Deep Learning for Parallel MRI Reconstruction: Overview, Challenges, and Opportunities

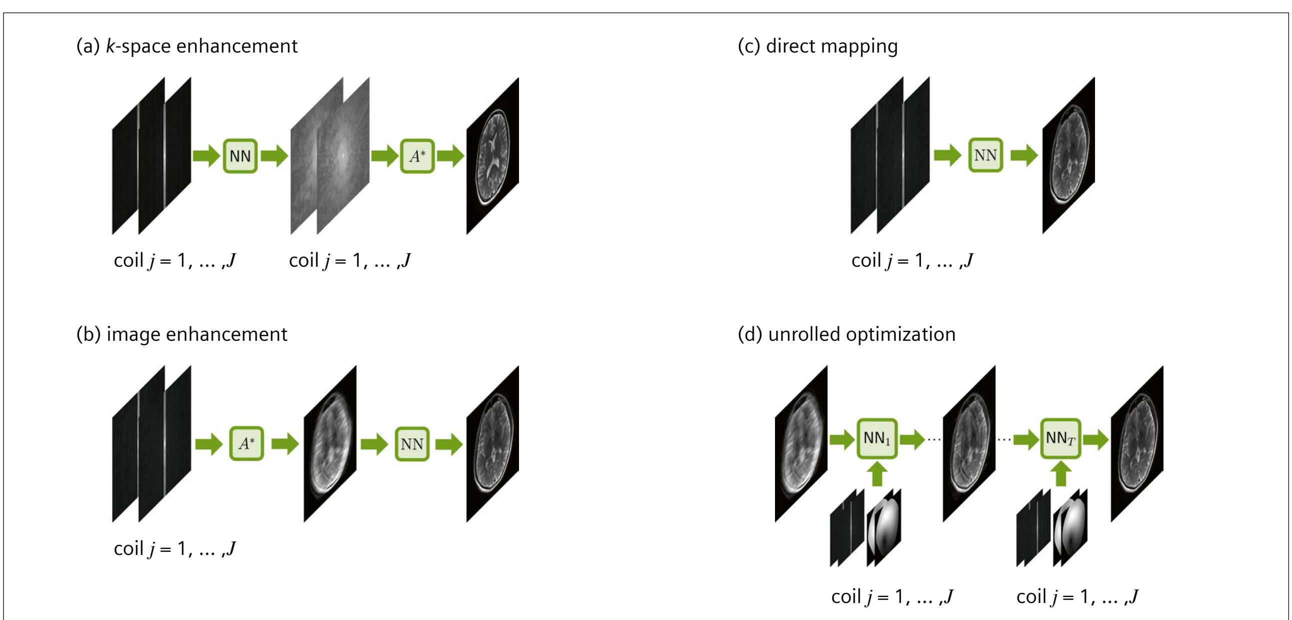
Kerstin Hammernik[1,2]; Florian Knoll[3]; Daniel Rueckert[1]

[1]Department of Computing, Imperial College London, United Kingdom
[2]Institute of Computer Graphics and Vision, Graz University of Technology, Graz, Austria
[3]Department of Radiology, New York University, New York, NY, USA

With the success of parallel imaging [1–4] and compressed sensing [5–7], we have achieved a breakthrough in the field of routine clinical MR imaging to tremendously accelerate the inherently slow acquisition process. However, with the currently available technologies, we have reached a plateau in terms of acquisition speed. The next paradigm shift is already looming: During the past years, we have seen a tremendous development and impressive results of deep learning [8] algorithms in the field of medical imaging. There are many opportunities how deep learning tools can change the world of clinical examinations, ranging from precision medicine, computer aided diagnosis, image classification and segmentation to data acquisition and image reconstruction. Deep learning leverages the potential to change the complete workflow of clinical imaging, however, many algorithms have been developed regardless of practical relevance. In this article, we focus on the application of deep learning tools for parallel MR image reconstruction. We show how the limits of acquisition speed in MR imaging can be pushed even further, with improved image quality and reduced image reconstruction times compared to current state-of-the-art methods and we provide insights into this highly demanding, clinical standard application from different perspectives.

