PARCEL: Physics-Based Unsupervised Contrastive Representation Learning for Multi-Coil MR Imaging

Shanshan Wang, Ruoyou Wu, Cheng Li, Juan Zou, Ziyao Zhang, Qiegen Liu, Yan Xi, and Hairong Zheng

Abstract—With the successful application of deep learning to magnetic resonance (MR) imaging, parallel imaging techniques based on neural networks have attracted wide attention. However, in the absence of high-quality, fully sampled datasets for training, the performance of these methods is limited. And the interpretability of models is not strong enough. To tackle this issue, this paper proposes a Physics-Based unsupervised Contrastive Representation Learning (PARCEL) method to speed up parallel MR imaging. Specifically, PARCEL has a parallel framework to contrastively learn two branches of model-based unrolling networks from augmented undersampled multi-coil k-space data. A sophisticated co-training loss with three essential components has been designed to guide the two networks in capturing the inherent features and representations for MR images. And the final MR image is reconstructed with the trained contrastive networks. PARCEL was evaluated on two vivo datasets and compared to five state-of-the-art methods. The results show that PARCEL is able to learn essential representations for accurate MR reconstruction without relying on fully sampled datasets. The code will be made available at https://github.com/ternencewu123/PARCEL.

Index Terms—Deep learning, parallel imaging, contrastive representation learning, magnetic resonance imaging (MRI)

1 INTRODUCTION

Parallel imaging is an essential technique for accelerating MR imaging [1], [2], [3]. It utilizes magnetic resonance physics and the sensitivity of multiple coils to reconstruct MR images from multi-coil measurements either directly in the k-space domain [1] or in the spatial domain [2]. For instance, Griswold et al. [1] proposed a partially parallel acquisition method, GRAPPA, to accelerate image acquisition. Pruessmann et al. [2] proposed a sensitivity encoding (SENSE) method, and Sodickson et al. [3] proposed a new fast-imaging technique (SMASH) to increase MR image acquisition speed. These methods have achieved great successes. However, the acceleration factor in classical parallel imaging is limited and its performance suffers from noise amplification effect [4]. To address this issue, compressed sensing along with different sparse prior knowledges have been introduced into parallel MR imaging, which can better remove aliasing artifacts and suppress noise [5], [6], [7]. However, it is difficult to determine the weight of these regularization terms in compressed sensing methods, and its inherent iterative reconstruction process is time-consuming [8], [9], [10].

To further promote MR imaging speed and automate weight parameter settings, deep learning has been introduced to parallel MR reconstruction from undersampled data. These methods can be roughly divided into data-driven and model-based methods [10]. Data-driven methods require a neural network model to learn the mapping between the artifact images and high-quality images, or undersampled k-space data and fully sampled k-space data [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. For
instance, Sriram et al. [17] proposed GrappaNet, which inte-
gerate traditional parallel imaging methods into deep neural
networks to generate high quality reconstructions. Pawar et al. [18] incorporates domain knowledge of parallel MR
imaging to augment the DL networks to achieve accurate
and stable image reconstruction. Feng et al. [20] proposed a
dual-octave (DONet) that can learn multi-scale spatial fre-
quency features of MRI data, further accelerating parallel
MRI reconstruction. Model-based methods based on the
compressed sensing reconstruction algorithm unroll the
iterative optimization process into a deep network, and
use data training to optimize the parameters [21], [22], [23],
[24], [25], [26], [27]. Generally, the model-based methods
have better physical interpretability and is more robust
compared to data-driven methods [10]. For example, Ham-
memirk et al. [8] combined the variational model with deep
learning and embedded compressed sensing reconstruction
into the gradient descent method to achieve the rapid recon-
struction of MR images. In addition, Aggarwal et al. [24]
proposed MoDL, which uses a deep neural network as a
regularization term and the conjugate gradient algorithm
to solve the inverse problem in combination with data consis-
tency. Aggarwal et al. [27] proposed J-MoDL, which intro-
duce a continuous strategy to optimize the sampling
pattern and network parameters jointly. Multi-modal imag-
ing is also an important branch of MR reconstruction [28],
[29]. Feng et al. [28] proposed a multi-stage integration net-
work for super-resolution of multi-modal MR images.

