LANTERN: LEARN ANALYSIS TRANSFORM NETWORK FOR DYNAMIC MAGNETIC RESONANCE IMAGING

SHANSHAN WANG
Paul C. Lauterbur Research Center for Biomedical Imaging
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
Shenzhen, Guangdong, China
1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, China

YANXIA CHEN
Paul C. Lauterbur Research Center for Biomedical Imaging
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
Shenzhen, Guangdong, China
1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, China
University of Chinese Academy of Sciences, Beijing, China
19 Yuquan Road, Shijingshan District, Beijing, China

TAOHUI XIAO, LEI ZHANG, XIN LIU AND HAIRONG ZHENG*
Paul C. Lauterbur Research Center for Biomedical Imaging
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
Shenzhen, Guangdong, China
1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, China

Abstract. This paper proposes to learn analysis transform network for dynamic magnetic resonance imaging (LANTERN). Integrating the strength of CS-MRI and deep learning, the proposed framework is highlighted in three components: (i) The spatial and temporal domains are sparsely constrained by adaptively trained convolutional filters; (ii) We introduce an end-to-end framework to learn the parameters in LANTERN to solve the difficulty of parameter selection in traditional methods; (iii) Compared to existing deep learning reconstruction methods, our experimental results show that our paper has encouraging capability in exploiting the spatial and temporal redundancy of dynamic MR images. We performed quantitative and qualitative analysis of cardiac reconstructions at different acceleration factors (2×-11×) with different undersampling patterns. In comparison with two state-of-the-art methods, experimental results show that our method achieved encouraging performances.

1. Introduction. Dynamic magnetic resonance imaging is able to provide important anatomical and functional information in a spatial-temporal manner. However, a fundamental challenge of MRI is its slow imaging speed a.k.a long imaging time, which hinders its wide applications. To address this challenge, various efforts have been made ranging from prompting hardware to software developments such as parallel imaging using phased array coils [28], fast imaging sequences [11], and reduced scan with advanced image reconstruction algorithms.

2020 Mathematics Subject Classification. Primary: 58F15, 58F17; Secondary: 53C35.
Key words and phrases. Dynamic MR imaging, Deep learning, CS-MRI.
S. Wang and Y. Chen contributed equally to this work.
* Corresponding author: Hairong Zheng.
Our specific focus here is the signal processing based MR image reconstruction from incomplete k-space data. Since the acquisition time of k-space is proportional to its amount, this strategy accelerates MR scan by undersampling or partial sampling of k-space. Undersampling introduces violations of the Nyquist sampling theorem and may cause aliasing and blurring issues by direct inverse Fourier transform \[3, 29\]. To solve these issues, prior knowledges are normally incorporated in the reconstruction formulation as regulations \[39, 16, 17, 13\]. Specifically, through the support of the well-known compressed sensing (CS) theory, researchers have developed a series of dynamic image reconstruction methods by exploiting either spatial or temporal redundancy or both with different sampling patterns. For example, with a Random k-t sampling pattern, k-t FOCUSS \[8\] was proposed, whose special cases included the celebrated k-t BLAST and k-t SENSE \[32\]. There was also a k-t iterative support detection (k-t ISD) method to improve the CS dynamic MR imaging methods \[14\]. These methods along others have explored different sparsifying transforms such as total variation (TV), wavelet and dictionary learning in spatial domain \[2, 23, 4, 33\], Fourier transform \[20, 21\], finite difference \[15\], and principal component analysis \[9, 10\] in the temporal domain, and 3D transforms such as wavelet-Fourier transform \[22\] or 3D wavelet transform in the spatial-temporal domain \[19\]. Besides the predefined transforms, dictionary learning has also been investigated. For example, temporal gradient sparsity was explored by \[3\] with adaptively learned dictionary and a patch-based 3-D spatiotemporal dictionary was trained for sparse representations of the dynamic image sequence \[34\]. In addition to the sparsifying transforms, low rank has been utilized to complete missing or corrupted entries for a matrix as well. For example, Bo Zhao et al proposed combining partial separable function and sparsity constraints to improve MR reconstruction performance \[42\] and the L+S \[24\], k-t SLR \[35\] methods. These methods all made great contributions to dynamic MR imaging. Nevertheless, the prior knowledge utilized is still limited to a few samples or reference images \[35\]. Furthermore, the iterative reconstruction can be time-consuming and have parameters that are hard to tune.

Deep learning based MR image reconstruction is an emerging field to accelerate MR scan \[36, 40, 37, 12\]. There are model-based deep learning methods that formulate the prior regularization iterative reconstruction process into network learning process. Typical examples include variational network (VN-net) \[6\], alternating direction method of multipliers network (ADMM-net) \[30, 41\] and Model DL \[1\], etc. \[18, 27\]. In addition to the model based methods, there are also direct end-to-end learning techniques that identify the mapping relationship between the undersampled and fully-sampled pairs. The end-to-end deep learning method uses a priori of big data, generally requiring large and complex networks to learn Complex structure in MR data. Instances consist of AUTOMAP \[43\], U-net \[7\], KIKI-net \[5\], recursive dilated net \[31\], and so on so forth \[25\]. Among all of them, there are only a few end-to-end learning networks for dynamic MR imaging \[29, 26\]. These works directly learn the mapping relationship and have shown great experimental results. Nevertheless, the explanation of this work is more empirical, which has not taken advantage of the theoretically explainable compressed sensing framework.

