

PAPER • OPEN ACCESS

Joint reconstruction and segmentation in undersampled 3D knee MRI combining shape knowledge and deep learning

To cite this article: A Kofler et al 2024 Phys. Med. Biol. 69 095022

View the article online for updates and enhancements.

You may also like

- <u>Brain Tumor Image Segmentation in MRI</u> Image Hapsari Peni Agustin Tjahyaningtijas
- ERSegDiff: a diffusion-based model for edge reshaping in medical image segmentation
- BaiJing Chen, Junxia Wang and Yuanjie Zheng
- <u>Research on Semantic Segmentation of</u> <u>Portraits Based on Improved Deeplabv3 +</u> Kaihui Zhang, Xianhui Liu and Yufei Chen

JOIN US | ESTRO 2024 In-Booth Talks, Demos, & Lunch Symposium

Browse talk schedule >



This content was downloaded from IP address 18.222.108.18 on 30/04/2024 at 16:32

Physics in Medicine & Biology

PAPER

OPEN ACCESS

CrossMark

RECEIVED 16 October 2023

REVISED 11 March 2024

ACCEPTED FOR PUBLICATION

25 March 2024 PUBLISHED

24 April 2024

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Joint reconstruction and segmentation in undersampled 3D knee MRI combining shape knowledge and deep learning

A Kofler¹, C Wald², C Kolbitsch¹, C V Tycowicz³ and F Ambellan³

- $^{\rm 1} \ \ {\rm Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany}$
- $^{\rm 2}$ Department of Mathematics, Technical University of Berlin, Berlin, Germany
- ³ Department of Visual and Data-Centric Computing, Zuse Institute Berlin, Berlin, Germany

E-mail: andreas.kofler@ptb.de

Keywords: deep learning, reconstruction, segmentation, statistical shape model

Abstract

Objective. Task-adapted image reconstruction methods using end-to-end trainable neural networks (NNs) have been proposed to optimize reconstruction for subsequent processing tasks, such as segmentation. However, their training typically requires considerable hardware resources and thus, only relatively simple building blocks, e.g. U-Nets, are typically used, which, albeit powerful, do not integrate model-specific knowledge. Approach. In this work, we extend an end-to-end trainable taskadapted image reconstruction method for a clinically realistic reconstruction and segmentation problem of bone and cartilage in 3D knee MRI by incorporating statistical shape models (SSMs). The SSMs model the prior information and help to regularize the segmentation maps as a final postprocessing step. We compare the proposed method to a simultaneous multitask learning approach for image reconstruction and segmentation (MTL) and to a complex SSMs-informed segmentation pipeline (SIS). Main results. Our experiments show that the combination of joint end-to-end training and SSMs to further regularize the segmentation maps obtained by MTL highly improves the results, especially in terms of mean and maximal surface errors. In particular, we achieve the segmentation quality of SIS and, at the same time, a substantial model reduction that yields a five-fold decimation in model parameters and a computational speedup of an order of magnitude. Significance. Remarkably, even for undersampling factors of up to R = 8, the obtained segmentation maps are of comparable quality to those obtained by SIS from ground-truth images.

1. Introduction and motivation

Knee osteoarthritis (OA) is a widely spread chronic and degenerative joint disease (Lawrence *et al* 2008) that can be assessed from quantitative image-based biomarkers such as the fraction of the apparent bone volume and the total bone tissue volume (Eckstein *et al* 2006), which can be obtained from magnetic resonance imaging (MRI). A Comparison of such biomarkers might give further insights about the prevention and treatment of OA. It is the most common joint disorder with over 250 million people affected worldwide (Vos *et al* 2012)—80% of which involve the tibiofemoral joint of the knee with a high socio-economic burden.

Commonly used x-ray imaging provides fast and relatively low-cost acquisition, but only yields 2D projections with marginal soft tissue contrast leading to limited inter-rater concordance. Recent studies have shown that MRI-derived 3D bone shape is a much better bio-physical parameter for the assessment of OA (Hanik *et al* 2020, Ambellan *et al* 2021b, Ambellan *et al* 2021c). A major challenge for its clinical application is the relatively long acquisition time of 3D MRI. Longer acquisition times can limit the achievable spatial resolution due to possible patients motion during the scan, increase patient discomfort and ultimately inevitably come with associated higher healthcare costs. Undersampling the images in the measurement space, the so-called *k*-space, can accelerate the data acquisition but the reconstruction typically requires advanced regularization methods to obtain images with diagnostic quality. In addition, precise segmentations of the anatomical features of interest

are required both for the determination of the latterly mentioned biomarkers as well as computer-based surgical planning of interventions. Furthermore, any impairment in visual acuity can carry over to downstream tasks such as tissue segmentation and, eventually, affect clinical decision-making. In the last years, deep learning (DL) methods based on unrolled optimization (Monga *et al* 2021) have attracted great attention in the research field of image reconstruction (Adler and Öktem 2018, Aggarwal *et al* 2018, Hammernik *et al* 2018, Schlemper *et al* 2018a, Sriram *et al* 2020, Kofler *et al* 2021). Using task-adapted unrolled neural networks (NNs), the obtained regularization methods cannot only be adapted to the data but also to the specific reconstruction methods as well as to a subsequent task such as image classification or image segmentation (Sun *et al* 2019, Calivá *et al* 2020, Adler *et al* 2021).

In this work, we propose a combination of a task-adapted NNs-based method and statistical shape models (SSMs) for joint reconstruction and segmentation of undersampled 3D knee MR images. We extend an approach similar to Sui *et al* (2021) to a high-dimensional 3D multi-coil MRI reconstruction and segmentation problem and further integrate SSMs to increase the quality of the obtained segmentation maps.

2. Methods

The overall method consists of two stages. First, an end-to-end trained reconstruction- and segmentation network is used to reconstruct the images from the undersampled *k*-space data as well as to deliver a first initial guess of the segmentation. In addition, as a second step, a statistical shape model in which information about shape anatomy is used is further employed to improve the initial segmentation.

2.1. Problem formulation and image reconstruction

In MRI, the data-acquisition process yields the Fourier transform of an image whose contrast is determined by the parameters of the MR sequence and the MR-related parameters of the image, such as relaxation times.

