ABSTRACT

Purpose: To propose COMPaS, a learning-free Convolutional Network, that combines Deep Image Prior (DIP) with transform-domain sparsity constraints to reconstruct undersampled Magnetic Resonance Imaging (MRI) data without previous training of the network. Methods: COMPaS uses a U-Net as DIP for undersampled MR data in the image domain. Reconstruction is constrained by data fidelity to k-space measurements and transform-domain sparsity, such as Total Variation (TV) or Wavelet transform sparsity. Two-dimensional MRI data from the public FastMRI dataset with Cartesian undersampling in phase-encoding direction were reconstructed for different acceleration rates \(R\) from \(R = 2\) to \(R = 8\) for single coil and multicoil data. Performance of the proposed architecture was compared to Parallel Imaging with Compressed Sensing (PICS). Results: COMPaS outperforms standard PICS algorithms by reducing ghosting artifacts and yielding higher quantitative reconstruction quality metrics in multicoil imaging settings and especially in single coil k-space reconstruction. Furthermore, COMPaS can reconstruct multicoil data without explicit knowledge of coil sensitivity profiles. Conclusions: COMPaS utilizes a training-free convolutional network as a DIP in MRI reconstruction and transforms it with transform-domain sparsity regularization. It is a competitive algorithm for parallel imaging and a novel tool for accelerating single coil MRI.

Keywords: Convolutional Prior, Magnetic Resonance Imaging, Deep Image Prior, DIP, Sparsity, Domain Transform Sparsity, Compressed Sensing, CNN, learning-free, Neural Network

1 Introduction

Increasing imaging speed is one major goal of ongoing research in Magnetic Resonance Imaging (MRI). Two common approaches are widely used in clinical routine to accelerate data acquisition. Both are based on extensive undersampling of the MR-data in the Fourier domain. The first one is parallel imaging (PI) which exploits the redundancy of MR-data received by several radio-frequency coils \([1, 2, 3, 4]\). The second one introduces sparsity regularization into the image reconstruction process and is known as compressed sensing \([5]\). In combination both techniques can achieve high acceleration rates \([6, 7]\). Nonetheless, the reconstruction problem in general is an ill-posed optimization problem. Solving it often requires computationally intense iterative algorithms, which limits its application for example in fields like real time imaging. Thus the need for fast image reconstruction has been one reason for the advent of neural networks and deep learning strategies in the application of MRI reconstruction \([8, 9, 10]\).

Abbreviations: BART, Berkeley Advanced Reconstruction Toolbox; COMPaS, Convolutional MRI Prior with Sparsity regulation CS, Compressed Sensing; DIP, Deep Image Prior; GT, Ground Truth; L1W, Wavelet L1 Norm; MRI, Magnetic Resonance Imaging; MSE, Mean Squared Error; PI, Parallel Imaging; PICS, Parallel Imaging combined with Compressed Sensing; RAKI, Scan-specific robust artificial-neural-networks for k-space interpolation; SSIM, Structural Similarity Index; TV, Total Variation
On a high abstraction level, one can distinguish two classes of methods, that use neural networks. Data-driven machine learning algorithms require large datasets in order to tune network parameters, so that they ‘learn’ to reconstruct high-quality images from undersampled measurements. Several studies have shown that, deep learning approaches outperform aforementioned state-of-the-art methods in terms of reconstruction speed but also in terms of image quality [11][12]. Despite remarkable results, data-based approaches have serious limitations, when generalizing from training data: Overfitting to particular features of the training data set can lead to bias and limit generalization when posed with previously unseen data in new applications. Upon inference on new data, this can lead to hallucination of features that are common in the training data set. Vice versa, it may cause removal of actual features, that were never seen before in the training data, which can be fatal for diagnostic imaging [12]. Moreover, reconstruction quality of a well-trained neural network deteriorates, when confronted with new imaging properties, like different sampling patterns or SNR of the underlying data [13].

The second approach that uses neural network algorithms does not require training data and fits the model parameters only to the individual scan’s data [14]. For example, the RAKI [15] architecture extends GRAPPA’s [1] linear convolution kernels by multiple layers of scan-specifically trained convolutions, linked by rectified linear units (ReLU). The authors point out, that non-linear functions with few degrees of freedom are better at approximating higher-dimensional linear functions and show that RAKI has superior noise-resilience and better reconstruction performance than GRAPPA. In both algorithms, the convolution kernel parameters are optimized to represent a mapping function from undersampled zero-filled k-space data \( k_0 \) to reconstructed \( k \), where the missing lines are calculated from local correlations to neighboring multi-channel k-space points. A similar approach is pursued by Wang et al. [16], who use an unrolled Convolutional Neural Network (CNN), with data-consistency layers to map from \( k_0 \) to \( k \).