To bridge the gap between the model-based dynamic imaging work and the empirical direct map learning framework, this work proposes a convolutional analysis transform network learning for dynamic magnetic resonance imaging with convolutional network dubbed as LANTERN. Compared to end-to-end deep learning
methods, the proposed method does not require much data or complex networks. The network focuses on learning the parameters for the compressed sensing models. Specifically, we initialize the discrete cosine transform (DCT) in the spatial domain and total variation (TV) in the temporal domain to fully exploit the redundancy of dynamic image sequences. To optimize the models, we use Alternating Direction Method of Multipliers (ADMM) which is a variable separable method. Our experiments demonstrate that the proposed scheme can effectively reconstruct dynamic MR images accurately and rapidly. Compared with the state-of-the-art methods, D5C5 [29] and k-t SLR [15], our method presents encouraging performance both quantitatively and qualitatively.

2. Methods.

2.1. Dynamic imaging model. In dynamic MR imaging, the measured signal $y_t \in \mathbb{C}^M$ at time $t$ can be described as follows, and $t = \{1, 2, 3...Q\}$.

$$y_t = F_t x_t + \eta_t$$

where $F_t \in \mathbb{C}^{M \times N}$ is the measurement matrix and $\eta_t \in \mathbb{C}^M$ is the measurement noise for the $t$-th vectorized cardiac phase image $x_t \in \mathbb{C}^N$, $F_t = P_t F_{2D}$, $P_t$ is an $M \times N$ undersampling matrix whose rows are extracted from an $N \times N$ identity matrix according to the k-space sampling locations at time $t$ $(M << N)$. $F_{2D} \in \mathbb{C}^{N \times N}$ is the unitary matrix representing the 2D discrete Fourier transform. Suppose a total of $Q$ cardiac phases are acquired, the entire acquisition process can be described as follows:

$$y = F_u x + e$$

where $x = \{x_1^H, x_2^H, \ldots, x_t^H, \ldots, x_Q^H\} \in \mathbb{C}^{NQ \times 1}$ represents the stacked $Q$ phase images, $e = \{\eta_1^H, \eta_2^H, \ldots, \eta_t^H, \ldots, \eta_Q^H\} \in \mathbb{C}^{MQ \times 1}$, $y = \{y_1^H, y_2^H, \ldots, y_t^H, \ldots, y_Q^H\} \in \mathbb{C}^{MQ \times 1}$ is the under-sampled k-space data, and $F_u = \text{diag}\{F_1, F_2, \ldots, F_t, \ldots, F_Q\} \in \mathbb{C}^{MQ \times NQ}$, $H$ is the hermitian transpose operation.

2.2. Sparse convolutional coding feature preserving prior model. The recovery of $x$ form $y$ underdetermined problem because $M << N$. To reconstruct $x$ we introduce a sparse coding convolutional feature preserving model to overcome the ill-posedness nature and propose the following model.

$$\arg\min_x \left\{ \frac{1}{2} \| F_u x - y \|_2^2 + \sum_{l=1}^L \lambda_l \text{Pri}(\Phi_l x) \right\}$$

where $\frac{1}{2} \| F_u x - y \|_2^2$ is the data fidelity term, and $\text{Pri}(\cdot)$ is a regularization prior, which can be defined as $l_q$ sparse regularizer, with $q \in (0, 1)$. $\Phi_l$ denotes a sparse coding convolutional operator that extracting image features. $\lambda_l$ is the regularization parameter. $l$ represents the $l$-th filter, and $l = \{1, 2, 3, \ldots, L\}$. To solve this, we introduce auxiliary variables $v = \{v_1^T, v_2^T, \ldots, v_l^T, \ldots, v_Q^T\} \in \mathbb{C}^{NQ \times 1}$ and get the following constrained formulation:

$$\arg\min_{x, v} \left\{ \frac{1}{2} \| F_u x - y \|_2^2 + \sum_{l=1}^L \lambda_l \text{Pri}(\Phi_l x) \right\} \quad \text{s.t.} \quad v = x$$
By adopting the augmented Lagrangian technique, the constrained problem in Eq. 4 can be transformed into the following unconstrained one:

\[
\mathcal{L}_\rho(x, v, \alpha) = \frac{1}{2} \| F_u x - y \|^2_2 + \sum_{l=1}^L \lambda_l Pri(\Phi_l x) - \langle \alpha, x - v \rangle + \frac{\rho}{2} \| x - v \|^2_2
\]

(5)

Where \( \alpha \) are Lagrangian multipliers, \( \rho (\rho > 0) \) represents the scaling factor, the above optimization problem can be further divided into three subproblems by using the alternating direction multiplier method with an assistant variable \( \beta = \frac{\alpha}{\rho} \) introduced.

2.3. Alternating direction minimization algorithm. To address the above problem, the alternating direction minimization algorithm can be adopted, reaching:

\[
\begin{align*}
& \arg \min_x \{ \frac{1}{2} \| F_u x - y \|^2_2 + \frac{\rho}{2} \| x + \beta - v \|^2_2 \\
& \arg \min_v \{ \frac{\rho}{2} \| x + \beta - v \|^2_2 + \sum_{l=1}^L \lambda_l Pri(\Phi_l v) \} \\
\beta & \leftarrow \beta + \tilde{\tau}(x - v)
\end{align*}
\]

(6)

where \( \tilde{\tau} \) denotes the learning rate.

2.4. Solve the above subproblems.

2.4.1. Solve the subproblems \( x \). Adopting least squares to solve the first subproblem in Eq. 6, we have:

\[
(F_u^H F_u + \rho I)x = F_u^H y + \rho(v - \beta)
\]

where the superscript has been eliminated for simplicity and clarity. Then further let \( F_u = PF \) with \( P = \text{diag}\{P_1, P_2, \ldots, P_t, \ldots, P_Q\} \in R^{MQ \times NQ} \) and \( F = \text{diag}\{F_{2D}, F_{2D}, \ldots, F_{2D}\} \in R^{MQ \times NQ} \), we have the following solution:

\[
x = F^H (P^H P + \rho I)^{-1}[P^H y + \rho F(v - \beta)]
\]

where \( P^H P \) is a diagonal matrix that can be efficiently computed.