Let $\mathbf{x}_{\text{true}} \in \mathbb{C}^N$ with $N = N_x \cdot N_y \cdot N_z$ denote the vector representation of such an unknown complex-valued 3D MR image and \mathbf{A}_I denote a 3D multi-coil MRI operator which maps the image to its corresponding undersampled *k*-space data. The considered forward problem is given by

$$\mathbf{y}_I = \mathbf{A}_I \mathbf{x}_{\text{true}} + \mathbf{e},\tag{1}$$

where, \mathbf{y}_I denotes the measured *k*-space data in presence of complex-valued Gaussian noise \mathbf{e} . The operator \mathbf{A}_I has the form

$$\mathbf{A}_{I} \coloneqq (\mathbf{I}_{N_{c}} \otimes \mathbf{F}_{I})\mathbf{C}, \tag{2}$$

where $\mathbf{C} = [\mathbf{C}_1, ..., \mathbf{C}_{N_c}]^{\mathsf{I}}$ with $\mathbf{C}_k = \text{diag}(\mathbf{c}_k)$, $\mathbf{c}_k \in \mathbb{C}^N$ contains the stacked coil-sensitivity maps, $\mathbf{F}_I \coloneqq \mathbf{S}_I \mathbf{F}$ denotes the composition of a 3D FFT operator and a binary sub-sampling mask that samples the *k*-space coefficients indexed by $I \subset J = \{1, ..., N\}$ with *J* denoting the entire set of *k*-space coefficients. The operation \otimes denotes the Kronecker product and \mathbf{I}_{N_c} an N_c -dimensional identity operator. Because problem (1) is ill-posed, image reconstruction approaches require the use of advanced regularization techniques. In the following, we describe the here considered regularization scheme in more details.

Let u_{Θ}^{Rec} denote a CNN with trainable parameters Θ which estimates a CNN-based image prior image \mathbf{x}_{CNN} from the undersampled image, i.e. $\mathbf{x}_{\text{CNN}} \coloneqq u_{\Theta}^{\text{Rec}}(\mathbf{x}_I)$, $\mathbf{x}_I \coloneqq \mathbf{A}_I^{\mathsf{H}} \mathbf{y}_I$. Similar as in other works, (Hyun *et al* 2018, Sui *et al* 2021, Kofler *et al* 2021), we formulate the reconstruction problem as

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}_I \mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2} \|\mathbf{x} - \mathbf{x}_{\text{CNN}}\|_2^2,$$
(3)

which, considering x_{CNN} to be fixed, has a unique solution that can be obtained by solving a linear system

$$\mathbf{H}\mathbf{x} = \mathbf{b},\tag{4}$$

$$\mathbf{H} = \mathbf{A}_{I}^{\mathsf{H}} \mathbf{A}_{I} + \lambda \mathbf{I}_{N},\tag{5}$$

$$\mathbf{b} = \mathbf{A}_I^{\mathsf{H}} \mathbf{y}_I + \lambda \, \mathbf{x}_{\text{CNN}}.\tag{6}$$

Approximately solving (4) can for example be achieved with a conjugate gradient method with *M* iterations which we here denote by $f_{M,\lambda}^{CG}$, see e.g. (Hyun *et al* 2018, Kofler *et al* 2021) for more details. By

$$\operatorname{Rec}_{\Theta,\lambda}: \mathbb{C}^{K \cdot N_c} \to \mathbb{C}^N, \tag{7}$$

$$\mathbf{y}_{I} \mapsto (f_{M,\lambda}^{\mathrm{CG}} \circ \boldsymbol{u}_{\Theta}^{\mathrm{Rec}} \circ \mathbf{A}_{I}^{\mathsf{H}})(\mathbf{y}_{I})$$

$$\tag{8}$$

we denote the entire reconstruction network which estimates a CNN-regularized and data-consistent solution \mathbf{x}_{Rec} from the multi-coil *k*-space data $\mathbf{y}_I \in \mathbb{C}^{K \cdot N_c}$, which serves as input for the subsequent multi-class segmentation network u_{Ψ}^{Seg} with parameters Ψ .

2.2. Network architectures

The basic component of the reconstruction and segmentation network is based on the U-Net (Ronneberger *et al* 2015) which has been extensively applied to image reconstruction (Jin *et al* 2017, Hyun *et al* 2018, Hauptmann *et al* 2019, Kofler *et al* 2018, Sriram *et al* 2020, Kofler *et al* 2021) as well as to image segmentation problems (Ambellan *et al* 2019b). As in Kofler *et al* (2018), we hyper-parameterize a U-Net by the number of encoding stages, the number of convolutional layers per stage and the number of filters which is initially applied to the input images by E, C and K, respectively. Based on preliminary experiments, we identified a model configuration of E3 C2 K16 to be powerful enough for the task of image reconstruction, while for the segmentation network, similar as in Ambellan *et al* (2019b), a network with more trainable parameters corresponding to E4C2 K32, was required to yield accurate segmentation maps. In addition, both networks differ in the shape of the kernels which are $3 \times 3 \times 3$ and $5 \times 5 \times 3$ for the reconstruction and the segmentation network, respectively. The number of CG-iterations for solving problem (3) in Rec Θ_{A} is set to M = 12.

The choice of the reconstruction and segmentation networks is based on a trade-off between the available computational hardware, i.e. a 48 GB GPU, and the expressiveness of the respective networks in terms of number of trainable parameters. In general, other studies have observed that increasing the number of trainable parameters is beneficial for the task of image reconstruction with unrolled methods (Kofler *et al* 2022). However, employing deeper networks increases the memory footprint of the end-to-end method. Thus, because the reconstruction module is model-based (due to the use data-consistency module), sacrificing trainable parameters for the reconstruction network seems to be favorable over reducing the number of trainable parameters for segmentation network that performs a more difficult task (semantic segmentation).

Note that the considered reconstruction method corresponds to a special case of the well-known MoDL approach (Aggarwal *et al* 2018), where the number of alternations is fixed to be one here due to the relatively high dimensionality of the considered problem, i.e. the presence of a 3D acquisition operator with multiple receiver coils. This choice is a necessary trade-off chosen between computational requirements to be able to train the entire model-based network $u_{\Psi}^{\text{Seg}} \circ \text{Rec}_{\Theta,\lambda}$ in an end-to-end fashion and using a sophisticated reconstruction network as in Sui *et al* (2021).