In the present paper we focus on a different reconstruction strategy, coined Deep Image Prior (DIP), which originates from the field of image restoration. Similar to MR-reconstruction tasks, image restoration tasks can be formulated as an optimization problem. Instead of using a gradient descent method to search for an optimal solution \( x \) directly in the image space, DIP transmits the search to the parameter space of the neural network (originally a U-Net) by optimizing the network weights until the network is able to reconstruct the final image output of a fixed random noise input sample. Ulyanov et al. showed, that a U-Net [17] architecture is biased to prefer the reconstruction of natural images over noise data. Therefore, training the network to reconstruct a corrupted image \( x_0 \), starting from randomly initialized parameters, the solution \( x \) is reached with lower impedance than \( x_0 \) during training. The architecture serves as a statistical prior of \( x \). DIP has shown competitive performance with data driven methods for image restoration tasks such as inpainting, denoising and super-resolution. The key idea of our work is to use a DIP-like approach for ‘inpainting’ of missing k-space data, i.e., to determine the Fourier coefficients \( k \) from the undersampled measurement \( k_0 \).

A similar approach can be found in Di Zhao et al. [18]. The authors use previously measured reference data to pretrain the DIP model’s network parameters before beginning the actual reconstruction task. This implies however, that the actual reconstruction task cannot be considered as a de novo image synthesis due to the use of a reference input. This approach therefore is not strictly training-free - although reduced - and does not eliminate the above-mentioned risks of bias towards training data.

A reference-free DIP approach was investigated in time-dependent radial MR-data [19]. By using latent mapping combined with generative CNNs (convolutional neural networks), Yoo et al. achieved promising results with time-dependent MR datasets. Time series of MR data in general show much higher informational redundancy as single images and thus are very suitable for high acceleration rates [20][21][22].

In our work we show that by augmenting the loss function and introducing a sparsity prior, DIP can also be applied in single image reconstruction and with random Cartesian sampling. We demonstrate that our approach outperforms state-of-the art PICS reconstructions and can also be applied in single channel reconstruction.

2 Methods

2.1 Reconstruction Pipeline

DIP [23] has shown promising results in various image reconstruction tasks, most notably for our context, image inpainting. The basic concept of our approach is to transfer this capability for image inpainting to k-space. However, convolutional networks, that are used as DIPs, favor natural images by imposing local spatial correlations on different length scales. The k-space results from an integral transform and does not show these local structures. Hence, the reconstruction \( k \in \mathbb{C}^{n_c \times N_x \times N_y} \), where \( n_c \) is the number of coil channels and \( N_x \times N_y \) the number of k-space coefficients in readout and phase encoding direction, is not searched directly in k-space by our approach. Instead, the U-Net [17] \( f \) operates in image space. It searches the parameter space of network weights \( \theta \), from which an image \( x = f_\theta(z) \) is built from a fixed but randomly initialized input \( z \in \mathbb{R}^{2 \times N_x \times N_y} \). The network architecture is shown in Fig. 1(a). Details on the layer parameters are provided in Tab. 1.

\( k \) is then given as the coil sensitivity-weighted Fourier transform \( k = FSx \), where \( F \) is the Fourier transform and \( S \) represents the coil sensitivity profiles obtained via ESPIRiT [24] calibration, which is implemented in BART [25]. To iteratively find the optimal k-space reconstruction \( k = FSx = FSf_\theta(z) \), the parameters \( \theta \) of the network \( f \) are optimized, such that

\[
\hat{\theta} = \arg \min_{\theta} \mathcal{L}(f_\theta(z), k_0)
\]  

(1)
Table 1 Parameters of the U-Net layers.

| Layer | input size | filters |
|-------|------------|---------|
| Down 1 | 2, 396² | 64 |
| Down 2 | 64, 198² | 128 |
| Down 3 | 128, 99² | 256 |
| Down 4 | 256, 49² | 512 |
| Bottom | 512, 24² | 1024 |
| Up 1 | 1024, 49² | 256 |
| Up 2 | 512, 99² | 128 |
| Up 3 | 256, 198² | 64 |
| Up 4 | 128, 396² | 2 |

Input sizes and number of filters for the convolution blocks in the U-Net. Downsampling layers i form the left side of the U-Net in Fig. 1(a), the upsampling layers form the right side. Each Layer Down i and Up i consists of two 2D Convolutions connected by ReLU and Batch Norm layer. From Down i to Down i + 1 down-sampling by 2D Maximum Pooling is used. For upsampling, the identical filter sizes are used in reverse order. On each level, the downsampling filter maps are concatenated to the upsampling filter maps Up i, before forwarding to Up i + 1. The final layer block (Up 4) returns an image with two channels, representing real and imaginary part.