2.4.2. Solve the subproblems \( v \). For the update of \( v \), we adopt the gradient descent method. Specifically, with the gradient \( \nabla v = \rho(v - x - \beta) + \sum_{l=1}^L \lambda_l \Phi_l^H \mathcal{H}_{Pri}(\Phi_l v) \), we have:

\[
v^{(i)} = v^{(i-1)} - l_r \nabla v = (1 - l_r \rho) v^{(i-1)} + l_r \rho (x + \beta) - \sum_{l=1}^L \lambda_l l_r \Phi_l^H \mathcal{H}_{Pri}(\Phi_l v)
\]

where \( \mathcal{H}_{Pri} \) represents the gradient (or sub-gradient) of the prior regulation function \( Pri(\cdot) \) and \( l_r \) is the step size.

2.4.3. Solve the subproblems \( \beta \). Adopts the gradient descent method, we have:

\[
\beta^{(i)} = \beta^{(i-1)} + \tilde{\tau}^{(i)}(x^{(i)} - v^{(i)})
\]
2.4.4. The solution of the subproblems. From the above solution process, we can easily get the following results:

\[
\begin{align*}
\mathbf{x}^{(i)} &= \mathbf{F}^H (\mathbf{P}^H \mathbf{P} + \rho^{(i)} \mathbf{I})^{-1} [\mathbf{P}^H \mathbf{y} + \rho^{(i)} \mathbf{F}(\mathbf{v}^{(i-1)} - \beta^{(i-1)})] \\
\mathbf{v}^{(i,k)} &= \mu_1^{(i,k)} \mathbf{v}^{(i-1)} + \mu_2^{(i,k)} (\mathbf{x}^{(i)} + \beta^{(i-1)}) - \sum_{l=1}^{L} \lambda_l \Phi_l^H \mathcal{H}_{Pr1}(\Phi_l \mathbf{v}^{(i-1)}) \\
\beta^{(i)} &= \beta^{(i-1)} + \tilde{\tau}^{(i)} (\mathbf{x}^{(i)} - \mathbf{v}^{(i)})
\end{align*}
\]

(7)

where \( \mu_1 = 1 - l_r \rho, \mu_2 = l_r \rho, \lambda = \lambda_l r, i = \{1, 2, \ldots, N_i\} \) denote \( i - \text{th} \) iteration and \( k = \{1, 2, \ldots, N_k\} \) represents the \( k - \text{th} \) iteration of the subproblem \( \mathbf{v} \). In experiments we found that when \( N_k = 1 \), there was a better reconstruction effect, which is consistent with the conclusion of reference [30]. After many experimental verifications, we set \( N_i = 13 \), which means that the number of iterations per epoch is 13.

3. Network architecture. Traditionally, these variables are directly updated with the undersampled data and minimization algorithms. However, it is hard to tune the weighting parameter and the prior information utilized is limited. Inspired by [30], we propose to use deep learning to adaptively learn sparse convolution kernels and the weighting parameters. Specifically, we rewrite the above Eq. 7 into the following Eq. 8 by splitting \( \mathbf{v}^{(i,k)} \) into the *Addition* layer, *Conv1* layer, *Nonlinear* layer and *Conv2* layer. In particular, we consider filter \( \Phi \) as convolution kernel. And \( \mathcal{H}_{Pr1} \) is approximated by learning a piecewise linear function \( S_{PLF}(\cdot) \).

\[
\begin{align*}
\text{Recon} : \quad & \mathbf{x}^{(i)} = \mathbf{F}^H (\mathbf{P}^H \mathbf{P} + \rho^{(i)} \mathbf{I})^{-1} [\mathbf{P}^H \mathbf{y} + \rho^{(i)} \mathbf{F}(\mathbf{v}^{(i-1)} - \beta^{(i-1)})] \\
\text{Addition} : \quad & \mathbf{v}^{(i,k)} = \mu_1^{(i,k)} \mathbf{v}^{(i-1)} + \mu_2^{(i,k)} (\mathbf{x}^{(i)} + \beta^{(i-1)}) - \sum_{l=1}^{L} \lambda_l \Phi_l^H \mathcal{H}_{Pr1}(\Phi_l \mathbf{v}^{(i-1)}) \\
\text{Conv1} : \quad & C_1^{(i,k)} = \sum_{l=1}^{L} (\mathbf{w}_{1,l}^{(i,k)} \ast \mathbf{v}^{(i,k-1)} + \mathbf{b}_{1,l}^{(i,k)}) \\
\text{Nonlinear} : \quad & h^{(i,k)} = S_{PLF}(C_1^{(i,k)}, \{p_n, q_n\}_{n=1}^{N_n}) \\
\text{Conv2} : \quad & C_2^{(i,k)} = \sum_{l=1}^{L} (\mathbf{w}_{2,l}^{(i,k)} \ast h^{(i,k)} + \mathbf{b}_{2,l}^{(i,k)}) \\
\text{Multi} : \quad & \beta^{(i)} = \beta^{(i-1)} + \tilde{\tau}^{(i)} (\mathbf{x}^{(i)} - \mathbf{v}^{(i)})
\end{align*}
\]

(8)

The iteration process is formulated into a network flow as shown in Fig. 1 (A). The network consists of a forward process to reconstruct the image and a backpropagation process to update the parameters in the network. For the deep learning process, the works go through two stages: the training state and testing state. The training loss is the mean squared error between reconstructed image and label.