2.3. Statistical shape models (SSMs)

As CNNs lack explicit shape knowledge, we opt for SSM-based post-processing to circumvent anatomical implausibilities in the segmentation masks. SSMs are geometric models that describe a collection of semantically similar shapes in a compact way. SSMs jointly represent an average shape of many three-dimensional objects as well as their variation thereof (Ambellan et al 2019a). Given a shape population, SSMs offer a powerful mechanism to efficiently capture the range of anatomical variation. To this end, main trends (a.k.a. modes) of variation around a population-average shape are learned, such that the spanned shape-subspace closely fits the training instances (Kainmüller 2014). More precisely, there exists no closer k-dimensional approximation of an input shape within the span of its collection than the unique representation as deviation from the mean by a linear combination of modes. During post-processing, we project the outcome of u_{Ψ}^{Seg} onto SSMs of the femur and tibia separately, thereby removing any unseen variation and returning the best approximation of the unseen shape within our training set. However, especially for femur and tibia we can assume that there is significantly more variation among individuals in the joint region than at the bone shafts, i.e we expect our shape approximation to fit the observed data better in the shaft than in the joint region. Therefore, when integrating the SSM prediction, we put more trust in the CNN in regions with high-frequent variations (specifically the cartilage interface rims), whereas we rely more on the SSM in regions, where we expect rather low-frequent variations in shape. We employ the common point distribution model that treats shapes as points in the highdimensional space of (stacked) vertex coordinates and captures their distribution via principal component analysis (PCA). Further details and implementation can be found in Kainmüller (2014), Ambellan et al (2019b), and Ambellan et al (2021a), respectively.

2.4. Dataset and experimental set-up

To evaluate the proposed method, we used the open-access OAI-ZIB dataset in Ambellan *et al* (2019b). It consists of 507 3D knee MR images from the OA initiative (http://nda.nih.gov/oai) of shape $N_x \times N_y \times N_z = 384 \times 384 \times 160$ together with their corresponding segmentation maps of the femoral bone and cartilage as well as the tibial bone and cartilage. The segmentations provided by the OAI-ZIB dataset were obtained as manual segmentations performed by experienced users of the Zuse Institute Berlin (ZIB) starting from automated segmentations which were obtained by the method in Seim *et al* (2010). The employed data covers the full spectrum of different OA grades, with a more pronounced focus towards severe cases. The images from the subjects (age 61.87 \pm 9.33 years, 262 male and 245 female) who were with, or at risk, for symptomatic femoral-tibial OA, were acquired on a Siemens 3T Trio scanner using a double echo steady state (DESS)

sequence and have a resolution of $0.36 \times 0.36 \times 0.7 \text{ mm}^3$. For further details, we refer to the supplementary material of the work (Ambellan *et al* 2019b).

Note that the SKM-TEA dataset (Desai et al 2021) also poses an interesting basis for evaluation as it contains raw k-space data. However, we opted to use the OAI-ZIB dataset from (Ambellan et al 2019b) as it additionally contains segmentations of femur and tibia bones and is thus not limited to articular soft tissues. These bones in fact pose distinct challenges (i) due to significantly lower signal strength and contrast to certain neighboring tissues like tendons as well as (ii) due to their physical size extending outside the field-of-view, thus, covering image regions with higher vulnerability to artefacts like signal decay and nonlinear distortions. Another interesting dataset is the recently published K2S-challange-dataset (Tolpadi et al 2023), which however, has two major limitations: first, ground-truth segmentation was carried out with CNNs rather than manually and a human reader quality rating served as an additional exclusion criterion. In other words, experts identified cases that are convenient for a CNN-based segmentation. Second, the subject characteristics suggest a bias of the dataset towards non-arthritic individuals that feature physiological bone and cartilage configurations. Contrary, all segmentation masks of OAI-ZIB are manually segmented and two thirds of them feature severe knee osteoarthritis yielding various unique disease patterns. In addition, by using the dataset from Ambellan et al (2019b), we are able to directly compare the obtained segmentation masks to the ones obtained from the ground-truth images in Ambellan et al (2019b) and therefore asses how much the data-acquisition could in principle be accelerated without sacrificing segmentation accuracy.

The phase information of the images was simulated similarly as in Schlemper *et al* (2018b). From these images, *k*-space data was retrospectively generated using three different undersampling factors $R \in \{8, 12, 16\}$. Undersampling masks along k_y and k_z were chosen according to a Poisson disk sampling pattern (Bridson 2007) which was implemented using SigPy (Ong and Lustig 2019). Coil-sensitivity maps for $N_c = 18$ coils were simulated by employing 3D Gaussian profiles using the function mrisensesim in Muckley *et al* (2020), version 1.0.0. Figure 1 shows a schematic representation of the retrospective *k*-space data simulation used for this study. Our dataset \mathcal{D} consisted of triples ($\mathbf{y}_I, \mathbf{x}_f, \mathbf{s}_t$) of undersampled *k*-space data \mathbf{y}_I , ground-truth images \mathbf{x}_f and target segmentation masks \mathbf{s}_t . As in Ambellan *et al* (2019b), the data was split into 227/26/254 images for training/validation/testing.

2.5. Network training

End-to-end training of the entire network $u_{\Psi}^{\text{Seg}} \circ \text{Rec}_{\Theta,\lambda}$ was carried out by minimizing a convex combination of the mean squared error (MSE) and the DICE similarity coefficient (DSC)-based loss function

$$\mathcal{L}_{\alpha}(\Theta, \Psi) = \frac{1}{|\mathcal{D}|} \sum_{(\mathbf{y}_{l}, \mathbf{x}_{t}, \mathbf{s}_{t}) \in \mathcal{D}} [(1 - \alpha) \operatorname{MSE}(\operatorname{Rec}_{\Theta}(\mathbf{y}_{l}), \mathbf{x}_{f}) + \alpha(1 - \operatorname{DSC}(u_{\Psi}^{\operatorname{Seg}}(\operatorname{Rec}_{\Theta, \lambda}(\mathbf{y}_{l})), \mathbf{s}_{t}))].$$
(9)

for $\alpha = 1/3$, 2/3, 1 which was also used in Sui *et al* (2021). Thereby, for $\alpha = 1$, the regularization is learned to be optimal with respect to the subsequent segmentation task given by u_{Ψ}^{Seg} , while for $0 < \alpha < 1$, also the reconstruction-error with respect to the L_2 -norm is minimized. Similar as in other works, e.g. (Sun *et al* 2019, Tolpadi *et al* 2023), Rec $_{\Theta,\lambda}$ and u_{Ψ}^{Seg} were separately pre-trained for the corresponding tasks, i.e. by minimizing the MSE- and the DSC-based loss functions, respectively, and by training for 36 epochs with the ADAM-optimizer (Kingma and Ba 2014) with learning rates of 10^{-4} and 10^{-5} , respectively. For each α , the entire network Rec $_{\Theta,\lambda} \circ u_{\Psi}^{\text{Seg}}$ was fine-tuned for further 16 epochs by decreasing the respective initial learning rates by a factor of five. Due to GPU-memory constraints, training had to be performed on image patches of shape $N'_x \times N'_y \times N'_z = 256 \times 256 \times 160$. Pre-training took about two days for each sub-network, while the end-to-end fine-tuning took further two days for each α . The GPU-memory allocation amounted to approximately 18 GB/27 GB for pre-training Rec $_{\Theta,\lambda}$ and u_{Ψ}^{Seg} , respectively, and 47 GB for fine-tuning the network Rec $_{\Theta,\lambda} \circ u_{\Psi}^{\text{Seg}}$. All experiments were run on an NVIDIA RTX A6000 with a 48 GB GPU. As no validation is needed to train SSMs, both were trained on the whole training set.