Here, \( \mathcal{L} \) is a cost function, which is further elaborated in Sec. 2.2. For optimization, the ADAM optimizer \( \theta \) with a learning rate \( l \) is used. The reconstruction algorithm is summarized in Algorithm 1. Fourier transforming the network output before passing it to \( \mathcal{L} \) allows to implement a loss functional that compares \( k \) against the measured \( k_0 \), whilst using the network’s convolutions in the image domain as an appropriate prior to the image statistics and spatial correlations.

After stopping the iteration, the final multi-coil reconstruction is given by \( \hat{k} \) after ensuring data-consistency in sampled k-space points \( k_0 \), which is formalized by

\[
\text{DC}(\hat{k}, k_0) = \begin{cases} 
  k_0 & \text{where } P = 1 \\
  \hat{k} & \text{else.}
\end{cases}
\]

Here, \( P \) is a matrix, representing the undersampling projection operator. Finally, the image space reconstruction is obtained by phase-sensitive combination of the Fourier transform of the multi-coil k-space data, as described by Roemer et al. [27].

Calibration-free reconstruction The number of output filter maps of the network can be adjusted to output \( X \in \mathbb{C}^{n_x \times n_y \times n_u} \) instead of \( x \in \mathbb{C}^{n_y \times n_u} \). This allows to build an algorithm that does not require previously calibrated coil sensitivity profiles, but inherently outputs a multi-coil k-space approximation \( k = FX \). Then, \( x \in \mathbb{C}^{n_y \times n_u} \) is given as the root sum of squares of the inverse Fourier transform.

Algorithm 1 k-Space reconstruction with COMPaS.

Require: Undersampled k-space \( k_0 \in \mathbb{C}^{n_x \times n_y \times n_u} \), number of epochs \( n \), learning rate \( l \), estimate coil sensitivity profiles \( S \) from \( k_0 \) with ESPIRiT

Randomly initialize \( z \in \mathbb{R}^{2 \times n_x \times n_y} \) and \( \theta = 0 \)

for \( n \) iterations do

\[
x_i = f_{\theta_i}(z) \quad \text{with} \quad x \in \mathbb{C}^{n_y \times n_u}
\]

\[
\theta_{i+1} = \theta_i - \text{ADAM}_l(\mathcal{L}(x_i, k_0))
\]

end for

\[
x = f_{\theta_n}(z)
\]

\[
k = DC(FSx, k_0)
\]

Output: \( k \in \mathbb{C}^{n_x \times n_y \times n_u} \)

2.2 Loss function

During the training process as given by Eq. 1, the trainable network parameters \( \theta \) are tuned by gradient backpropagation from the multi-term loss-function \( \mathcal{L} \). It is depicted in Fig. 1(b) and given as

\[
\mathcal{L}(x, k_0) = \eta_1 \|PFx - k_0\|_1 \\
+ \eta_2 \|F^{-1}PFx - F^{-1}k_0\|_2 \\
+ \rho \mathcal{R}[F^{-1}DC(Fx, k_0)]
\]

The first two terms of \( \mathcal{L} \) assert data consistency of the reconstruction with \( k_0 \). The last term enforces transform-domain sparsity of \( x \).

Sparsity Regularization For the sparsity regularization term \( \mathcal{R}[x] \), Total Variation (TV), as well as L1 norm of the wavelet coefficients (L1W) have proven useful [28]. For TV, we used a finite gradient operator with a one-pixel-shift in all spatial dimensions. For L1W, we used a wavelet transformation \( \Psi \) with Daubechies-Wavelets (D4) as basis functions and decomposition level of 5.