The above methods have enabled great progress in
accelerating MR imaging. However, they rely heavily on
high quality, fully sampled MR images [25]. To decrease
the dependence on full reference data, self-supervised
learning has been introduced for MRI [30], [31], [32], [33],
[34]. Especially, Yaman et al. [30] proposed a self-super-
vised learning method (SSDU) for physics-guided deep
learning reconstruction that divides measured data into
two disjoint subsets, one used for training and the other
used for loss function. This inspired method introduces the
training of neural networks without fully sampled refer-
cence data. However, the performance of the method has
room for improvements due to the under-utilization of the
measurement data. They therefore further split the scanned
undersampled measurements into multiple groups with each group consisting two sets of disjoint k-
space data, for more effective deep learning MR recon-
struction [32]. In addition, Gan et al. [33] proposed a
method named SS-JIRCS, which is a self-supervised
model-based deep learning method for image reconstruc-
tion that is equipped with automated coil sensitivity map
(CSM) calibrations. These methods have made encourag-
ing contributions. Nevertheless, it is still worth investiga-
tion of unsupervised deep learning for parallel MR
imaging. Contrastive representation learning [35] is a
widely known unsupervised co-training framework that
is very effective for obtaining essential and accurate rep-
resentations for target samples.

To induce higher reconstruction accuracy, we propose a
physics-based unsupervised contrastive representation
learning (PARCEL) method, which investigates and integra-
tes the strengths of contrastive representation learning [35],
[36], [37] and model-based deep learning MR reconstruction
models [24]. Specifically, we make the following key con-
tributions in this study:

1) A PARCEL imaging framework is proposed which
introduces unsupervised contrastive representation learn-
ing into parallel MR imaging. It simultaneously learns two
model-based unfolding networks from augmented multi-
coil k-space data and then uses them for the final accurate
MRI reconstruction.

2) A sophisticated co-training loss with three essential
components has been designed to guide the two networks
in capturing the similarity of inherent features and repre-
sentations and eliminating the information by which the
two representations differ. Specifically, it has the under-
sampled calibration loss, reconstructed calibration loss and
contrastive representation loss.

3) We compare PARCEL to five state-of-the-art methods
with different sampling masks. The results show that PAR-
CEL achieves good results in both qualitative and quantita-
tive evaluations, which closely approaches the results
achieved by supervised learning methods. In addition,
PARCEL achieves better reconstructions than existing self-
supervised methods.

The remainder of this paper is organized as follows. Sec-
tion 2 introduces parallel magnetic resonance imaging and
a brief recap of MoDL. Section 3 introduces the proposed
method PARCEL. Section 4 summarizes the experimental
details and results. Section 5 is the discussion part and Sec-
ction 6 concludes the paper.

2 PRELIMINARY

2.1 Compressed Sensing Based Parallel MR
Imaging

In parallel MR imaging, multiple receiver coils are used to
accelerate MR imaging. Specifically, for the ith coil mea-
surement, we have

$$y_i = \Omega F S_i x + \varepsilon_i, \quad i = 1, 2, \ldots, C,$$

where $y_i \in \mathbb{C}^M$ represents the k-space measurement corre-
sponding to the ith coil, $\Omega$ indicates the sampling mask, $\varepsilon_i \in \mathbb{C}^M$ represents measurement noise, $C$ represents the num-
ber of coils to be measured, $F$ is the normalized Fourier
transform, $x \in \mathbb{R}^N$ is the to be reconstructed image, and $S_i$ rep-
resents the sensitivity map of the ith coil. $S_i$ can be esti-
mated using the k-space region corresponding to low fre-
quencies (also known as the autocalibration signal, or ACS),
which is fully sampled. In this experiment, we use the
ESPIRiT [38] algorithm to obtain the sensitivity maps. When
applying compressed sensing to parallel MR imaging, the
minimization formula can be described as follows:

$$x^* = \text{arg min}_{x} \frac{1}{2} \| A x - y \|^2 + \lambda \mathcal{R}(x),$$

where $A = \Omega F S$ represents the encoding matrix with the
diagonal matrix $S$ denoting the stack of all the coil sensitivi-
ties $S = \text{diag}(S_i)$. The first term represents the data consist-
tency term; $\mathcal{R}(x)$ represents the regularization term and $\lambda$ rep-
resents the regularization coefficient.
2.2 A Brief Recap of MoDL