2.6. Experiments and evaluation

In the following we evaluate the NNs-based method in terms of reconstruction quality as well as accuracy of the estimated segmentation masks. More precisely, we investigate the impact of the single components of the proposed pipeline as well as the training scheme, i.e.

- (1) The importance of end-to-end training over model-agnostic decoupled pre-training.
- (2) The importance of employing the statistical shape models.



Figure 1. A schematic illustration of the retrospective k-space data generation process as well as the used end-to-end trainable reconstruction and segmentation network Rec $_{\Theta,\lambda} \circ u_{\Psi}^{Seg}$ (red) and the SSMs-based post-processing. The undersampled k-space data is retrospectively simulated according to (1). The input of the network is the zero-filled reconstruction. After a few warm-up iterations for approximately solving the normal equations, the 3D U-Net u_{Θ}^{Rec} is applied to obtain the CNN-prior \mathbf{x}_{CNN} . The image \mathbf{x}_{CNN} is used to regularize the reconstruction problem (3) which is then approximately solved using a conjugate gradient module. From this data-consistent solution, a further 3D segmentation network u_{Ψ}^{Seg} is applied to estimate a segmentation mask. The reconstruction and segmentation network can be trained to jointly reconstruct the images from the undersampled k-space data and subsequently segment the reconstructions, while the SSM-based post-processing (blue) is a subsequent step that improves the obtained segmentations.

For the obtained reconstructions, we report the peak signal-to-noise ratio (PSNR), normalized root mean squared error (NRMSE) and structural similarity index measure (SSIM) (Wang *et al* 2004), while for the segmentation maps, we provide surface distance-based metrics which provide more insightful differentiation of tissue shape in addition to the DSC. More precisely, the metrics used to assess the quality of the obtained segmentations are given by

$$DSC = 100 \cdot \frac{2|B \cap A|}{|B| + |A|}$$
(10)

$$ASD = \frac{1}{N_{\partial A} + N_{\partial B}} \left(\sum_{i=1}^{N_{\partial A}} \min_{b \in \partial B} |a_i - b|_2 + \sum_{j=1}^{N_{\partial B}} \min_{a \in \partial A} |b_j - a|_2 \right)$$
(11)

$$MSD = \max\left(\max_{a \in \partial A} \min_{b \in \partial B} |a - b|_2, \max_{b \in \partial B} \min_{a \in \partial A} |b - a|_2\right),$$
(12)

where A denotes the set of ground-truth voxels and B denotes the resulting segmentation mask, ∂A and ∂B represent the boundary of A and B, i.e. the set of voxels with at least one neighbor being not part of the respective segmentation mask. The number of voxels on the boundaries ∂A and ∂B is written as $N_{\partial A}$ and $N_{\partial B}$, respectively. The segmentation accuracy is evaluated for femoral/tibal bone (FB/TB) and femoral/tibial cartilage (FC/TC)

	,	C · · ·			
		DT	E2E, $\alpha = \frac{1}{3}$	E2E, $\alpha = \frac{2}{3}$	E2E, $\alpha = 1$
	PSNR	38.89 ± 2.99	39.12 ± 3.06	$\textbf{38.84} \pm \textbf{2.99}$	$\textbf{38.18} \pm \textbf{3.06}$
R = 8	NRMSE	$\textbf{0.18} \pm 0.01$	$\textbf{0.18} \pm 0.01$	$\textbf{0.18}\pm0.01$	0.21 ± 0.01
	SSIM	0.82 ± 0.04	$\textbf{0.82}\pm0.04$	$\textbf{0.82}\pm0.04$	0.80 ± 0.04
	PSNR	36.16 ± 2.75	$\textbf{36.35} \pm 2.84$	36.12 ± 2.76	35.52 ± 2.87
R = 12	NRMSE	0.21 ± 0.01	0.21 ± 0.01	$\textbf{0.20}\pm0.01$	0.23 ± 0.01
	SSIM	0.78 ± 0.05	0.78 ± 0.05	$\textbf{0.79}\pm0.05$	0.76 ± 0.05
	PSNR	34.23 ± 2.59	$\textbf{34.46} \pm 2.68$	34.26 ± 2.57	33.74 ± 2.73
R = 16	NRMSE	$\textbf{0.22}\pm0.01$	0.23 ± 0.01	0.22 ± 0.01	0.25 ± 0.01
	SSIM	0.76 ± 0.06	$\textbf{0.76} \pm 0.06$	$\textbf{0.76}\pm0.05$	0.74 ± 0.06

Table 1. Reconstruction results for undersampling factors R = 8, 12, 16 obtained with the investigated baseline MTL by decoupled training (DT) as well as the joint end-to-end training (E2E) with the loss function in (9) for $\alpha \in \left\{\frac{1}{2}, \frac{2}{3}, 1\right\}$.

using the DSC, average surface distance (ASD) and maximum distance (MSD). All these measures are symmetric and allow to assess the global, volumetric (DSC) as well as the local, boundary-related (ASD, MSD) segmentation quality. As methods of comparison, we use the multitasking learning (MTL) approach in Sui *et al* (2021) and SIS (Ambellan *et al* 2019b). Since our work extends MTL by incorporating SSMs, we abbreviate it by MTL+SSM. Further, similar to Calivá *et al* (2020), we evaluate the impact of the joint end-to-end training (E2E) over the decoupled training (DT). Note that while MTL and our approach have the same overall number of 17821797 trainable parameters for the segmentation network, the SIS pipeline has 81203846 trainable parameters.

3. Results

3.1. Reconstruction results

Table 1 summarizes the results of the obtained image reconstructions for the baseline MTL method with DT and E2E training for $\alpha \in \{1/3, 2/3, 1\}$ and $R \in \{8, 12, 16\}$. These reconstructions also serve as input for the subsequent SSM and the entire SIS pipeline. By comparing E2E with $\alpha = 1$ to DT, we see that the quality of the obtained reconstructions tends to decrease with increasing α . For E2E $\alpha = 1/3$ and $\alpha = 2/3$, the quality of the obtained reconstructions is on the other hand comparable to DT.