1) *Recon* stands for reconstruction layer \( \mathbf{x}^{(i)} \), the input of which layer is the undersampled k-space data \( \mathbf{y}, \mathbf{v}^{(i-1)} \) and \( \beta^{(i-1)} \).

\[
\text{R}_{org} : \quad \mathbf{x}^{(1)} = \mathbf{F}^H (\mathbf{P}^H \mathbf{P} + \rho^{(1)} \mathbf{I})^{-1} (\mathbf{P}^H \mathbf{y})
\]

\[
\text{R}_{mid} \text{ and } \text{R}_{final} : \quad \mathbf{x}^{(i)} = \mathbf{F}^H (\mathbf{P}^H \mathbf{P} + \rho^{(i)} \mathbf{I})^{-1} [\mathbf{P}^H \mathbf{y} + \rho^{(i)} \mathbf{F}(\mathbf{v}^{(i-1)} - \beta^{(i-1)})]
\]
Note that the **Recon** layer includes three different forward propagation equations $R_{org}$, $R_{mid}$ and $R_{final}$ due to when $i = 1$ the input $y^{(i-1)}$, $\beta^{(i-1)}$ is 0. Although these are different iteration formulas, they are all derived from Eq. 8.

2) **Addition** stands for the addition layer $v^{(i,k)}$, which performs simple weight summation operation. The input is $v^{(i,k-1)}$, $x^{(i)}$, $\beta(i-1)$ and $C_2^{(i,k)}$.

$$A_{org} : v^{(1,0)} = \mu_2^{(i,0)} x^{(1)}, \text{ initialization } C_2^{(1,0)} = 0$$

$$A_{mid} : v^{(1,k)} = \mu_1^{(1,k)} v^{(1,k-1)} + \mu_2^{(1,k)} x^{(1)} + C_2^{(1,k)}$$

$$A_{1org} : v^{(i,0)} = \mu_2^{(i,0)} (x^{(i)} + \beta(i-1))$$

$$A_{1mid} : v^{(i,k)} = \mu_1^{(i,k)} v^{(i,k-1)} + \mu_2^{(i,k)} (x^{(i)} + \beta(i-1)) - C_2^{(i,k)}$$

3) **Conv1** and **Conv2** stand for convolution layers. The input is $v^{(i,k-1)}$ and $h^{(i,k)}$. In which we used 3D convolution is relatively large. In order to reduce the parameter quantity, we use 2D convolution in the time domain TV filter.

$$\text{Conv1} : C_1^{(i,k)} = \sum_{l=1}^L (w_{1,l}^{(i,k)} \ast v^{(i,k-1)} + b_{1,l}^{(i,k)})$$

$$\text{Conv2} : C_2^{(i,k)} = \sum_{l=1}^L (w_{2,l}^{(i,k)} \ast h^{(i,k)} + b_{2,l}^{(i,k)})$$

4) **Nonlinear** Stands for nonlinear layer $h^{(i,k)}$.

$$h^{(i,k)} = S_{PLF}(C_1^{(i,k)}, \{p_n, q_n^{(i,k)}\}_{n=1}^{N_n})$$

Where input is $C_1^{(i,k)}$, $\{p_n, q_n^{(i,k)}\}_{n=1}^{N_n}$, $\{p_n\}_{n=1}^{N_n}$ are $N_n$ points uniformly located within $[-1, 1]$, where $N_n = 101$, and $\{q_n^{(i,k)}\}_{n=1}^{N_n}$ are the values at different points $n$ for $i-th$ iterate.

5) **Multi** Stand for multiplier update layer $\beta^{(i)}$, whose input is $\beta^{(i-1)}$, $x^{(i)}$ and $v^{(i)}$.

$$M_{org} : \beta^{(1)} = \tilde{\beta}(1)(x^{(1)} - v^{(1)})$$

$$M_{mid} and M_{final} : \beta^{(i)} = \beta^{(i-1)} + \tilde{\beta}(i)(x^{(i)} - v^{(i)})$$

Similarly, the **Addition** and **Multi** layers are also different in the formula because of the number of iterations of the input, which can be derived from Eq. 8.

6) **Loss** The loss layer can be calculated from the reconstructed dynamic MR image (dMRI) and the original fully sampled dMRI by a normalized mean square error (NMSE).

$$\text{Loss} : E(\Theta) = \frac{1}{|A|} \sum_{(y, \text{Ref}) \in A} \sqrt{\|\text{Rec}(y, \Theta) - \text{Ref}\|_2^2}$$

where, $A$ represents the number of training sets. $\text{Rec}(y, \Theta)$ is the LANTERN network output based on under-sampled data $y$ in k-space and parameter $\Theta$ and $\text{Ref}$ is fully sampled dMRI. With the final trained network, the forward process can reconstruct a high-quality, artifact-free image that is close to Ground Truth with the undersampled k-space data. Fig. 1 (B) shows the specific process of the $i-th$ iteration, each layer corresponding to the following forward propagation formula, wherein the detail process of the **prior** layer is described in Fig. 1 (C). The **prior**
layer includes Conv1, Nonlinear and Conv2, where Conv1 and Conv2 learn filter operators that make the image sparse. A piecewise linear function is used to approximate the derivative of the regularization function.

The method proposed in this paper mainly uses deep learning to find the optimal parameters for CS-MRI reconstruction. These parameters are small in size compared to end-to-end deep learning works.