2.3 Reconstruction data

Reconstruction performance of the suggested design was evaluated on an axial T2-weighted image from the FastMRI brain dataset [29] (file name file_brain_AXT2_201_2010066). The lowest slice from the fully sampled multi-slice dataset was taken as the ground truth (GT) and retrospectively undersampled. To save memory the field of view was cropped to 396² pixels. For Cartesian undersampling we pseudo-randomly selected readout lines from a Gaussian probability density distribution with maximum in k-space center. The sampling algorithm was designed to assure, that the central 18 lines were always sampled. The used sampling patterns are sketched in Fig. 2. GT and zero-filling reconstructions for all reported acceleration rates \( R \) are shown in Fig. 3. Multicoil (\( n_u = 12 \)) reconstruction performance was assessed for \( R = 3, 5, 8 \). The robustness of our results was investigated on a subset of 200 measurements, selected randomly from the FastMRI Brain dataset. Different contrasts (FLAIR, T1, T2) were present in the subset. From each multi-slice dataset, the lowest slice was selected.
and cropped to shape $N^2$, where $N = \min(N_x, N_y)$ and $N_{x,y}$ is the number of pixels in the two spatial dimensions. For each sample, an undersampling pattern was randomly generated as described in Sec. 2.2.3 with acceleration rates $R = 3, 5, 8$. Reconstruction was performed as described for the single sample described above and quantitative metrics of PICS and COMPaS reconstructions were compared.

Furthermore, single coil ($n_c = 1$) reconstruction for $R = 2$ was investigated. Single coil data were generated synthetically by using virtual coil compression (as provided by BART [25]) of the multicoil ($n_c = 12$) data described above.

### 2.4 Performance metrics

For quantitative evaluation of our results we used

$$L1[a, b] = \frac{1}{N} \sum_{i} |a_i - b_i|,$$  \hspace{1cm} (4)

and Mean Squared Error (MSE)

$$MSE[a, b] = \frac{1}{N} \sum_{i} (a_i - b_i)^2,$$  \hspace{1cm} (5)

where $i$ runs over all $N$ values of the input images $a$ and $b$. Results were also compared to the GT with the Structural Similarity Index

$$SSIM[a, b] = \frac{(2\mu_a \mu_b + c_1)(2\sigma_{ab} + c_2)}{\mu_a^2 + \mu_b^2 + c_1(\sigma_a^2 + \sigma_b^2 + c_2)}.$$  \hspace{1cm} (6)

where $\mu_a$, $\mu_b$, $\sigma_a$, $\sigma_b$, $\sigma_{ab}$ are the averages, $\sigma_a$, $\sigma_b$, $\sigma_{ab}$ the variances and $\sigma_{ab}$ is the covariance between $a$ and $b$.

The parameters $c_1 = (k_1 L)^2$ for $i = 1, 2$ are determined by the dynamic range $L$ of the images’ pixel values and fixed parameters $k_1 = 0.01$ and $k_2 = 0.03$.

### 2.5 Implementation

Python 3 was used for all algorithms and simulations described in this work. COMPaS network and the training pipeline were built using the PyTorch library [30]. The U-Net implementation builds on publicly available code [31]. All complex-valued tensors were replaced by tensors in $\mathbb{R}^2$. Iteration of COMPaS and PICS reconstructions were performed on Linux system with Intel(R) Xeon(R) Silver 4214R CPU and Nvidia Titan RTX GPU.

### 3 Results

k-Space reconstructions were generated by 1000 epochs of each iteration step. The hyperparameters $l$, $\eta_{1,2}$ and $\rho$ introduced in Sec. 2.2.3 as well as the number of epochs were tuned heuristically. The best results, described in the following sections, were obtained with the configurations reported in Tab. 2. Reconstruction quality was assessed both visually and quantitatively, using image error measures introduced in Sec. 2.4, which were calculated from magnitude images of the reconstructions against the GT. All results were compared to PICS reconstructions of $k_0$.

In PICS, reconstruction is formalized by the optimization problem

$$\arg \min_{x} \frac{1}{2} ||PFSx - k_0||^2_2 + \lambda S[x],$$  \hspace{1cm} (7)

where $S$ is a sparsity regularization term, in our case TV or L1W. For all reported comparisons, the PICS algorithm implemented in SigPy [32] was iterated 500 times with heuristically tuned weighting parameter $\lambda$ of the sparsity regularization term.

### 3.1 Multicoil Parallel Imaging

#### 3.1.1 Total Variation Sparsity

COMPaS and PICS parameters for reconstruction with TV sparsity functional are reported in Tab. 2. As can be seen in Fig. 4 our architecture achieves higher performance on quantitative error measures L1 and MSE in all acceleration regimes, while SSIM is similar for PICS and COMPaS. Especially for $R = 8$, COMPaS achieves approx. 24% lower L1 and approx. 47% lower MSE. Lowest relative difference in L1 and MSE (approx. 10%) is achieved for $R = 5$. However, COMPaS reduces ghosting artifacts, which are still visible in PICS reconstructions in Fig. 4 (red arrows).