3.2. Segmentation results

Table 2 summarizes the quality of the obtained segmentation maps for MTL, our proposed approach MTL +SSM and MTL+SIS. Obtaining the segmentations with the different methods takes $\ll 0.1$ s, 30 s and 10 m, respectively. Figure 2 shows an example of images and corresponding segmentation masks obtained with the NNs-based method, once for the DT and for the E2E for $\alpha = k/3$ for k = 1, 2, 3. We again clearly see how the joint end-to-end training for $\alpha = 1$ improves the segmentation masks over the ones obtained by DT although the obtained reconstruction exhibit a slightly lower PSNR, NRMSE and SSIM. In this particular example, the wrong misclassification in the region of the tibia, which is visible for DT (actually for all *R*) could be corrected by E2E for all α , which however introduced some artefacts towards the border of the image.

By evaluating table 2 and comparing the segmentations obtained from MTL for DT and E2E with $\alpha = 1$, we see that although the segmentation quality tends to slightly increase in terms of DSC (except for TC), the ASD and MSD quite consistently increased.

Our proposed combination of MTL+SSMs significantly improved the obtained segmentation maps and shows a consistent improvement of E2E compared to DT with respect to all measures. An example showcasing typical differences between results of MTL, MTL+SIS and the proposed MTL+SSM is given in figure 3 clearly showing that shape knowledge can be used to enhance the outcome of MTL. Further, by comparing the segmentation maps provided by the MTL+SSM for E2E to the ones obtained by MTL+SIS, we see that the results are very similar, meaning that MTL+SSM shows nearly the same performance as SIS, which is takes 20 times longer than MTL+SSM for obtaining the segmentations. Finally, we point out that the segmentation metrics obtained by the SIS for an undersampling factor of R = 8 are close to the ones which were reported in table 5 in Ambellan *et al* (2019b), i.e. on the ground-truth images. This result remarkably implies that with the proposed MTL+SSM, it is possible to accelerate the scanning process by a factor of R = 8 on the one hand and to accelerate the segmentation process compared to SIS on the other hand, without sacrificing segmentation accuracy.

Table 2. Segmentation accuracy of MTL (Sui *et al* 2021), our MTL+SSM and MTL+SIS (Sui *et al* 2021)+(Ambellan *et al* 2019b) for both DT and E2E training strategies, where $\alpha = 1$ for undersampling factor R = 8 for E2E. The first row shows results of the segmentation network u_{Ψ}^{Seg} of MTL trained along with the reconstruction NN. The middle part shows results with u_{Ψ}^{Seg} as above followed by SSM-based post-processing as in SIS, correcting errors as shown in figure 3. The lower part shows results for the SIS segmentation pipeline incorporating statistical shape knowledge applied to the reconstructed images.

		DT			E2E, $\alpha = 1$		
R = 8		DSC (%)	ASD (mm)	MSD (mm)	DSC (%)	ASD (mm)	MSD (mm)
MTL	FB	95.7 ± 2.82	2.81 ± 2.21	57.20 ± 15.20	98.3 ± 0.94	3.36 ± 3.56	55.87 ± 18.66
	TB	92.8 ± 3.65	2.80 ± 1.47	45.05 ± 20.28	94.1 ± 1.87	15.51 ± 1.87	63.50 ± 5.58
	FC	87.6 ± 2.40	0.21 ± 0.09	8.75 ± 12.21	89.2 ± 2.53	0.60 ± 0.72	13.50 ± 19.61
	TC	80.7 ± 5.84	3.41 ± 3.41	41.64 ± 26.77	66.0 ± 7.43	29.77 ± 5.03	70.14 ± 5.69
MTL+SSM(Ours)	FB	98.1 ± 0.47	0.25 ± 0.08	4.21 ± 2.24	$\textbf{98.5}\pm0.32$	0.19 ± 0.05	3.36 ± 1.78
	TB	93.7 ± 3.75	0.88 ± 0.54	14.10 ± 6.16	97.6 ± 1.28	0.32 ± 0.18	5.24 ± 3.33
	FC	87.6 ± 2.40	0.21 ± 0.06	5.47 ± 2.45	89.2 ± 2.53	0.18 ± 0.05	$\textbf{5.37} \pm 2.43$
	TC	83.8 ± 4.27	0.25 ± 0.11	5.03 ± 2.12	$\textbf{85.5} \pm 4.45$	0.22 ± 0.10	$\textbf{4.86} \pm 1.98$
MTL+SIS	FB	98.4 ± 0.31	0.21 ± 0.05	3.88 ± 1.51	98.4 ± 0.31	0.20 ± 0.05	$\textbf{3.35} \pm 1.52$
	TB	$\textbf{98.1}\pm0.59$	0.26 ± 0.16	$\textbf{4.77} \pm 3.22$	$\textbf{98.1}\pm0.67$	$\textbf{0.27} \pm 0.20$	4.87 ± 3.31
	FC	89.4 ± 3.94	0.17 ± 0.07	$\textbf{5.37} \pm 2.25$	$\textbf{89.5}\pm3.85$	0.17 ± 0.07	$\textbf{5.37} \pm 2.27$
	TC	84.5 ± 4.21	0.23 ± 0.10	5.45 ± 2.62	84.6 ± 4.29	$\textbf{0.22}\pm0.10$	5.32 ± 2.32



Figure 2. Images and segmentation for different undersampling factors R = 8, 12, 16 obtained by MTL with DT as well as joint endto-end training (E2E) using the the loss function in (9) for $\alpha \in \left\{\frac{1}{3}, \frac{2}{3}, 1\right\}$. E2E partially improves the obtained segmentation compared to DT in terms of partially wrongly segmented regions in the tibia (yellow arrows). However, residual artefacts at the lower tibia shaft, which we attribute to the necessity to train on patches, were either not entirely removed correctly or further amplified (orange arrows). However, using the proposed MTL+SSM—as shown in a second step later—can easily account for this issue and successfully removes the remaining segmentation artefacts.





4. Discussion

In the following, we discuss the results obtained with the proposed MTL+SSM method as well as similarities and differences to some other recently published works.