**Figure 1.** The proposed LANTERN network architecture for dMRI reconstruction. In (A) and (B), the blue arrow indicates forward process. The pink arrow indicates the process of back-propagation to update network parameters, where \( i \) represents the \( i^{th} \) iteration and \( N_i \) represents a total of \( N_i \) iterations. \( k \) expresses that the priori loop for \( k \) times and \( (i, k) \) means that in the \( i^{th} \) iteration, the a priori loops \( k \) times.
Table 1. Experimental masks and acceleration factors

| 1D Random | 2D Radial |
|-----------|-----------|
| 2X        | 2X        |
| 3X        | 3X        |
| 4X        | 4X        |
| 5X        | 5X        |
| 7X        | 7X        |
| 9X        | 9X        |
| 11X       | 11X       |
| 15X       | 15X       |

4. Experiments and results.

4.1. Performance evaluation. To get a quantitative assessment, we used common image quality evaluation, Peak signal-to-noise ratio (PSNR), Structural Similarity Index (SSIM) and High-frequency error norm (HFEN).

\[
\text{PSNR} = 20 \log_{10} \frac{\text{max}(\text{Ref}) \sqrt{N}}{\| \text{Ref} - \text{Rec} \|_2^2}
\]

\[
\text{SSIM} = \frac{(2\mu_{\text{Ref}}\mu_{\text{Rec}} + c_1)(2\sigma_{\text{RefRec}} + c_2)}{(\mu_{\text{Ref}}^2 + \mu_{\text{Rec}}^2 + c_1)(\sigma_{\text{Ref}}^2 + \sigma_{\text{Rec}}^2 + c_2)}
\]

\[
\text{HFEN} = \frac{1}{N} \sum_{i=1}^{N} \frac{\| \text{LoG}(\text{Ref}_i) - \text{LoG}(\text{Rec}_i) \|_p^2}{\| \text{LoG}(\text{Ref}_i) \|_F^2}
\]

where \( \text{Ref} \) is fully sampled dMRI, \( \text{Rec} \) denotes the LANTERN network output and \( N \) is the total number of image pixels. In SSIM the \( \mu_{\text{Ref}} \) represents the average of \( \text{Ref} \), \( \sigma_{\text{Ref}}^2 \) represents the variance of \( \text{Ref} \), \( \sigma_{\text{RefRec}} \) represents the covariance of \( \text{Ref} \) and \( \text{Rec} \). \( c_1 = (k_1L), c_2 = (k_2L), L \) the dynamic range of pixel values, \( k_1 = 0.01 \) and \( k_2 = 0.03 \) (details shown in [38]). In HFEN, the \( \text{LoG}(\text{Ref}) \) is a Laplacian Gaussian filter function with a template size of 15 and a filter standard deviation of 1.5.

4.2. Configuration. We collected 101 fully sampled k-space data using the 3T SIEMENS scanner with the FLASH acquisition sequence, whose parameter configurations are 25 temporal frames, FOV 330 × 330 mm, acquisition matrix 192 × 192, slice thickness = 6 mm, TR = 50 ms, TE = 3 ms and 24 receiving coils. The raw multi-coil data of each frame was combined through the adaptive coil combination method to produce a combined-channel complex-valued image. Data augmentation strategies have been applied. We shear the original images along the \( x, y \) and \( t \) direction. The sheared size is 126 × 126 × 16 \((x \times y \times t)\). We removed the blurred data and the data containing motion artifacts or relatively small FOVs. We transferred the data into image space to make sure our images lie in the center of the FOV. So the data cutting is in image space and then we transfer it into k-space for data normalization purposes. By applying corresponding masks, images with different under-sampling patterns were obtained (See Table. 1 for details). Finally, 150 cardiac data were generated with 80 data for network training and 20 data for network validation, and 50 for network testing. The online model training took 45 hours on an Intel Xeon (R) CPU E5-2640 V4 @2.40GHz × 40, 64G.

Table. 1 shows the experimental setup in this paper, with a maximum acceleration factor of 11× in 1D Random undersampling mode and a maximum acceleration factor of 15× in 2D Radial undersampling mode. It is worth noting that the acceleration factors in the table are net acceleration factors.
4.3. Impact of dataset size. Fig. 2 shows the visual results comparison of the proposed method when the number of training data were increased from 50 to 120. And the average reconstruction quantitative results comparison of the 50 test data as shown in Table 2. The proposed method does not require a large amount of data and can reconstruct a good quality image even in the case of fewer data samples. From the results, it can be seen that when the training data volume is 100, there is already a relatively high reconstruction quality. When the training data is further increased, the reconstruction quality is not greatly improved which also suggests that the network is close to the optimal solution in the case of small data. In this article, we selected 100 images for experiments in subsequent works.

4.4. Effect of initializations: Gauss, TV, DCT, LANTERN. We used the methods of Random Gauss, TV, DCT, and LANTERN(DCT+TV) to initialize the proposed model and compare the experiments. Fig. 3 shows the reconstruction visual results under several different initialization methods. It can be seen that the reconstruction result initialized by Gaussian noise is the worst and the reconstruction result under LANTERN initialization is slightly better than DCT. The PSNR value of LANTERN is also the highest. Table 3 is the average quantization index corresponding to the reconstruction result under the data volume of 100 and 1D Random 4× acceleration, which also proves that the proposed LANTERN method is superior to the other initialization methods.

4.5. Comparison with state-of-the-art methods. To further verify the feasibility of the proposed model based dynamic MR image reconstruction algorithm, we compare the performance of the proposed method with the compressed sensing based k-t SLR technique and the data-driven based D5C5 algorithm. Considering that D5C5 is a method based on big data sets, if we use 100 data to train, the reconstruction results are not good enough and the experimental results also verify this conjecture. So for the D5C5 method, we cut our data to 126×126×16 and
Table 2. Quantitative results comparison for the sensitivity to the training data size. The average quantitative indicator values of the results reconstructed for the 50 test data with the network trained from different different amount of data with 1D Random sampling pattern at an accelerated factor of 4.

| 1D Random | NMSE | PSNR/dB | SSIM  | HFEN  |
|-----------|------|---------|-------|-------|
| data50    | 0.0413 | 40.8047 | 0.8943 | 0.8333 |
| data60    | 0.0397 | 41.1515 | 0.9   | 0.7939 |
| data80    | 0.0388 | 41.3589 | 0.9034 | 0.7729 |
| data100   | 0.0385 | 41.4391 | 0.9043 | 0.7633 |
| data120   | 0.0386 | 41.4402 | 0.9035 | 0.7685 |

Figure 3. The comparison of the three initialization modes of Random Gaussian, TV, DCT and LANTERN based on the proposed method with 1D Random sampling at an acceleration factor of 4. PSNR value are given under the results.

