$	$0.69 \pm 0.47^*$	0.22 ± 0.05
	S_2	0.13 ± 0.04	0.51 ± 0.40	$0.47 \pm 0.34^*$	0.16 ± 0.05

Table 2. Cartilage T_2 relaxation times for MRSR, TCI, FI, and ScSR compared to the ground-truth using differences and coefficients of variation (CV%) in test cohort 2.

Subject	Ground-truth	MRSR	TCI	FI	ScSR
1	35.2	35.8	36.4	36.1	42.4
2	42.6	44.1	44.4	44.5	50.1
3	27.9	29.1	29.8	29.4	35.9
4	35.3	38.5	39.5	39.0	58.3
5	36.6	38.0	39.0	39.2	46.7
Average	35.5 ± 5.2	37.1 ± 5.4	37.8 ± 5.3	37.6 ± 5.5	46.7 ± 8.4
CV %	N/A	3.1 ± 1.8	4.5 ± 2.2	4.1 ± 2.0	18.8 ± 9.3
Difference	N/A	1.6 ± 1.0	2.3 ± 1.1	2.1 ± 1.1	11.2 ± 6.7
CCC	N/A	0.93	0.87	0.89	0.21

5 Discussion and Conclusion

In this study, we demonstrated that transfer learning can be effectively used to perform SR on MRI sequences with varied contrasts that are used clinically and in epidemiological studies, even with a small training dataset. The dual-contrast DESS sequence was able to maintain a considerably higher resolution and detail than the comparison methods. It is important to note that since the SR was carried out only in one dimension of the 3D dataset, the image enhancements in Figs. 3 and 4 are more prominent in the left-right direction anatomically, which is also the same direction of the displayed images.

The MRSR approach maintained comparable T_2 relaxation times between the ground-truth. A pixel-wise CV of 3% has shown to be adequate for use in OA studies and a CCC of over 0.90 indicated excellent reproducibility compared to the ground-truth [9]. With MRSR, slices can be acquired with a higher section thickness for accurate T_2 measurement, while enabling super-resolution for performing high-resolution MRI scans, which was not possible previously due to SNR limitations. Interestingly enough, all methods over-estimated T_2 values, likely because the thin cartilage has two major divisions (deep and superficial), where the deep cartilage has lower signal. Blurring from the superficial cartilage would increase signal in the deeper layer, leading to a higher T_2 value. Performing layer-wise T_2 values will be important in future studies.

In conclusion, we demonstrated how SR enhanced through-plane resolution in MRI and maintained quantitative accuracy of the T_2 relaxation time biomarker. MRSR outperforms conventional and state-of-the-art resolution enhancement methods and has potential for use in clinical and research studies.

References

1. Chaudhari, A.S., et al.: Five-minute knee MRI for simultaneous morphometry and T_2 relaxometry of cartilage and meniscus and for semiquantitative radiological assessment using double-echo in steady-state at 3T. *J. Magn. Reson. Imaging* **47**, 1328–1341 (2017)
2. Mosher, T.J., Dardzinski, B.J.: Cartilage MRI T_2 relaxation time mapping: overview and applications. *Semin. Musculoskelet. Radiol.* **8**, 355–368 (2004)
3. Peterfy, C.G., Schneider, E., Nevitt, M.: The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis Cartilage* **16**(12), 1433–1441 (2008)
4. Sveinsson, B., Chaudhari, A., Gold, G., Hargreaves, B.: A simple analytic method for estimating T_2 in the knee from DESS. *Magn. Reson. Imaging* **38**, 63–70 (2017)
5. Monu, U.D., Jordan, C.D., Samuelson, B.L., Hargreaves, B.A., Gold, G.E., McWalter, E.J.: Cluster analysis of quantitative MRI T_2 and $T_{1\rho}$ relaxation times of cartilage identifies differences between healthy and ACL-injured individuals at 3T. *Osteoarthritis Cartilage* **25**(October), 1–8 (2016)
6. Kim, J., Kwon Lee, J., Mu Lee, K.: Accurate image super-resolution using very deep convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1646–1654 (2016)

7. Wang, Y.H., Qiao, J., Li, J.B., Fu, P., Chu, S.C., Roddick, J.F.: Sparse representation-based MRI super-resolution reconstruction. *Measurement* **47**, 946–953 (2014)
8. Chaudhari, A.S., et al.: Super-resolution musculoskeletal MRI using deep learning. *Magn. Reson. Med.* (2018)
9. Baum, T., Joseph, G.B., Karampinos, D.C., Jungmann, P.M., Link, T.M., Bauer, J.S.: Cartilage and meniscal T2 relaxation time as non-invasive biomarker for knee osteoarthritis and cartilage repair procedures. *Osteoarthritis Cartilage/OARS* **21**(10), 1474–84 (2013)