4.1. Image reconstruction

From table 1 we can assess the effect of the end-to-end training in the obtained intermediate reconstructions which serve as input for the subsequent segmentation pipelines. Table 1 indicates that the joint training of the reconstruction and the segmentation networks has a slight negative influence on the obtained reconstruction in terms of PSNR, SSIM and NRMSE. The extent of this influence clearly depends on the hyper-parameter α which, as can be seen from the loss-function in (9), balances between the importance of the segmentation mask and the quality of the reconstruction. However, we see that the influence of α is perhaps less pronounced than expected, since the obtained reconstructions only slightly differ in terms of the reported measures as well as visually, see figure 2. This result can most probably be explained by the fact that the reconstructions are constraint by the chosen model used for regularization, i.e. they are the solutions of the variational problem (3). Additionally, an interplay of the the regularization parameter λ and the hyper-parameter α could be responsible for the perhaps unexpectedly small difference between the obtained reconstructions for E2E for the different choices of α . On the other hand, from the point-wise error images in figure 2 we clearly see that the fully task-adapted joint training, i.e. E2E with $\alpha = 1$, yields somewhat noisier reconstructions with potentially sharper edges which might thus be better suitable for the subsequent segmentation task, which suggests that the obtained reconstructions at least up to some point differ in terms of local image features.

4.2. Image segmentation

From figure 2 it is visible how the joint end-to-end training improves the obtainable segmentation in terms of underestimation of the tibia (R = 8) as well as misclassification of the tibia (R = 12 and R = 16). However, residual artefacts in the region of the lower tibia shaft seem to remain despite the joint end-to-end training. These residual artefacts can possibly be attributed to the necessity to train on patches, which is a compromise necessary to be able to use highly expressive networks for the reconstruction/segmentation. However, from figure 3, we see that these residual artefacts can easily be corrected by employing the SSMs as a final regularization step. By carefully evaluating table 2, we can make the following observations. First, by comparing MTL for DT versus E2E with $\alpha = 1$, we can see that employing end-to-end training quite consistently improves the obtainable segmentation maps with respect to all measures except for TC (which is also consistent with the visual results shown in figure 2). This result proves the superiority of the segmentation maps obtained by end-toend training over decoupled training. Second, by comparing MTL with E2E, $\alpha = 1$ to the proposed combination of MTL+SSM with E2E, $\alpha = 1$, we see that the employed SSM can compensate for the residual artefacts in the lower shaft of the tibia (66% DSC versus 85.5% DSC) and additionally significantly improves the distance-based metrics ASD and MSD. Finally, the comparison of the proposed MTL+SSM with E2E to MTL+SIS with E2E, both for $\alpha = 1$, reveals that both methods perform comparably well in terms of all metrics. However, MTL +SSM yields the segmentation maps in only approximately 30 s compared to the 10 m required by the computationally more demanding SIS pipeline, which alternates betweend the application of 2D and 3D CNNs and the fitting of statistical shape models. Thus, the proposed method yields an undersampling of approximately 20 times in terms of image processing time. Figure 3 shows an example that visually confirms the observations made from table 2.

4.3. Related works

The approach presented in Calivá *et al* (2020) shares our application example but assumes a single-coil data acquisition. It consists of a single 3D U-Net which processes a zero-filled reconstruction with a shared encoder path and two distinct decoder branches, which address the reconstruction and the segmentation, respectively. However, the employed reconstruction network (merely a 3D U-Net) does not employ any method for ensuring data consistency (DC) of the solution, which is typically used in today's state-of-the-art methods for image reconstruction, see e.g. Ongie *et al* (2020).

The work in Sun *et al* (2019) uses a joint data-consistent reconstruction and segmentation approach for 2D brain MRI, where the physics-informed (single-coil) reconstruction method (Schlemper *et al* 2018a) is followed by a 2D U-Net for segmentation. They report improved results in both the segmentation and the reconstruction task when compared to decoupled training. Similarly, the authors in (Karkalousos *et al* 2023) propose a multi-task learning network for joint image reconstruction and segmentation in 2D brain MRI, by also involving the use of multiple receiver coils. Thereby, the application of a data-consistent reconstruction network based on cascades of independently recurrent inference machines (Karkalousos *et al* 2022) and an Attention U-Net

(Oktay *et al* 2018) are alternated several times to yield a final reconstruction and segmentation. However, endto-end training of reconstruction networks similar to (Schlemper *et al* 2018a), (Aggarwal *et al* 2018), (Hammernik *et al* 2018), (Karkalousos *et al* 2022), (Geng *et al* 2023) that are based on algorithm unrolling (Monga *et al* 2021) and which often alternate several times between CNN- and DC-blocks, becomes computationally prohibitive for large reconstruction problems, e.g. the here considered 3D and multi-coil acquisitions. As an alternative, one can consider methods in which only one single NN-based image-prior is obtained and used for regularization, see e.g. (Hyun *et al* 2018, Kofler *et al* 2021). This choice of architecture for the reconstruction module was for example used in the work (Sui *et al* 2021), followed by a U-Net for image segmentation of liver and renal lesions in 2D MRI.

The work in Acar *et al* (2022) compares several reconstruction and segmentation networks for joint training. They introduce a training stabilization technique by successively increasing undersampling across training epochs. In their experiments, the global MRI reconstruction suffers from joint training but the quality in the area of diagnostic interest improves. However, none of the discussed methods incorporates shape knowledge into their frameworks.

Finally, the winners of the K2S-challenge (Tolpadi *et al* 2023) use a similar approach to Sui *et al* (2021), but with no DC-module and with two nnU-Nets (Isensee *et al* 2021) instead of U-Nets. Interestingly, the challenge organizers reported that the winning approach yields the best metrics with respect to the segmentation task, while the intermediate reconstructions, which, by not employing any DC-module, are not enforced to be data-consistent, were outperformed by several other teams. This suggests that, aiming for accurate segmentations with the method in Isensee *et al* (2021) seems to be somewhat necessarily contradictory to obtaining data-consistent reconstructions that can be additionally analyzed by radiologists. Further, we note that all mentioned methods employ a relatively simple U-Net for the task of segmentation, while there nowadays exits more sophisticated and complex models. For example, the method in Ambellan *et al* (2019b) alternatingly applies 2D/3D-U-Nets and fittings of SSMs which act as a regularization method and allow for an effective region of interest restriction. These SSMs solely rely on the shape of the anatomy under study but not on the underlying imaging modality.