Table 3. Quantitative results comparison for the sensitivity to the initialization. The average quantitative indicator values of the results reconstructed for the 50 test data with the network trained with different initialization with 1D Random sampling pattern at an accelerated factor of 4.

| 1D Random | Gaussian | TV | DCT | LANTERN |
|-----------|---------|----|-----|---------|
| AVE       | PSNR    | HFEN | PSNR | HFEN    | PSNR | HFEN    |
|           | 39.8089 | 0.9459 | 40.5884 | 0.8514 | 40.9971 | 0.8064 | 41.4391 | 0.7633 |

get 3200 data, 2900 data for training, 300 data for testing, and then calculated the average PSNR, SSIM and HFEN. It can be seen from the results of 1D Random 4× and 5× acceleration factors in Fig. 4 and Fig. 5 that the reconstructed image by the k-t SLR method is somewhat blurred. The reconstruction visual effect of the D5C5 algorithm is close to the proposed method, but the PSNR value is the highest from the proposed method. Table. 4 compares the quantitative indicators of 7×, and 11× acceleration factors under 1D Random sampling. We can see that
the proposed method has the best index. We also plotted a quantitative index plot of the reconstruction results from $2\times$ to $11\times$ acceleration factors, as shown in Fig. 6. The performance of all methods decline as the acceleration factor increases. However, our approach has always maintained optimal performance.

![Figure 4](image-url)

**Figure 4.** The comparison of k-t SLR, D5C5 and the proposed method with 1D Random sampling at an acceleration factor of 4. PSNR value is given under the results.

![Figure 5](image-url)

**Figure 5.** The comparison of k-t SLR, D5C5 and the proposed method with 1D Random sampling at an acceleration factor of 5. PSNR value is given under the results.

In addition, we also conducted an experimental comparison of the radial undersampling trajectory. Table. 5 shows the average quantified index values of the test data with the radial sampling at $11\times$ and $15\times$ acceleration factors. It shows that our method is far superior to the comparison algorithm in PSNR, SSIM or HFEN. The comparison algorithm D5C5 and k-t SLR have their own advantages. The PSNR and HFEN values of D5C5 are better than k-t SLR, but the SSIM value of k-t SLR is relatively better than that of D5C5. We give a visual comparison of the reconstruction results of the radial sampling under the $11\times$ acceleration factor 7. The reconstruction results of the k-t SLR and D5C5 algorithms are both fuzzy. The proposed method is closest to the original image, and can also be seen from the
Table 4. Average reconstruction quantitative metrics with standard deviation of the 50 test data based on various methods with 1D Random sampling at a different accelerated factor.

| Methods     | Random 7X       |           | Random 11X      |           |
|-------------|-----------------|-----------|-----------------|-----------|
|             | PSNR | SSIM  | HFEN | PSNR | SSIM | HFEN |
| Zero-filling| 29.14±2.2 | 0.57±0.04 | 2.73±0.65 | 27.58±2.08 | 0.51±0.04 | 3.12±0.74 |
| Kt-SLR      | 33.50±2.70 | 0.77±0.03 | 1.83±0.53 | 32.44±2.61 | 0.73±0.03 | 2.01±0.63 |
| D5C5        | 36.76±2.00 | 0.78±0.03 | 1.40±0.33 | 35.22±2.00 | 0.73±0.03 | 1.82±0.50 |
| Proposed    | 37.48±2.45 | 0.82±0.02 | 1.31±0.36 | 35.40±2.60 | 0.77±0.03 | 1.67±0.53 |

Figure 6. The comparison of various methods between average quantification index of the 50 test data and acceleration factor based on 1D Random sampling.

error map. Fig. 8 is a graph showing the variation of the average quantization index of the test data of several methods with the increase of the acceleration factor, and the results are consistent with those described in Table 5.

Figure 7. The comparison of k-t SLR, D5C5 and the proposed method with 2D Radial sampling at an acceleration factor of 11. PSNR value is given under the results.

4.6. Convergence analysis. The results in the previous section demonstrate that our method has the best test performance. Due to the small number of samples used, over-fitting problems may occur in machine learning studies. To make sure that our method has properly converged, we give the training and validation error
Table 5. Average reconstruction quantitative metrics with standard deviation of the 50 test data based on various methods with radial sampling at a different accelerated factor.

| Methods | Radial 11X | Radial 15X |
|---------|------------|------------|
|         | PSNR/dB    | SSIM       | HFEN  | PSNR/dB    | SSIM       | HFEN  |
| Zero-filling | 22.269±1.37 | 0.345±0.06 | 5.198±0.72 | 20.153±1.27 | 0.275±0.05 | 5.986±0.67 |
| Kt-SLR  | 31.961±2.34 | 0.718±0.03 | 2.179±0.51 | 31.518±2.36 | 0.707±0.04 | 2.229±0.54 |
| D5C5    | 34.954±2.08 | 0.761±0.03 | 1.735±0.42 | 34.248±2.04 | 0.677±0.03 | 1.907±0.45 |
| Proposed | 38.874±2.28 | 0.831±0.03 | 1.019±0.26 | 38.115±2.28 | 0.808±0.03 | 1.164±0.30 |

Figure 8. The comparison of various methods between average quantification index of the 50 test data and acceleration factor based on Radial sampling.

curves of the proposed model to monitor the training. As shown in Fig. 9, a total of 400 epochs are trained, and the training and validation errors are at the same level, which have not created the over fitting issue by seeing the training and validation errors are almost at the same level.