Deep learning for image reconstruction

Image reconstruction aims at recovering a clean, high-quality MR image from a set of acquired *k*-space measurements from multiple receiver coils. This process involves inverse Fourier transforms to map the measured *k*-space data to the image space. However, this is an ill-posed problem due to measurement errors, low signal-to-noise ratios, sparsely sampled data and limitations of



1 Overview of deep learning for parallel MR image reconstruction.

the hardware itself. Two great categories exist for MR image reconstruction, which are in fact closely related: Sensitivity encoding (SENSE) [1, 2] operates in image domain, while generalized autocalibrating partial parallel acquisitions (GRAPPA) [3] fills the missing information of undersampled acquisitions in *k*-space. Similar to the question if one would prefer SENSE or GRAPPA for image reconstruction, deep learning can improve image reconstruction both in *k*-space and image space. In this article, we introduce the basic idea of how deep learning can be used for parallel MR image reconstruction, which is illustrated in Figure 1. We refer the interested reader to [9, 10] for a more detailed insight into this topic.

Learning *k*-space enhancement

To learn improved *k*-space enhancement, both supervised learning methods that depend on training data and self-supervised methods, that learn an interpolation function from the fully sampled autocalibration lines, are reported. DeepSPIRiT [11] is based on *k*-space convolutional neural networks (CNN) that are trained on a reference database and does not depend on explicit coil sensitivity maps, *k*-space interpolation kernels or collection of autocalibration lines. In contrast, GRAPPA-based methods learn a relationship between the coils from an autocalibration signal. While GRAPPA, which is the most clinically used reconstruction method, is based on linear kernel methods and known to suffer from severe noise amplification at higher acceleration rates, it was improved by non-linear kernel methods (RAKI) [12]. RAKI is based on training a CNN from the autocalibration signal to interpolate missing *k*-space lines, resulting in less severe noise amplification compared to GRAPPA as illustrated in Figure 2.

Learning image enhancement

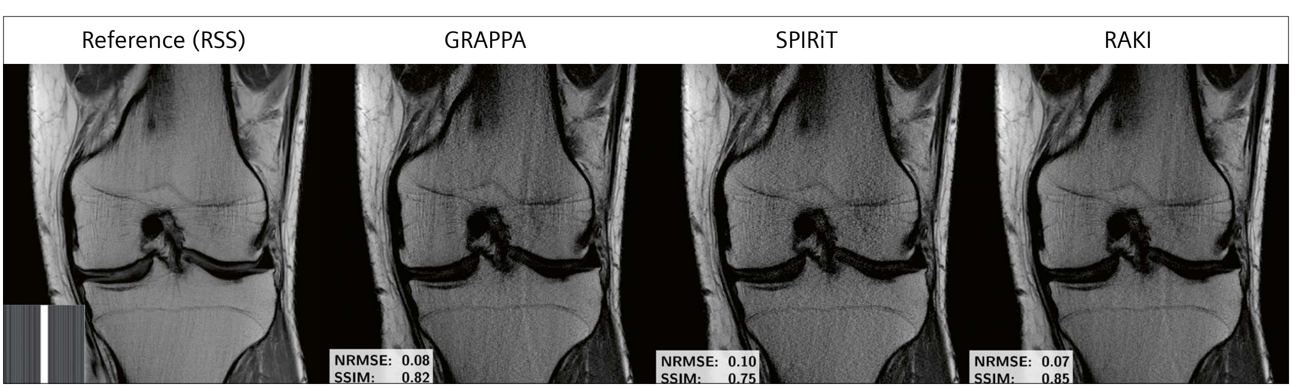
The first group of image-based algorithms learns to enhance a, possibly coil-sensitivity-weighted, zero-filling solution [13–18]. The zero-filling solution is mapped to a fully-sampled reference and does not require any further prior knowledge, hence, consistency to the measured *k*-space data is not ensured. To train these architectures successfully, large amounts of training samples and huge network architectures are required.

Learning the direct transform

The second group of image-based algorithms directly learn a transformation from the undersampled *k*-space data to the fully sampled image. This approach was presented as AUTOMAP by Zhu et al. [19] and is especially useful to overcome errors in the physical model, i.e., imperfect forward operator. The AUTOMAP architecture is characterized by a combination of fully-connected layers with convolutional layers on top and can be applied to any sampling trajectory in MR. The training requires a huge amount of memory due to the fully connected layers, hence, AUTOMAP is limited to small image size. This scaleability issue can be improved by decomposing AUTOMAP as shown in [20]. Although preliminary results are promising, it is practically not applicable yet as the trained networks are limited to specific input sizes which do not meet the requirements for heterogeneous training data in clinical practise.