#### 3.1.2 Wavelet L1 Norm Sparsity

Hyperparameter settings for COMPaS (L1W) reconstruction are given in Tab. 2. Results are compared with PICS (L1W) in Fig. 5 for different $R$. Considering PICS reconstruction, L1W removes ghosting artifacts better than TV regularization (cf. Fig. 5). However, as for TV, COMPaS outperforms PICS, judged by visual appearance and quantitative metrics. The largest relative difference can be observed for $R = 8$, where L1 shows a relative difference of approx. 12% and MSE of approx. 22%.

#### 3.1.3 Statistical Analysis

Statistical analysis was performed on a 200 samples subset of the FastMRI dataset, as described in Sec. 2.3. Reconstruction was performed for PICS and COMPaS with the hyperparameters reported in Tab. 2. For simplicity and to illustrate the robustness of COMPaS, the settings for $R = 3$ were used for all $R = 3, 5, 8$ in this sample analysis. As can be seen in Fig. 6 lower image errors and higher SSIM can be achieved with the COMPaS reconstruction. Statistical significance of group differences between PICS and COMPaS at different $R$ was evaluated with a Mann-Whitney U rank test. Differences in samples are statistically significant for almost all settings (no significant difference could be shown for MSE in $R = 3$).

In general, error metrics are higher, than in the preceding sections, for both PICS and COMPaS. This implies, that the previously presented sample ranks among those with the lowest reconstruction error. It should be noted though, that the FastMRI dataset consists of routine clinical scans, including noisy and flawed measurements or low resolution localization images. The sample file_brain_AXT2_201_2010066, which was presented above, was selected by visual appearance, which already
ascertains a certain level of image quality. This was not done for the full data set, which can explain the higher average error. Furthermore, especially for noisy GT images, the practical significance of high error metrics between reconstruction and GT can be questioned. A smooth reconstruction might not approximate well to the noisy GT but might actually be preferable in terms of diagnostic quality.

### 3.1.4 Calibration-free Multicoil Imaging

As noted in Sec. 2.2, the COMPaS method can be adapted to allow calibration-free reconstruction, i.e., without requiring ESPiRiT-estimated coil sensitivity profiles. In this case, magnitude images were obtained from the multicoil images by calculating the root sum of squares. Hyperparameters for calibration-free multicoil reconstruction with TV and L1W sparsity regularization are reported in Tab. 2.

To filter random noise in the low signal-region outside the subject’s anatomy, a signal threshold was applied to the images. Reconstructions are shown in Fig. 7. Comparing the calibration-free approach to the algorithms demonstrated in the previous section, similar L1 can be observed. Relative differences range between approx. −1% in L1W for \( R = 5 \) and +17% in TV for \( R = 8 \). Also SSIM is similar, whereas MSE is much larger than for the algorithms that use ESPiRiT estimates, with differences ranging between approx. 25% and 176%. Undersampling artifacts in the TV reconstructions are more prominent, which is shown exemplarily at \( R = 5 \) (red arrow in Fig. 7(a)). There, ghosting is still present, whereas it is removed in the COMPaS reconstruction, that explicitly uses coil sensitivity information (cf. Sec. 3.1.1 and 3.1.2). As Fig. 7(b) shows (green arrow), ghosting can be removed for L1W. Yet, judging from the visual appearance of the error maps, blurring of fine details seems to be stronger in the calibration-free algorithm.

### 3.2 Single Coil Imaging

Hyperparameters for single coil reconstruction are reported in Tab. 2 and results of PICS and COMPaS are compared in Fig. 8. For both algorithms, random noise from low signal-regions was removed, by zeroing values below a threshold value. PICS (TV) reconstruction does not remove ghosting artifacts for low \( \lambda \). For high values (\( \lambda \) was set to 5), artifacts are removed, but fine structure details are smoothed out, as can be seen from the residuals in the error map. Our architecture on the other hand eliminates ghosting artifacts while preserving finer structures, resulting in approx. 35% lower L1 and 55% lower MSE. For L1W, the COMPaS reconstruction exhibits approx. 13% smaller L1 and 30% smaller MSE. Comparing COMPaS (L1W) to COMPaS (TV), the former exhibits slightly higher noise, both visually and by quantitative metrics. However, high detail resolution appears better in L1W, as illustrated in the error maps in Fig. 8.
4 Discussion and Conclusion

The application of DIP for reconstruction of undersampled k-space data was demonstrated for 2D brain images from the FastMRI dataset. Our approach requires no training data. It exploits a CNN’s inherent capabilities for image reconstruction and generation without introducing the downsides of data-driven machine learning applications, such as limitations in generalizability and transformability. Learning-free neural networks have also been investigated by other groups for mapping of $k_0$ to k. In contrast, our approach uses the capabilities of a U-Net for image generation, as illustrated by Ulyanov et al., and transfers it to a k-space reconstruction setting.