Figure 9. The training and validation loss curves of the proposed model.
5. Discussion. Our method simultaneously sparsely constrains the spatial and temporal domains of dynamic magnetic resonance images and uses the end-to-end framework to learn the parameters. By comparing with the most advanced dynamic magnetic resonance image reconstruction algorithm, as shown in Fig. 4, Fig. 5 and Fig. 7, it can be seen from the visual effect that the proposed method can reconstruct clearer results and the image detail content recovery is more accurate. Under the 4x acceleration factor of 1D Random sampling, the contrast algorithm D5C5 is slightly better than the k-t SLR, and under the 1D Random sampling 5x acceleration factor, D5C5 is significantly better than the k-t SLR. The reconstruction results of the two contrast algorithms under the 11x acceleration factor of radial sampling have their own advantages and disadvantages. Although D5C5 is higher in PSNR value, D5C5 is better for image smoothing, while k-t SLR contains some noise-like artifacts. Comparatively, the reconstruction accuracy of the proposed method is higher. Fig. 5 and Fig. 7 further demonstrate that the results of the proposed method are superior to the comparison algorithm in terms of various quantitative indicators.

Therefore, due to sparse constraints in space and time and combining traditional constraint methods with deep learning, the proposed algorithm improves the image reconstruction ability and compensates for the limitations of pure deep learning algorithms that require significant amounts of data. We can train a better reconstruction model with less sample data than D5C5 algorithm which requires a large amount of data.

6. Conclusion. This work proposed a model-based convolutional dynamic MR imaging framework. The framework is able to fully exploit the spatial and temporal redundancy of dynamic MR images. The experiment results have shown that the proposed method can reconstruct comparable and even better MR images than the state-of-the-art algorithms in a shorter time under the same acceleration factors. In the future, we may investigate unsupervised learning for dynamic MR imaging.

Acknowledgments. This research was partly supported by the National Natural Science Foundation of China (61871371, 81830056, 81801691, 61671441), Key-Area Research and Development Program of GuangDong Province (2018B010109009), Science and Technology Planning Project of Guangdong Province (2017B020227012), the Basic Research Program of Shenzhen (JCYJ20180507182400762), Youth Innovation Promotion Association Program of Chinese Academy of Sciences (2019351). We would like to acknowledge Ian Zhang, our intern, for helping with the proofreading of our paper.

REFERENCES

[1] H. K. Aggarwal, M. P. Mani and M. Jacob, Modl: Model-based deep learning architecture for inverse problems, *IEEE Transactions on Medical Imaging*, 38 (2018), 394–405.

[2] K. T. Block, M. Uecker and J. Frahm, Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint, *Magnetic Resonance in Medicine*, 57 (2007), 1086–1098.

[3] J. Caballero, A. N. Price, D. Rueckert and J. V Hajnal, Dictionary learning and time sparsity for dynamic MR data reconstruction, *IEEE Transactions on Medical Imaging*, 33 (2014), 979–994.
[4] L. Chaari, J. C. Pesquet, A. Benazza-Benyahia and P. Ciuciu, A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging, *Med. Image Anal.*, 15 (2011), 185–201.

[5] T. Eo, Y. Jun, T. Kim, J. Jang, H. Lee and D. Hwang, KIKI-Net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images, *Magnetic Resonance in Medicine*, 80 (2018), 2188–2201.

[6] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock and F. Knoll, Learning a variational network for reconstruction of accelerated MRI data, *Magnetic Resonance in Medicine*, 79 (2018), 3055–3071.

[7] Y. Han, J. Yoo, H. H. Kim, H. J. Shin, K. Sung and J. C. Ye, Deep learning with domain adaptation for accelerated projection–reconstruction MR, *Magnetic Resonance in Medicine*, 80 (2018), 1189–1205.

[8] H. Jung, K. Sung, K. S. Noyak, E. Y. Kim and J. C. Ye, K-T FOCUSS: A general compressed sensing framework for high resolution dynamic MRI, *Magnetic Resonance in Medicine, An Off. J. Int. Soc. Magn. Reson. Med.*, 61 (2009), 103–116.

[9] H. Jung, J. Yoo and J. C. Ye, Generalized kt BLAST and kt SENSE using FOCUSS, 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, (2007), 145–148.

[10] H. Jung and J. C. Ye, Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques, *Int. J. Imaging Syst. Technol.*, 20 (2010), 81–98.

[11] W. A. Kaiser and E. Zeitler, MR imaging of the breast: Fast imaging sequences with and without Gd-DTPA. Preliminary observations, *Radiology*, 170 (1989), 681–686.

[12] F. Knoll, T. Murrell, A. Sriram, N. Yakubova, J. Zbontar, M. Rabbat, A. Defazio, M. J. Muckley, D. K. Sodickson, C. L. Zitnick and M. P. Recht, Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge, arXiv preprint, arXiv:2001.02518, 2020.

[13] D. Liang, J. Cheng and Z. Ke Z, Deep magnetic resonance image reconstruction: Inverse problems meet neural networks, *IEEE Signal Processing Magazine*, 37 (2020), 141–151.

[14] D. Liang, E. V. R. DiBella, R. R. Chen and L. Ying, K-t ISD: Dynamic cardiac MR imaging using compressed sensing with iterative support detection, *Magnetic Resonance in Medicine*, 68 (2012), 41–53.

[15] S. G. Lingala, Y. Hu, E. DiBella and M. Jacob, Accelerated dynamic MRI exploiting sparsity and low-rank structure: K-t SLR, *IEEE Transactions on Medical Imaging*, 30 (2011), 1042–1054.

[16] Q. Liu, Q. Yang, H. Cheng, S. Wang, M. Zhang and D. Liang, highly undersampled magnetic resonance imaging reconstruction using autoencoder priors, *Magnetic Resonance in Medicine*, 83 (2020), 322–336.