4.4. Limitations

Clearly, the main limitation of the proposed method is given by the required hardware components necessary for end-to-end training. Despite the use of a relatively powerful GPU with 48 GB of memory, training could only be carried out on patches of the entire images, which we believe to be the reason for the residual artefacts visible in figures 2 and 3 for MTL. However, we point out that there might be several possibilities to address this issue, both from a implementation point of view, e.g. applying activation checkpointing, or by different choices in the architecture design, e.g. 2.5D U-Nets (Zimmermann *et al* 2023) instead of 3D U-Nets for the reconstruction network. Additionally, we point out that due the lack of raw *k*-space measurements in the OAI-ZIB data, the *k*-space data for this study had to be retrospectively simulated, which limits the potential impact of this study from a practical and clinical perspective.

4.5. Outline and future work

Although the presented approach yields accurate results in terms of segmentation accuracy which are close to the ones that can be obtained by SIS from the ground-truth images, we can identify several research directions which might be worth to be pursued. Based on the visible success of the end-to-end training strategy, it might be desirable to include the regularization by the employed SSMs as an additional block which can be be backpropagated through and thus allows for the end-to-end training of the reconstruction network, the segmentation network as well as the regularization by the SSMs. To be able to achieve this, which necessarily comes at the cost of additional hardware requirements, it might be necessary to further investigate what compromises can or have to be made for the choice of the reconstruction and segmentation networks. Addressing these questions would reduce the needed hardware requirements to be able to train such methods and ultimately increase their applicability to different problems as well as to different datasets.

Further, we note that our baseline segmentation network consists of a relatively simple 3D U-Net. Other recent and more evolved methods, such as the anomaly-aware 3D U-Net presented in Woo *et al* (2024), could also be applied to further extend the method.

Another aspect that could be addressed is the employed sampling pattern. While in this work we employed a sampling pattern based on Poisson disk sampling (Bridson 2007), other patterns, e.g. pseudo-Gaussian sampling could also be employed (Pandit *et al* 2016). Nevertheless, note that while an adaptation to different sampling patterns on a Cartesian grid can easily be accomplished, for an extension to non-Cartesian sampling patterns such as radial (Lauterbur 1973) or spiral (Meyer *et al* 1992), one would need to appropriately address the resulting challenges concerning the more complex image reconstruction block. Last, we note that although the

approach was presented for the reconstruction and the segmentation of MR images focusing on femoral as well as tibial bone and cartilage, the method's structure is general. Thus, we expect it to be applicable for the reconstruction and segmentation of other organs and structures as well.

5. Conclusion

We proposed a combination of task-adapted NNs-based approach and statistical shape models for a realistic 3D multi-coil knee MRI reconstruction and segmentation problem from undersampled *k*-space data and compared it to a MTL (Sui *et al* 2021) as well as to a shape-informed segmentation (SIS) pipeline (Ambellan *et al* 2019b).

Our experiments suggest that it is still possible to further improve over end-to-end trainable task-adapted NNs such as MTL by incorporating SSMs. Further, the segmentation masks of the proposed MTL+SSM are on par with the ones of the segmentation pipeline (SIS) but are obtained in only 30 s compared to 10 m. Last, we report that for an undersampling factor of R = 8, MTL+SIS as well as the proposed MTL+SSM are able to segment the undersampled images nearly as accurately as SIS from the ground-truth images (see (Ambellan *et al* 2019b) table 5).

Acknowledgments

The authors acknowledge the support of the German Research Foundation (DFG) under Germany's Excellence StrategyThe Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689) as this work was initiated during the Hackathon event 'Maths Meets Image', Berlin, March 2022, which was part of the MATH+ Thematic Einstein Semester on 'Mathematics of Imaging in Real-World Challenges'.

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors. https://pubdata.zib.de.

References

Acar M, Çukur T and Öksüz İ 2022 5th Int. Workshop, MLMIR 2022, Held in Conjunction with MICCAI 2022, Singapore, September 22, 2022, Proc.Segmentation-aware MRI reconstruction machine learning for medicalimage reconstruction (Springer) (https://doi.org/10.1007/ 978-3-031-17247-2_6)

Adler J, Lunz S, Verdier O, Schönlieb C B and Öktem O 2021 Task adapted reconstruction for inverse problems *Inverse Problems* 38 075006 Adler J and Öktem O 2018 Learned primal-dual reconstruction *IEEE Trans. Med. Imaging* 37 1322–32

Aggarwal H K, Mani M P and Jacob M 2018 Modl: model-based deep learning architecture for inverse problems *IEEE Trans. Med. Imaging* 38 394–405

Ambellan F, Hanik M and von Tycowicz C 2021a Morphomatics: geometric morphometrics in non-euclidean shape spaces

Ambellan F, Lamecker H, von Tycowicz C and Zachow S 2019a Biomedical Visualisation ed P M Rea (Springer) Vol 3, pp 67-84

Ambellan F, Tack A, Ehlke M and Zachow S 2019b Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: data from the osteoarthritis initiative *Med. Image Anal.* **52** 109–18

Ambellan F, Zachow S and Tycowicz C v 2021b Geodesic B-score for improved assessment of knee osteoarthritis *Int. Conf. on Information Processing in Medical Imaging* (Springer) pp 177–88

Ambellan F, Zachow S and von Tycowicz C 2021c Rigid motion invariant statistical shape modeling based on discrete fundamental forms Med. Image Anal. 73 102178

Bridson R 2007 Fast poisson disk sampling in arbitrary dimensions ACM SIGGRAPH Sketches 10 22-es

Calivá F, Leynes A P, Shah R, Bharadwaj U U, Majumdar S, Larson P E and Pedoia V 2020 Breaking speed limits with simultaneous ultra-fast MRI reconstruction and tissue segmentation *Medical Imaging with Deep Learning PMLR* pp 94–110

Desai A D et al 2021 Skm-tea: a dataset for accelerated MRI reconstruction with dense image labels for quantitative clinical evaluation 35th Conf. on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)

- Eckstein F, Burstein D and Link T M 2006 Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis NMR in Biomedicine: an International *J. Devoted Dev. Appl. Magn. Resona. In vivo* 19 822–54
- Geng C, Jiang M, Fang X, Li Y, Jin G, Chen A and Liu F 2023 Hfist-net: high-throughput fast iterative shrinkage thresholding network for accelerating MR image reconstruction *Comput. Methods Programs Biomed.* 232 107440

Hammernik K, Klatzer T, Kobler E, Recht M P, Sodickson D K, Pock T and Knoll F 2018 Learning a variational network for reconstruction of accelerated MRI data *Magn. Reson. Med.* **79** 3055–71