Reconstruction performance of the proposed architecture was examined for undersampled multicoil k-space data. Our DIP architecture outperforms PICS in terms of different quantitative image metrics and visual artifact reduction. A statistical analysis of 200 images from the FastMRI dataset showed the consistency of these results in almost all investigated modalities. In general, larger image errors were observed on this subset. We used reconstruction hyperparameters, that were heuristically found during testing on a single image. Although COMPaS already outperforms PICS for this direct approach, this could hint to the additional benefit that could come with optimizing hyperparameters to specific scan properties or with further exploration of more general parameter settings on larger samples with variable imaging parameters.

Additionally, we assessed the reconstruction of under-sampled single coil measurements. In this case, the first term of the PICS optimization problem in Eq. \(7\) is under-determined. Therefore, the matrix inversion in the PICS algorithm fails, leading to strong residual artifacts. Suppressing the latter with a high $\lambda$-parameter leads to mainly sparsity-regulated reconstructions that lack detailed structures. In contrast, COMPaS allows reconstructions with better preservation of fine structures. This could be due to the higher number of degrees of freedom in the U-Net, as well as the fact that COMPaS uses architectural priors for the image reconstruction. Therefore, our approach is particularly interesting for accelerating single coil imaging, when multichannel sampling is not feasible, for example in preclinical laboratories.

In their work on DIP, Ulyanov et al. illustrated that CNNs are appropriate priors for natural images [23], that allow to recover non-sampled image information due to the statistical redundancy that lies in spatial structure and correlations of images. Our approach utilizes these advantages by reconstructing $k_0$ in the image domain while imposing both data fidelity in k-space and domain-transform sparsity, as known from CS. Prevailing acceleration strategies use information redundancy in 3D, temporal or multicoil data. Usage of DIP extends these strategies with convolutional networks as priors for natural images. Thereby, it allows higher reconstruction quality than PICS alone. Since the neural network architecture is crucial for this task, future research on new approaches are of considerable interest.

As for now, a limitation of our work is the increase in reconstruction time compared to PICS. On our system, 12-coil COMPaS (TV) takes approx. 98ms per iteration vs. 4ms for PICS (TV). PICS algorithm though, is well-established and optimized, whereas our DIP implementation is still at an experimental state with room for development. For single coil reconstruction, we observed a computation time of approx. 87ms. The reduction from 98ms does not scale with the reduction in computational effort from $n_c = 12$ to $n_c = 1$. This hints at potential for code-optimization, which is beyond the scope of this work. Moreover, future development of advanced hardware is to be expected and will allow faster and more extensive computations in image reconstruction.

Finally, it should be pointed out, that artifact correction by sparsity regularization requires incoherent ‘noise-like’ artifacts [28][33]. Therefore, as for PICS, better performance can be expected for less coherent k-space trajectories, for example with Poisson, radial or spiral sampling. Also 3D or spatio-temporal data, that bring further redundancy in the acquired data should allow for higher acceleration rates. Thus, applying our approach to these modalities can be a promising field of future research.

Acknowledgments

The authors want to thank Peter Dawood for many discussions on related topics. This work was partially supported by Deutsche Forschungsgemeinschaft (DFG) grant 396923792.

Financial disclosure

None reported.

Conflict of interest

The authors declare no potential conflict of interests.

References

[1] Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med.* 2002;47(6):1202–1210.

[2] Sodickson DK, Manning WJ. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn Reson Med.* 1997;38(4):591–603.

[3] Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med.* 1999;42(5):952–962.

[4] Deshmane A, Gulani V, Griswold MA, Seiberlich N. Parallel MR imaging. *J Magn Reson Imaging.* 2012;36(1):55–72.

[5] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med.* 2007;58(6):1182–1195.

[6] Otazo R, Feng L, Chandarana H, Block T, Axel L, Sodickson DK. Combination of compressed sensing
and parallel imaging for highly-accelerated dynamic MRI. In: *IEEE International Symposium on Biomedical Imaging*. 2012;980–983.