Learning unrolled optimization

The third and largest group of image-based algorithms learns a fixed unrolled scheme in a supervised end-to-end manner. These fixed, unrolled schemes alternately update the image and impose data consistency. Data consistency can be realized in various ways: By performing gradient steps as in Variational Networks [21] or by solving the proximal mapping [22, 23]. We also find various other optimization schemes, not only for parallel MRI reconstruction, but for medical imaging in general. These schemes range from ADMM-net [24] to variable-splitting schemes [25] and primal-dual optimization [26].



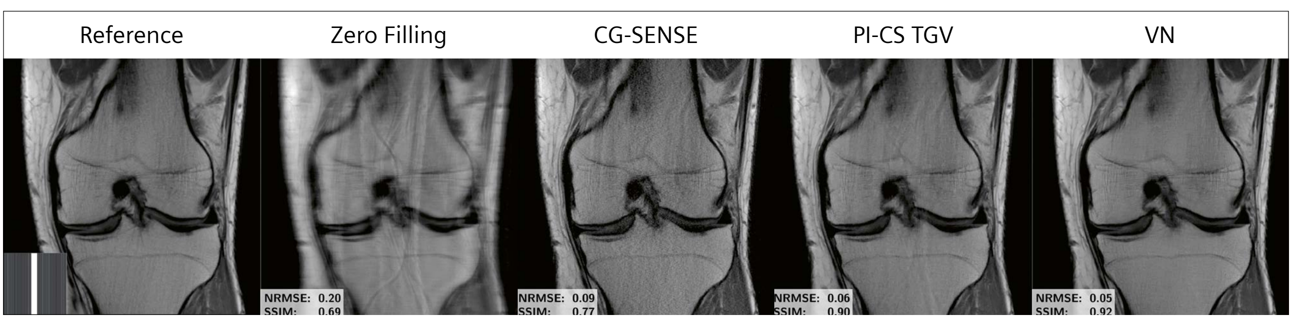
2 GRAPPA-type reconstructions for 4-fold Cartesian undersampling: Comparison of classic GRAPPA, SPIRiT [32] and learning-based RAKI [12] to the root-sum-of-squares (RSS) reference. While GRAPPA suffers from residual artifacts and SPIRiT from noise amplification, RAKI reconstructions achieve both better noise suppression and less residual artifacts. This observation is supported by the quantitative values.

An example for the impact of learning unrolled optimization for accelerated Cartesian imaging is depicted in Figure 3. Here, a variational network reconstruction [21] is compared to a linear CG-SENSE reconstruction [1] and a combined parallel-imaging-compressed sensing (PI-CS) approach [27]. We can clearly see the benefits of the learning-based reconstruction algorithm.