 Hanik M, Hege H C, Hennemuth A and von Tycowicz C 2020 Nonlinear regression on manifolds for shape analysis using intrinsic Bézier splines International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2020 (Springer) pp 617–26
 Hauptmann A, Arridge S, Lucka F, Muthurangu V and Steeden J A 2019 Real-time cardiovascular MR with spatio-temporal artifact

suppression using deep learning-proof of concept in congenital heart disease Magn. Reson. Med. 81 1143–56

Hyun C M, Kim H P, Lee S M, Lee S and Seo J K 2018 Deep learning for undersampled MRI reconstruction *Phys. Med. Biol.* 63 135007 Isensee F, Jaeger P F, Kohl S A, Petersen J and Maier-Hein K H 2021 NNU-Net: a self-configuring method for deep learning-based biomedical image segmentation *Nat. Methods* 18 203–11 Jin K H, McCann M T, Froustey E and Unser M 2017 Deep convolutional neural network for inverse problems in imaging IEEE Trans. Image Process. 26 4509–22

Kainmüller D 2014 Deformable Meshes for Automatic Segmentation Deformable Meshes for Medical Image Segmentation (Springer) (https://doi.org/10.1007/978-3-658-07015-1_3)

- Karkalousos D, Isgum I, Marquering H and Caan M W 2023 Multitask learning for accelerated-MRI reconstruction and segmentation of brain lesions in multiple sclerosis Medical Imaging with Deep Learning
- Karkalousos D, Noteboom S, Hulst H E, Vos F M and Caan M W 2022 Assessment of data consistency through cascades of independently recurrent inference machines for fast and robust accelerated MRI reconstruction *Phys. Med. Biol.* **67** 124001

Kingma D P and Ba J 2014 Adam: a method for stochastic optimization arXiv:1412.6980

- Kofler A, Haltmeier M, Kolbitsch C, Kachelrieß M and Dewey M 2018 A u-nets cascade for sparse view computed tomography *Int. Workshop* on Machine Learning for Medical Image Reconstruction pp 91–9
- Kofler A, Haltmeier M, Schaeffter T and Kolbitsch C 2021 An end-to-end-trainable iterative network architecture for accelerated radial multi-coil 2D cine MR image reconstruction *Med. Phys.* 48 2412–25
- Kofler A, Schaeffter T and Kolbitsch C 2022 The more the merrier?—on the number of trainable parameters in iterative neural networks for image reconstruction *Proc. of the Joint Annual meeting of ISMRM-ESMRMB and SMRT 31st Annual Meeting, London*
- Lauterbur P C 1973 Image formation by induced local interactions: examples employing nuclear magnetic resonance *Nature* 242 190–1 Lawrence R C *et al* 2008 Estimates of the prevalence of arthritis and other rheumatic conditions in the united states: part ii *Arthritis Rheumatism* 58 26–35

Meyer C H, Hu B S, Nishimura D G and Macovski A 1992 Fast spiral coronary artery imaging Magn. Reson. Med. 28 202–13

- Monga V, Li Y and Eldar Y C 2021 Algorithm unrolling: interpretable, efficient deep learning for signal and image processing *IEEE Signal Process Mag.* 38 18–44
- Muckley M J, Stern R, Murrell T and Knoll F 2020 Torchkbnufft: a high-level, hardware-agnostic non-uniform fast fourier transform in ISMRM *Workshop on Data Sampling and Image Reconstruction*
- Oktay O et al 2018 Attention U-net: Learning where to look for the pancreas arXiv:1804.03999
- Ong F and Lustig M 2019 Sigpy: a python package for high performance iterative reconstruction *Proc. of the ISMRM 27th Annual Meeting, Montreal, Quebec, Canada* vol 4819
- Ongie G, Jalal A, Baraniuk C A M R G, Dimakis A G and Willett R 2020 Deep learning techniques for inverse problems in imaging *IEEE J. Sel. Areas Inf. Theory* **1** 39–56
- Pandit P, Rivoire J, King K and Li X 2016 Accelerated t1 ρ acquisition for knee cartilage quantification using compressed sensing and datadriven parallel imaging: a feasibility study *Magn. Reson. Med.* 75 1256–61
- Ronneberger O, Fischer P and Brox T 2015 Int. Conf. on Medical image computing and computer-assisted interventionU-net: convolutional networks for biomedical image segmentation (Springer) pp 234–41
- Schlemper J, Caballero J, Hajnal J V, Price A N and Rueckert D 2018a A deep cascade of convolutional neural networks for dynamic MR image reconstruction *IEEE Trans. Med. Imaging* 37 491–503
- Schlemper J, Castro D C, Bai W, Qin C, Oktay O, Duan J, Price A N, Hajnal J and Rueckert D 2018b Bayesian deep learning for accelerated MR image reconstruction in International Workshop on *Machine Learning for Medical Image Reconstruction* (Springer) pp 64–71
- Seim H, Kainmueller D, Lamecker H, Bindernagel M, Malinowski J and Zachow S 2010 Model-based auto-segmentation of knee bones and cartilage in MRI data *Proc. MICCAI Workshop Medical Image Analysis for the Clinic* pp 215–23
- Sriram A, Zbontar J, Murrell T, Defazio A, Zitnick C L, Yakubova N, Knoll F and Johnson P 2020 Int. Conf. on Medical Image Computing and Computer-Assisted InterventionEnd-to-end variational networks for accelerated MRI reconstruction (Springer) pp 64–73
- Sui B, Lv J, Tong X, Li Y and Wang C 2021 Simultaneous image reconstruction and lesion segmentation in accelerated MRI using multitasking learning Med. Phys. 48 7189–98
- Sun L, Fan Z, Ding X, Huang Y and Paisley J 2019 Information Processing in Medical Imaging: 26th Int. Conf., IPMI 2019, Hong Kong, China, June 2–7, 2019, Proc. 26Joint CS-MRI reconstruction and segmentation with a unified deep network (Springer) pp 492–504
- Tolpadi A A et al 2023 K2s challenge: from undersampled k-space to automatic segmentation Bioengineering 2 267
- Vos T *et al* 2012 Years lived with disability (ylds) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the global burden of disease study 2010 *The lancet* 380 2163–96
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* 13 600–12
- Woo B, Engstrom C, Baresic W, Fripp J, Crozier S and Chandra S S 2024 Automated anomaly-aware 3d segmentation of bones and cartilages in knee MR images from the osteoarthritis initiative *Med. Image Anal.* **93** 103089
- Zimmermann F F, Kolbitsch C, Schuenke P and Kofler A 2024 Pinqi: an end-to-end physics-informed approach to learned quantitative MRI reconstruction *IEEE Transactions on Computational Imaging* (https://doi.org/10.1109/TCI.2024.3388869)