In order to solve (2), the regularization term can be a denoising regularization. A typical example is MoDL proposed in [24], which attempts to solve the following optimization problem:

$$x^* = \arg \min_x \frac{1}{2} \|Ax - y\|^2 + \lambda \|N_w(x)\|^2,$$

where \(\lambda\) represents a trainable regularization parameter and \(N_w(x) = x - D_w(x)\) denotes a learned convolutional neural network (CNN) of noise and alias corresponding to the "denoised" version \(D_w(x)\) of \(x\). With the alternating minimization algorithm, the model was solved as the following iteration process:

$$z^k = D_w(x^k)$$

$$x^{k+1} = (A^H A + \lambda I)^{-1}(A^H y + \lambda z^k)$$

where \(k\) is the iteration number and \(z^k\) is the intermediate denoised version of the image \(x^k\). This iteration process is unrolled into a fully supervised learning process, which has two main modules, namely the residual learning based denoiser module \(z = D_w(x)\) and the data consistency module for updating the image \(x^k\). Here, the data consistency constraint is solved by conjugate gradient method. Fig. 1 shows the specific MoDL architecture. The CNN based denoiser block \(D_w(x)\) is shown in Fig. 1a and the unrolled neural network architecture is shown in Fig. 1b, whose weights are shared at different iterations.

They used fully sampled datasets to train the regularization parameters and the unrolled neural network. The final image is generated as the outputs of the neural network.

3 THE PROPOSED METHOD

3.1 The Overall Framework of PARCEL

We proposed a physics-based unsupervised contrastive representation learning model PARCEL, whose overall framework is shown in Fig. 2. The specific information of co-training loss can be found in Section 3.2. Its training phase has two branches of model-based networks unfolded with the conjugate gradient algorithm for solving the following formula:

$$x^* = \arg \min_x \left\{ \frac{1}{2} \|Ax_j - y\|^2 + \lambda_j \|N_{w_j}(x_j)\|^2 \right\}, j = 1, 2,$$

Fig. 1. The originally developed fully supervised MoDL [24] architecture. (a) shows the CNN based denoiser block \(D_w(x)\). (b) is the unrolled architecture of \(K\) iterations. \(D_w(x)\) share the weights across all the \(K\) iterations.

Fig. 2. The overall framework of PARCEL. (a) Training phase. Reundersample \(y\) to obtain \(y_1\) and \(y_2\) for parallel training. The representations \(x_1\) and \(x_2\) are then obtained through the trained parallel network architecture. The co-training loss is specially designed to constrain the contrastive learning process. \(S\) denotes sensitivity map. (b) Testing phase. Undersampled k-space data are fed into the trained model to obtain the output representation \(x_1\) and \(x_2\). The average of \(x_1\) and \(x_2\) is used as the final reconstruction result.
where $A_j = \mathbf{O}_j FS$. Specifically, the re-undersampled mask $\mathbf{O}_j (j = 1, 2$ is the contrastive representation learning branch number), which needs to meet the following conditions: 1) two parallel networks use different selection masks, 2) the input of the network contains mostly low-frequency signals. During the training phase, two re-undersampled mask $\mathbf{O}_1$ and $\mathbf{O}_2$ were used to perform secondary undersampling on the original undersampled data $y_1$ to obtain $y_1$ and $y_2$. Thus, the input data for the contrastive learning framework is obtained, where the two parallel network has similar architecture to MoDL. Among them, the $D_u(x)$ module in our model is shown in Fig. 3. The parallel network is used to train the data, and the co-training loss function was specially designed to constrain the learning process of the model. During the testing phase, undersampled data are fed into the trained model to obtain representations $x_1$ and $x_2$. Then the average of $x_1$ and $x_2$ is used as the final reconstruction result.