[7] Liang D, Liu B, Wang J, Ying L. Accelerating SENSE using compressed sensing. *Magn Reson Med.* 2009;62(6):1574–1584.

[8] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys.* 2019;29(2):102–127.

[9] Liang D, Cheng J, Ke Z, Ying L. Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks. *IEEE Signal Process Mag.* 2020;37(1):141–151.

[10] Ahishakiye E, van Gijzen MB, Tumwiine J, Wario R, Obungoloch J. A survey on deep learning in medical image reconstruction. *Intell Med.* 2021;1(3):118–127.

[11] Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans Med Imaging.* 2018;37(2):491–503.

[12] Knoll F, Hammernik K, Zhang C, et al. Deep-Learning Methods for Parallel Magnetic Resonance Imaging Reconstruction: A Survey of the Current Approaches, Trends, and Issues. *IEEE Signal Process Mag.* 2020;37(1):128–140.

[13] Knoll F, Hammernik K, Kobler E, Pock T, Recht MP, Sodickson DK. Assessment of the generalization of learned image reconstruction and the potential for transfer learning. *Magn Reson Med.* 2019;81(1):116–128.

[14] Yazdanpanah AP, Afacan O, Warfield SK. Non-learning based deep parallel MRI reconstruction (NLdPMRI). *Proc. SPIE 10949, Medical Imaging*. 2019;(10949):1094904.

[15] Açıkakaya M, Moeller S, Weingärtner S, Uğurbil K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. *Magn Reson Med.* 2019;81(1):439–453.

[16] Wang AQ, Dalca AV, Sabuncu MR. Neural Network-based Reconstruction in Compressed Sensing MRI Without Fully-sampled Training Data. 2020; [http://arxiv.org/abs/2007.14979v1](http://arxiv.org/abs/2007.14979v1)

[17] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention*. 2015;234–241.

[18] Di Zhao , Huang Y, Zhao F, Qin B, Zheng J. Reference-Driven Undersampled MR Image Reconstruction Using Wavelet Sparsity-Constrained Deep Image Prior. *Comput Math Methods Med.* 2021;8865582.

[19] Yoo J, Jin KH, Gupta H, Yerly J, Stuber M, Unser M. Time-Dependent Deep Image Prior for Dynamic MRI. *IEEE Trans Med Imaging.* 2021;40(12):3337–3348.

[20] Feng L, Srichai MB, Lim RP, et al. Highly accelerated real-time cardiac cine MRI using k-t SPARSE-SENSE. *Magn Reson Med.* 2013;70(1):64–74.

[21] Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUS: a general compressed sensing framework for high resolution dynamic MRI. *Magn Reson Med.* 2009;61(1):103–116.

[22] Tsao J, Kozerke S. MRI temporal reconstruction techniques. *J Magn Res Imaging*. 2012;36(3):543–560.

[23] Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018;9446–9454.

[24] Uecker M, Lai P, Murphy MJ, et al. ESPIRIT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magn Reson Med.* 2014;71(3):990–1001.

[25] Uecker M, Tamir J. bart: version 0.4.00. [https://zenodo.org/record/2149128](https://zenodo.org/record/2149128)

[26] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. 2014; [http://arxiv.org/abs/1412.6980v9](http://arxiv.org/abs/1412.6980v9)

[27] Roemer PB, Edelstein WA, Hayes CE, Souza SP, Mueller OM. The NMR phased array. *Magn Reson Med.* 1990;16(2):192–225.

[28] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med.* 2007;58(6):1182–1195.

[29] Zbontar J, Knoll F, Sriram A, et al. fastMRI: An Open Dataset and Benchmarks for Accelerated MRI. 2019; [http://arxiv.org/abs/1811.08839v2](http://arxiv.org/abs/1811.08839v2)

[30] Paszke A, Gross S, Chintala S, et al. Automatic differentiation in PyTorch. 2017; [https://openreview.net/forum?id=BJJsrmfC2](https://openreview.net/forum?id=BJJsrmfC2)

[31] GitHub . Package pytorch-unet. [https://github.com/milesial/Pytorch-UNet/pkgs/container/pytorch-unet](https://github.com/milesial/Pytorch-UNet/pkgs/container/pytorch-unet) (Accessed: 2022-09-13).

[32] SigPy . sigpy 0.1.23 documentation. [https://sigpy.readthedocs.io/en/latest/](https://sigpy.readthedocs.io/en/latest/) (Accessed: 2022-08-09).