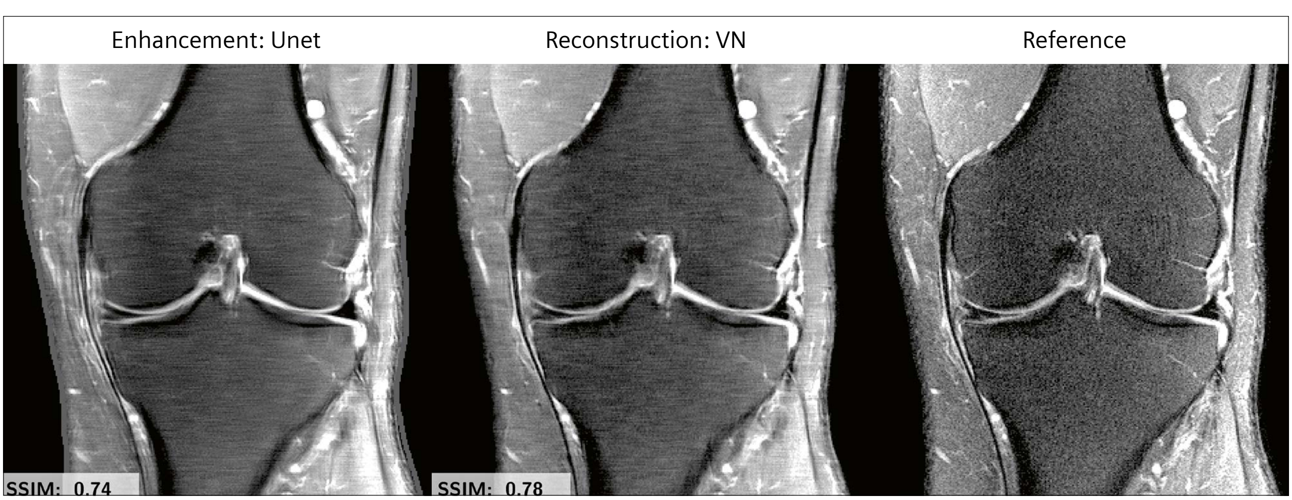
Another question which arises at this stage is if it is really necessary to incorporate the original *k*-space data into the learning-based reconstruction process or if learning image enhancement is enough. A first answer to this question is depicted in Figure 4. Here, a Unet [28] architecture is trained to enhance an initial sensitivity-combined zero-filling solution [16]. This architecture has about 14 million parameters. In comparison, a simple variational network with 140,000 parameters was trained as an unrolled gradient descent scheme with 10 iterations according to [21]. Indeed, we observe that the learned unrolled scheme outperforms the image enhancing network, hence, it is beneficial to include any available prior knowledge in the reconstruction process, which makes the learning task easier and might require less training data to achieve descent results.

High demands for learning-based MRI reconstruction approaches

Many research papers show promising results for learning-based MR image reconstruction, however, these results are often presented for a specific sequence and a simulated environment, e.g., single-coil MR data. We also feel that the image content for a specific sequence is very similar over a wide range of data. In fact, the data is highly inhomogenous and the radiologists' expectations differ from the researchers' perspective in terms of evaluation.



3 SENSE-type reconstructions for 4-fold Cartesian undersampling. Comparison of linear CG-SENSE reconstruction, a parallel imaging-compressed sensing combined Total Generalized Variation (PI-CS TGV) reconstruction [27], and a learning-based Variational Network (VN) reconstruction [21]. The learning-based VN approach reaches superior image quality and reduced artifacts.



4 Comparison of learning-based image enhancement with a Unet to learning-based image reconstruction with a variational network (VN). The VN that uses the acquired *k*-space data achieves better SSIM scores and has only a fraction (1%) of the parameters compared to the Unet.

More than inverse Fourier transforms

When addressing MR data we often have a simplified picture in mind, telling us that we just have to perform an inverse Fourier transform to obtain the reconstructed image. Indeed, inverse Fourier transforms are the main ingredient for image reconstruction, however, many more aspects have to be considered when building learing-based solutions. The data are acquired in Fourier domain, hence, are complex-valued, which has to be addressed. A common approach here is to handle the complex-valued images as a two-channel real image. Further research points towards addressing the complex-valued issue by complex convolutions and complex activations [29].

In the case of SENSE-based approaches, the explicit estimation of coil sensitivity maps is still a pre-processing step and the final reconstruction quality highly depends on the quality of the coil sensitivity maps. The joint estimation of reconstructed image and coil sensitivity maps is still an open question in deep learning.