[17] F. Liu, D. Li, X. Jin, W. Qiu, Q. Xia and B. Sun, Dynamic cardiac MRI reconstruction using motion aligned locally low rank tensor (MALLRT), *Magnetic Resonance in Medicine*, 66 (2020), 104–115.

[18] Y. Liu, Q. Liu, M. Zhang, Q. Yang, S. Wang and D. Liang, IFR-net: Iterative feature refinement net-work for compressed sensing MRI, *IEEE Transactions on Computational Imaging*, 6 (2019), 434–446.

[19] M. Lustig, J. M. Santos, D. L. Donoho and J. M. Pauly, KT sparse: high frame-rate dynamic magnetic resonance imaging exploiting spatio-temporal sparsity, *U.S. Patent*, 7 (2009), 183.

[20] A. Majumdar, Improved dynamic MRI reconstruction by exploiting sparsity and rank-deficiency, *Magn. Reson. Imaging*, 31 (2013), 789–795.

[21] A. Majumdar, R. K. Ward and T. Aboulnaasr, Non-convex algorithm for sparse and low-rank recovery: Application to dynamic MRI reconstruction, *Magn. Reson. Imaging*, 31 (2013), 448–455.

[22] A. Majumdar, Improving synthesis and analysis prior blind compressed sensing with low-rank constraints for dynamic MRI reconstruction, *Magn. Reson. Imaging*, 33 (2015), 174–179.

[23] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin, An iterative regularization method for total variation-based image restoration, *Multiscale Model. Simul.*, 4 (2005), 460–489.
[24] R. Otazo, E. Cands and D. K. Sodickson, Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components, Magnetic Resonance in Medicine, 73 (2015), 1125–1136.

[25] T. M. Quan, T. Nguyen-Duc and W.-K. Jeong, Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss, IEEE Transactions on Medical Imaging, 37 (2018), 1488–1497.

[26] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V Hajnal and D. Rueckert, Convolutional recurrent neural networks for dynamic MR image reconstruction, IEEE Transactions on Medical Imaging, 38 (2019), 280–290.

[27] M. Rizkinia and M. Okuda, Evaluation of primal-dual splitting algorithm for MRI reconstruction using spatio-temporal structure Tensor and L1-2 norm, Makara Journal of Technology, 23 (2020), 126–130.

[28] D. k. Sodickson and W. J. Manning, Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays, Magnetic Resonance in Medicine, 38 (1997), 591–603.

[29] J. Schlemper, J. Caballero, J. V Hajnal, A. N. Price and D. Rueckert, A deep cascade of convolutional neural networks for dynamic MR image reconstruction, IEEE Transactions on Medical Imaging, 37 (2018), 491–503.

[30] J. Sun, H. Li and Z. Xu, Deep ADMM-Net for compressive sensing MRI, Advances in Neural Information Processing Systems, (2016), 10–18. http://papers.nips.cc/paper/6406-deep-admm-net-for-compressive-sensing-mri.

[31] L. Sun, Z. Fan, Y. Huang, X. Ding and J. Paisley, Compressed sensing MRI using a recursive dilated network, Thirty-Second AAAI Conference on Artificial Intelligence, (2018). http://www.columbia.edu/~jwp2128/Papers/SunFanetal2018.pdf

[32] J. Tsao, P. Boesiger and K. P. Pruessmann, k-t BLAST and k-t SENSE: Dynamic MRI With High Frame Rate Exploiting Spatiotemporal Correlations, Magnetic Resonance in Medicine, 50 (2003), 1031–1042.

[33] S. Wang, Y. Xia, Q. Liu, P. Dong and D. Feng, Fenchel duality based dictionary learning for restoration of noisy images, IEEE Transactions on Image Processing, 22 (2013), 5214–5225.

[34] Y. Wang, Y. Zhou and L. Ying, Undersampled dynamic magnetic resonance imaging using patch-based spatiotemporal dictionaries, 2013 IEEE 10th International Symposium on Biomedical Imaging, (2013), 294–297.

[35] S. Wang, Z. Ke, H. Cheng, S. Jia, Y. Leslie, H. Zheng and D. Liang, Dimension: Dynamic mr imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training, NMR in Biomedicine, (2019), e4131.

[36] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng and D. Liang, Accelerating magnetic resonance imaging via deep learning, 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), (2016).

[37] S. Wang, H. Cheng, L. Ying, T. Xiao, Z. Ke, H. Zheng and D. Liang, DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution, Magnetic Resonance in Medicine, 68 (2020), 136–147.

[38] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, 13 (2004), 600–612. https://ece.uwaterloo.ca/~z70wang/publications/ssim.pdf

[39] J. Yao, Z. Xu, X. Huang and J Huang, Accelerated dynamic MRI reconstruction with total variation and nuclear norm regularization, International Conference on Medical Image Computing and Computer-Assisted Intervention, (2015), 635–642.

[40] G. Yang, S. Yu and H. Dong, DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction, IEEE Transactions on Medical Imaging, 37 (2018), 1310–1321.

[41] Y. Yang, J. Sun, H. Li and Z Xu, ADMM-CSNet: A deep learning approach for image compressive sensing, IEEE Transactions on Pattern Analysis and Machine Intelligence, 42 (2018), 521–538.
[42] B. Zhao, J. P. Haldar, A. G. Christodoulou and Z.-P. Liang, Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints, *IEEE Transactions on Medical Imaging, 31* (2012), 1809–1820.

[43] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen and M. S. Rosen, Image reconstruction by domain-transform manifold learning, *Nature, 555* (2018), 487–492.

Received December 2019; revised May 2020.

E-mail address: sophiasswang@hotmail.com  
E-mail address: yx.chen@siat.ac.cn  
E-mail address: th.xiao@siat.ac.cn  
E-mail address: l.zheng@siat.ac.cn  
E-mail address: xin.liu@siat.ac.cn  
E-mail address: hr.zheng@siat.ac.cn