3.2 The Proposed Co-Training Loss

We have designed a sophisticated co-training loss with three essential components to guide the two networks in capturing the inherent features and representations for MR images. Specifically, it has the undersampled calibration loss, reconstructed calibration loss and contrastive representation loss. Compared with self-supervised learning of a single network, contrastive learning of parallel network allows for more rigorous and inherent features to be captured. In this way, the network is expected to avoid learning erroneous information [37]. The mathematical formula of the total co-training loss function is as follows:

$$
\xi_{ct} = \frac{1}{L} \left( \sum_{i=1}^{L} \ell_{uc}(Ax_1^i, y_1^i) + \sum_{i=1}^{L} \ell_{uc}(Ax_2^i, y_1^i) \right) \\
+ \frac{1}{L} \left( \sum_{i=1}^{L} \ell_{rc}(x_1^i) + \sum_{i=1}^{L} \ell_{rc}(x_2^i) \right) \\
+ \frac{1}{L} \sum_{i=1}^{L} \ell_{cl}(z_1^i, z_2^i),
$$

(6)

where $L$ is the total number of training samples, $i$ is the $i$th training sample, $\xi_{ct}$ represent co-training loss; $\ell_{uc}$ represent undersampled calibration loss; $\ell_{rc}$ represent reconstructed calibration loss; $\ell_{cl}$ represent contrastive representation loss; $x_1^i$ and $x_2^i$ represent the output representation of the two networks; $z_1^i$ and $z_2^i$ represent the embedding features with the respect of $x_1^i$ and $x_2^i$, respectively.

1) Undersampled Calibration Loss: The undersampled calibration loss is mainly concerned with the k-space points that have been sampled, which ensures that the reconstruction results of the sampling elements are consistent with the zero-filled k-space data from the measurement. Specifically, the output representation is undersampled by the encoding matrix $A$ and then compared with the original undersampled data. It was used to calculate the difference between the undersampled version of the network prediction and the directly zero-filled one. The specific calculation formula is as follows:

$$
\ell_{uc}(Ax_1^i, y_1^i) = \ell_{mse}(F^{-1}Ax_1^i, F^{-1}y_1^i)
$$

(7)

$$
\ell_{mse}(a, b) = \frac{1}{M} \sum_{i=1}^{M} (a_i - b_i)^2,
$$

(8)

where $F^{-1}$ represents the two-dimensional inverse Fourier transform; $M$ represents size of the input data. This operation is trying to calibrate the consistency between the undersampled portion of the output representation and the undersampled k-space data of the original input.

2) Reconstructed Calibration Loss: The reconstructed calibration loss (obtained by applying an affine projection based on the undersampling mask) is constructed and applied to the output representation $x_1$ and $x_2$ of the parallel networks, which not only ensures that the reconstruction results do not deviate from the measurement results but also improves the signal-to-noise ratio of the reconstructed image [16]. Specifically, the output representation is first transformed into k-space by the encoding matrix $E$ and multiplied with $1 - \mathbf{O}$. Then the undersampled k-space data of the original input is added. Then it is transformed to the image domain using the inverse encoding matrix $E^H$. Finally, the mean squared error is calculated with the output representation $x$. The specific calculation formula is as follows:

$$
\ell_{rc}(x) = \ell_{mse}(x, E^H(Ex(1 - \mathbf{O}) + y)),
$$

(9)

where $x$ represents the output representation of Net1 or Net2, $E = FS$ denotes the encoding matrix, $E^H = SF^{-1}$ denotes inverse encoding matrix, and $I$ denotes a matrix with the same size as $\mathbf{O}$ and all elements are 1.

3) Contrastive Representation Loss: We construct a contrastive representation loss based on contrastive representation learning, as shown in Fig. 2a, data are expanded by generating different inputs through two transformations, and the inputs are encoded into representations $x_1$ and $x_2$. Finally, the similarity of the two output representations is maximized to ensure that the outputs of the upper and lower network are close enough. It is expected to more effectively recover high frequency information. Specifically, it’s to calculate the loss between the embedding features $z_1$ and $z_2$ with the respect to the output representations $x_1$ and $x_2$. The specific calculation formula is as follows:

$$
\ell_{cl}(z_1, z_2) = -\log \frac{\exp(\text{sim}(z_1, z_2))}{\exp(\text{sim}(z_1, z_2)) + \gamma}
$$