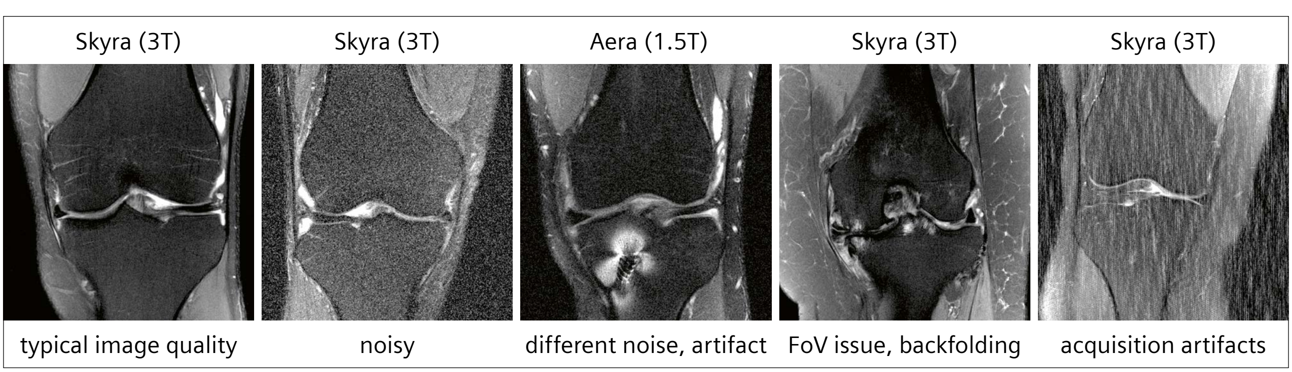
Heterogenous MRI data acquisition

Case study

A 42-year-old female patient had to undergo a clinical knee examination. The patient was referred to institution A, where a full clinical protocol (consisting of coronal PDw, coronal PDw with fat saturation, sagittal PDw, sagittal T2w with fat-saturation, axial T2w with fat saturation) was acquired using a 3T MAGNETOM Skyra (Siemens Healthcare, Erlangen, Germany) and a 15-channel knee coil. The patient moved to another country and had to undergo further treatment due to reappearing medical issues at institution B, where the same clinical protocol was acquired using a 1.5T MAGNETOM Aera (Siemens Healthcare, Erlangen, Germany) and a 15-channel knee coil. The patient, who is a computer scientist, was astonished when comparing the two protocols: Why do the images have different size? Why are some parts of the knee cropped and folded back on the other side in certain images of institution A? Why is so much noise in the images of institution B? You can hardly see anything! Why are there some artifacts in the images of institution A?

Using this case study, we can already identify common challenges in every day clinical MRI examinations: There is no universal acquisition scheme and even standard acquisition protocols vary not only from institution to institution, but also within the institution. These variations include changes in sequence parameters, matrix size, and base resolution. Furthermore, the radiographers acquiring the images have to adapt characteristic parameters individually for each patient. This includes for example the setting for phase encoding oversampling to ensure that the entire field-of-view is considered during acquisition and no backfolding occurs. Not only the sequence setting but also the hardware setting itself has a huge impact on the final image quality. The image quality depends on the coil load and the field strength of the MRI scanner. Other sources that influence the image quality are any kind of artifacts, including patient motion. The fastMRI dataset [16] is a great example for a heterogeneous dataset. The protocols for this dataset were adapted to fit the individual scanner hardware and imaged patient as optimal as possible. Figure 5 shows examples from this dataset for varying field strengths, noise levels and artifacts. The images have a different dimension in phase encoding (x) direction. The number of phase encoding steps are individually adjusted during acquisition, hence, introduce another degree of inhomogeneity for learning-based reconstruction approaches.

We see that learning-based approaches have to deal with highly heterogeneous data. Most of the currently available approaches are tested on high SNR data of the same scanner and sequence. Furthermore, for supervised learning approaches, we have to define a "ground truth" reconstruction. This becomes even more challenging, when an additional dynamic component is added or quantitative imaging is performed using learning-based approaches.



5 Examples of coronal PD-weighted knee scans with fat saturation from the fastMRI dataset [16]. The dataset contains heterogeneous data in terms of varying field strength, number of phase encoding steps and artifacts.

Researchers' evaluation

From a researchers' perspective, quantitative evaluation metrics are required to benchmark different approaches objectively. Common evaluation metrics here are the Peak-Signal-to-Noise ratio (PSNR) or the Structural Similarity Index (SSIM) [30]. Furthermore, supervised machine learning approaches require a qualitative image metric for network training which is reflected in the final image quality. The major drawback of most commonly used quantitative image metrics is their over-smoothing behaviour and the incapability to reflect small, subtle details in the metric. The images might have high quantitative scores, but appear unpleasant from a radiologists' perspective, hence, the insights into the true nature of MR images are still limited. Another obvious issue is to compare GRAPPA-based and SENSE-based algorithms, especially when no additional noise measurement is available to obtain the optimal weighting between the individual coils. Although Figures 2 and 3 show the same image slice, they cannot be compared directly as in this case the SENSE-based algorithms require explicit coil sensitivity maps and GRAPPA-based methods use implicit coil sensitivities for reconstruction. Hence, it is still an open question how to compare these algorithms from a researchers' perspective and it might require a more thorough evaluation of radiologists.