[33] Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory.* 2006;52(2):489–509.
The COMPaS architecture uses a U-Net with 2D convolutions \( \text{conv } 3 \times 3 \), where \( 3 \times 3 \) describes the 2D convolutional kernel’s size. The number of filter maps is depicted above the layers and reported in Tab. 1. 2D normalization layers and Rectified Linear Units (ReLU) are placed between the convolutional layers. Maximum pooling (max pool \( 2 \times 2 \)) with kernel size is used for down-sampling. Bilinear interpolation is used for up-sampling (up-conv \( 2 \times 2 \)). Zero-padding is used in convolutional layers to preserve the input images’ sizes. U-Net’s final output layer returns a \( \mathbb{R}^{2 \times N_x \times N_y} \) tensor. The first dimension’s two channels are treated as real and imaginary part of the image \( x \in \mathbb{R}^{N_x \times N_y} \).

b) In the loss function \( C \), the transformed network output \( k = FSx \) is compared to \( k_0 \) in the measured data-points and subjected to a sparsity constraint (here, L1 norm of wavelet coefficients is shown exemplary).
Figure 2 Pseudo-random Cartesian sampling patterns with undersampling in phase-encoding direction, for retrospective undersampling of FastMRI data, as described in Sec. 2.3.

Figure 3 Fourier transformed and phase-sensitively combined data for the fully-sampled Ground Truth (GT), as well as retrospectively undersampled data for various acceleration factors $R$. A single slice from file_brain_AXT2_201_2010066 of the FastMRI brain dataset is selected and shown here.

Figure 4 Comparison of PICS and COMPaS reconstruction with Total Variation (TV) sparsity regularization for different acceleration rates $R$. Error Map shows relative error ($\times 10$) to the Ground Truth (GT) reconstruction shown in Fig. 3. Quantitative error measures L1, MSE and SSIM compared to the GT are reported next to each reconstruction image. Inlay image shows zoom-in of the periventricular region (blue frame). In PICS, residual artifacts are especially prominent in the occipital region and are suppressed more efficiently in DIP (red arrows).
Figure 5  PICS and COMPaS reconstruction with L1 Wavelet (L1W) sparsity regularization for different acceleration factors $R$. Error Map shows relative error (×10) to the Ground Truth (GT) reconstruction displayed in Fig. 3. Quantitative error measures L1, MSE and SSIM compared to the GT are reported next to each reconstruction. Inlay image shows zoom-in of the periventricular region (blue frame). While SSIM is similar for both PICS and DIP, the latter achieves lower L1 and MSE error metrics.
Figure 6  Error metrics for PICS and COMPaS reconstructions with TV and L1W sparsity on a random subset of 200 samples from the FastMRI dataset. Pn in the legend indicates the n-th percentile. Asterisks indicate significance level of difference between COMPaS and PICS at same R, determined by a Mann-Whitney U rank test: * : P < 0.05, ** : P < 0.01, *** : P < 0.0001.
**Figure 7**  
**a)** Root sum of squares of calibration-free COMPaS reconstruction with Total Variation (TV) sparsity regularization for different acceleration rates \( R \). No knowledge of coil sensitivity maps is used for reconstruction. Error Map shows relative error (x10) to the Ground Truth (GT) reconstruction displayed in Fig. 3. Quantitative error measures \( L_1 \), MSE and SSIM compared to GT are reported next to each reconstruction. Inlay image shows zoom-in of the periventricular region (blue frame). Residual artifacts (red arrow) are more prominent than in reconstructions using coil sensitivity information (cf. Fig. 4 and 5).  
**b)** Root sum of squares of calibration-free COMPaS reconstruction with L1 Wavelet (L1W) sparsity regularization for different acceleration rates \( R \). No knowledge of coil sensitivity maps is used for reconstruction. Error Map shows relative error (x10) to the GT reconstruction shown in Fig. 3. Quantitative error measures \( L_1 \), MSE and SSIM compared to GT are reported next to each reconstruction. Inlay image shows zoom-in of the periventricular region (blue frame). Comparing to TV-regularized reconstruction (cf. panel a)), higher removal of ghosting artifacts is achieved (green arrow).
Figure 8  2-fold undersampled single coil reconstruction for PICS and COMPaS, both with TV and L1W sparsity. Quantitative image metrics L1, MSE and SSIM are reported next to the images. Relative error (x10) to Ground Truth (see Fig. 3) is shown right to each reconstruction. Inlay image shows zoom-in of the periventricular region (blue frame).