(10)

where $z_1 = h_1(x_1)$ and $z_2 = h_2(x_2)$, $h_1$ and $h_2$ are the expander, which consists of a fully connected layer of size
1,024 and a ReLU activation function. \( \text{sim}(z_1, z_2) = \frac{z_1^T z_2}{\|z_1\|_2 \|z_2\|_2} \), \( T \) represent the transpose of the matrix, \( \gamma \) is a regulating parameter used to prevent the \( \ell_2 \) from falling to 0 and \( \text{sim}(z_1, z_2) \) is maximized by minimizing \( \ell_\alpha(z_1, z_2) \) to make the two output representations \( x_1 \) and \( x_2 \) as close as possible. The expander eliminates the information by which the two representations differ.

4 Experiments and Results

4.1 Experiment Setup

4.1.1 Datasets

We use two datasets for evaluating the method with a public knee dataset and an in-house brain dataset. The datasets were selected based on two considerations: 1) The knee and brain MR datasets are two of the most widely used datasets in existing MR reconstruction studies. It can be regarded as a routine selection for most papers [10], [22]; 2) Knee imaging and brain imaging are very important for the diagnosis of relevant diseases, including meniscus tear, bone lesion and dural thickening, etc. Therefore, reconstructing high-quality knee/brain MR images is worth of investigation. The public knee data were obtained from the NYU fastMRI database [39] and approved by the NYU School of Medicine Institutional Review Board. Fully sampled MRI data were acquired on one of three clinical 3 T systems (Siemens Magneton Skyra, Prisma and Biograph mMR) or one clinical 1.5 T system (Siemens Magneton Aera). Data acquisition was achieved with a 15-channel knee coil array and a conventional Cartesian 2D TSE protocol. The dataset includes data from two pulse sequences, yielding coronal proton-density weighting with (PD-FS) and without (PD) fat suppression. As per standard clinical protocol, the sequence parameters were matched as closely as possible between the two systems. The specific sequence parameters used were: echo train length 4, matrix size 320 \( \times \) 320, in-plane resolution 0.5 mm \( \times \) 0.5 mm, slice thickness 3 mm, and no gap between slices. The timing varied from system to system, with a repetition time (TR) of between 2,200 and 3,000 ms, and an echo time (TE) between 27 and 34 mm. The shape of the k-space tensor is slices \( \times \) coils \( \times \) height \( \times \) weight. The in-house brain dataset was obtained with 3D TSE protocol by the United Imaging system, uMR 790. The dataset contains different contrasts such as T1, T2 and PD. All data were cropped to 256 \( \times \) 256. For T1-weighted images, TR = 928 ms, TE = 11 ms, voxel resolution = 0.9 \( \times \) 0.9 \( \times \) 0.9 mm. For T2-weighted images, TR = 2500 ms, TE = 149 ms, voxel resolution = 0.9 \( \times \) 0.9 \( \times \) 0.9. For PD-weighted images, TR = 2000 ms, TE = 13 ms, resolution = 1.0 \( \times \) 1.1 \( \times \) 1.1 mm. Informed consents were obtained from the imaging subject in compliance with the Institutional Review Board policy [22]. The knee dataset contains 245 volumes, and the brain dataset contains 22 volumes. The training, validation, and testing sets are divided randomly, with 6: 2: 2. In the experiment, both one and two-dimensional random undersampled masks were tested; the corresponding masks are shown in Fig. 4.