Radiologists' evaluation

In a diagnostic setting, the evaluation of diagnostic content, hence, the presence or absence of small subtle structures is inevitable. This requires large-scale studies to prove if all information are still available for the correct diagnosis if the acquisitions are accelerated. This includes a subjective evaluation of the image quality itself: Many learning-based approaches suffer from blurred images and residual artifacts, however, are these degraded images sufficient for correct diagnosis? Up to now, only small-scale studies in terms of image quality [21] and diagnostic content [31] were performed. Future opportunities include an evaluation on more diverse MR data, sample size and imaging exams.

Opportunities

A major concern in deep learning for medical imaging in general is the need for big data itself. While for computer vision applications, large databases and benchmarks exist, big data are only slowly arriving in medical imaging, mostly associated with dedicated image challenges. In the field of image reconstruction, the fastMRI dataset [16] provides a huge step towards a more generalized raw data archive, currently containing about 1000 fully sampled knee training datasets acquired with different Siemens Healthineers scanners. This rises the questions if we can train a universal network that is able to deal with the heterogeneous data, various anatomies and intra-/inter-vendor hardware settings. Are we able to train a universal network? Or might semi-supervised and unsupervised approaches provide a way to adapt to patient- and acquisition-specific clinical scenarios?

We experience that deep learning, and artificial intelligence in general, have the capability to change the complete imaging workflow in radiology. However, many of the existing approaches so far are based on simulated scenarios and have limited clinical value. Considering parallel imaging in image reconstruction provides a first step towards clinical applicability. While we focus here only on imaging data, there is also a vast amount of medical records, patients' history and even other sensor data available, which might be included and improve the image acquisition and reconstruction workflow.

Conclusion

The current developments in deep learning for medical image reconstruction reminds one of the hype we have experienced with compressed sensing. While compressed sensing started to change image reconstruction almost 20 years ago, it is only now a commercial product and established in clinical routine. Similarly, deep learning approaches for image reconstruction are not yet established in clinical examinations and will require a thorough evaluation, but they already provide a huge potential for the future of MR image reconstruction.

Acknowledgements

We want to thank Jo Schlemper for providing the Unet reconstruction and Mehmet Akçakaya for providing the GRAPPA, SPIRiT, and RAKI reconstructions. We also acknowledge research support by the EPSRC Programme Grant EP/P001009/1 and the National Institutes of Health under grants R01EB024532 and P41EB017183.

References

1 Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. Magn Reson Med, 42(5):952–962, 1999.
2 K P Pruessmann, M Weiger, P Boernert, and P Boesiger. Advances in sensitivity encoding with arbitrary k-space trajectories. Magn Reson Med, 46(4):638–651, 2001.
3 Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn ResonMed, 47(6):1202–1210, 2002.
4 D K Sodickson and W J Manning. Simultaneous Acquisition of Spatial Harmonics (SMASH): Fast Imaging with Radiofrequency Coil Arrays. Magn ResonMed, 38(4):591–603, 1997.
5 Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med, 58(6):1182–1195, 2007.

6 Kai Tobias Block, Martin Uecker, and Jens Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. Magn Reson Med, 57(6):1086–1098, 2007.
7 Michael Lustig, David L. Donoho, Juan M. Santos, and John M. Pauly. Compressed Sensing MRI. IEEE Signal Processing Magazine, 25(2):72–82, 2008.
8 Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep Learning. Nature,521(7553):436–444, 2015.
9 Kerstin Hammernik and Florian Knoll. Machine learning for image reconstruction. In S. Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger, editors, Handbook of Medical Image Computing and Computer Assisted Intervention. Academic Press, 2019.
10 Florian Knoll, Kerstin Hammernik, Chi Zhang, Steen Moeller, Thomas Pock, Daniel K. Sodickson, and Mehmet Akcakaya. Deep Learning Methods for Parallel Magnetic Resonance Image Reconstruction. Accepted for publication in IEEE Signal Processing Magazine, 2019.
11 Joseph Y. Cheng, Morteza Mardani, Marcus T. Alley, John M. Pauly, and Shreyas S. Vasanawala. Deep-SPIRiT: Generalized parallel imaging using deep convolutional neural networks. In Proc. 26th Annual Meeting of the ISMRM, Paris, France, 2018.
12 Mehmet Akçakaya, Steen Moeller, Sebastian Weingärtner, and Kâmil Uğurbil. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. Magn Reson Med, 81(1):439–453, 2019.
13 Yoseob Han, Jaejun Yoo, Hak Hee Kim, Hee Jung Shin, Kyunghyun Sung, Jong Chul Ye. Deep learning with domain adaptation for accelerated projection-reconstruction MR. Magn Reson Med, 80:1189–1205, 2018.
14 Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin KeunSeo. Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology, 63(13), 2018.
15 Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xu-jiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, and David Firmin. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Transactions on Medical Imaging, 2018.
16 Florian Knoll, Jure Zbontar, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, ErichOwens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fast MRI: a publicly available raw *k*-space and DICOM dataset for accelerated MR image reconstruction using machine learning. Radiology AI, page in press, 2019.
17 Jong Chul Ye, Yoseob Han, and Eunju Cha. Deep Convolutional Framelets: A General Deep Learning Framework for Inverse Problems. SIAM Journal in Imaging Sciences, 11(2):991–1048, 2018.
18 Kyong Hwan Jin, Michael T. McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. IEEE Trans Image Proc, 26(9):4509–4522, 2017.
19 Bo Zhu, Jeremiah Z. Liu, Stephen F. Cauley, Bruce R. Rosen, and Matthew S. Rosen. Image reconstruction by domain-transform manifold learning. Nature 555 (7697), pp. 487-492, 2018.
20 Jo Schlemper, Ilkay Oksuz, James Clough, Jinming Duan, Andrew King, Julia Schnabel, Joseph Hajnal, and Daniel Rueckert. dautomap: Decomposing automap to achieve scalability and enhance performance. In ISMRM 27th Annual Meeting, page 0658, 2019.
21 Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. Mag Reson Med, 79:3055–71, 2018.
22 Jo Schlemper, Jose Caballero, Joseph Hajnal, Anthony Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for MR image reconstruction. In Proceedings of the International Society of Magn Reson Med (ISMRM), page 643, 2017.
23 Hemant K. Aggarwal, Merry P. Mani, and Mathews Jacob. MoDL: Model-Based Deep Learning Architecture for Inverse Problems. IEEE Transactions on Medical Imaging, 2019.
24 Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Deep ADMM-Net for Compressive Sensing MRI. In Advances in Neural Information Processing Systems (NIPS), pages 10–18, 2016.
25 Jinming Duan, Jo Schlemper, Chen Qin, Cheng Ouyang, Wenjia Bai, Carlo Biffi, Ghalib Bello, Ben Statton, Declan P O'Regan, and Daniel Rueckert. Vs-net: Variable splitting network for accelerated parallel mri reconstruction. Medical Image Computing and Computer Assisted Intervention (MICCAI), pp. 713-722, 2019.
26 Jonas Adler and Ozan Öktem. Learned Primal-Dual Reconstruction. IEEE Transactions on Medical Imaging, 37(6):1322–1332, 2018.
27 Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. Second order total generalized variation (TGV) for MRI. Magnetic Resonance in Medicine, 65(2):480–491, 2011.
28 Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp 234-241, 2015.
29 Chiheb Trabelsi, Olexa Bilaniuk, Dmitriy Serdyuk, Sandeep Subramanian, João Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J. Pal. Deep complex networks. arXiv preprint arXiv:1705.09792,2017.
30 Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Proc, 13(4):600–612, 2004.
31 Florian Knoll, Kerstin Hammernik, Elisabeth Garwood, Anna Hirschmann, Leon Rybak, Mary Bruno, Tobias Block, James Babb, Thomas Pock, Daniel Sodickson, and Michael Recht. Accelerated Knee Imaging Using a Deep Learning Based Reconstruction. In Proceedings of the International Society of Magn Reson Med (ISMRM), page 645, 2017.
32 Michael Lustig and John M Pauly. SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary *k*-space. Magn Reson Med, 64(2):457–471, 2010.

Contact

Kerstin Hammernik
Imperial College London
Department of Computing
South Kensington Campus
London SW7 2AZ
United Kingdom
k.hammernik@imperial.ac.uk