4.1.2 Implementation Details

We use a deep neural network with 5 layers, each with 64 convolution kernels to achieve \( D_w(x) \), and the size of the convolution kernel is 3 \( \times \) 3, except for the last layer. Each layer contains two continuous convolution operations and a linear activation function, ReLU (rectified linear unit, \( f(x) = \max(0, x) \)). The last layer has only one convolution operation. Among them, we have replaced BN [40] with the convolution operation, as shown in Fig. 3. This is similar to the practice in literature [30]. We extract the output of block \( D_w(x) \) to the data consistency layer. In the experiment, the number of alternate iterations of network unfolding, \( K_s \), is 5 [24]. The input and output of the data consistency layer are complex values. Module \( D_w(x) \) provides input by superimposing the real and imaginary parts in a channel. Among these, coil sensitivity is estimated from the central k-space region of each slice using the ESPRIT [38] algorithm, with all assumptions known in the experiment. During training, we used ADAM [41] optimization, and the momentum was (0.9, 0.999). The initial rate was \( 10^{-4} \), and the learning rate attenuation strategy was adopted [42]. The SSIM metric value of the validation set was taken as the monitoring indicator, and the learning rate was multiplied by 0.3 when the SSIM metric no longer decreased within 10 epoch periods. If the SSIM metric did not change within 50 epochs, the training ends. The total epoch was 200, and the batch size was 4. We train the network under two random masks and different acceleration factors to explore the reconstruction effect of the model under different sampling methods. The model is implemented in Pytorch and the code can be downloaded from this link: https://github.com/termencewu123/PARCEL.

4.1.3 Evaluation Metrics

In the experiment, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [43] were used to quantitatively evaluate the experimental results. The SSIM index is the product of the luminance, contrast and structure measure functions. The corresponding calculation formula is

\[
\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{\mu_x^2 + \mu_y^2 + C_1(\sigma_x^2 + \sigma_y^2 + C_2)},
\]

where \( \mu_x \) is the average of \( x \), \( \mu_y \) is the average of \( y \), \( \sigma_x^2 \) is the variance of \( x \), \( \sigma_y^2 \) is the variance of \( y \), \( \sigma_{xy} \) is the covariance of \( x \) and \( y \), \( C_1 = (k_1L)^2 \) and \( C_2 = (k_2L)^2 \) are constants used to ensure stability, \( L \) is the dynamic range of pixel values, \( k_1 = 0.01 \), \( k_2 = 0.03 \).

4.1.4 Comparison Methods

We compared PARCEL with five methods, SENSE [2], Variational-Net [8], U-Net-256, SSDU [30] and Supervised-MoDL [24] methods under different acceleration rates and sampling patterns. SENSE is a classical parallel imaging method based on coil sensitivity encoding; Variational-Net learns a variational network to accelerate MRI reconstruction, the loss function is the mean-squared error (MSE); U-Net-256 is a classical U-Net model trained in a supervised manner, where the number of channels of the last encoder layer is 256, and the loss function is the mean-squared error; SSDU is a model-based method in which an MoDL model trained in a self-supervised manner as in [30], and the loss function is a normalized \( \ell_1 - \ell_2 \) loss; Supervised-MoDL are...
trained using pairs of measurement data and ground truth images, and the loss function is the mean-squared error. The code of all methods are implemented in Python.

4.2 Evaluation With Different Sampling Masks

4.2.1 One-Dimensional Random Sampling Results

To evaluate the reconstruction performance of PARCEL for the one-dimensional random sampling, the public knee dataset is used in this section, which consists of 15 coils and is complex-valued data.

Fig. 5 demonstrates the reconstruction results of coronal PD images with acceleration rate of 3 using SENSE, Variational-Net, U-Net-256, SSDU, Supervised-MoDL and the proposed unsupervised model PARCEL. The first row shows the reconstructed images. The second row represents local detail graphs of the reconstructed images, and the third row presents error maps using a jet color map (blue: low, red: high error) between the reconstructed and the reference images. The SENSE, Variational-Net, and U-Net-256 approaches suffer from visible residual artifacts and do not recover detailed structures due to over-smoothing problems, as shown in the error maps. SSDU, Supervised-MoDL, and our approach PARCEL remove residual artifacts and recover detailed structure. Compared with SSDU, our method achieves better reconstruction results. This is mainly due to contrastive representation learning’s ability to dig deep information along with the co-training loss constraint on the network. Furthermore, the performance of our method is close to the Supervised-MoDL and may have benefited from the constraint of reconstructed calibration loss on the reconstruction result to avoid deviation from the measured data. The quantitative metrics and error maps shown in Fig. 5 are consistent with these observations.

Similar results can be observed for coronal PD-FS images with acceleration rate of 3, as depicted in Fig. 6. PARCEL and Supervised-MoDL methods achieve similar performance while improving the suppression of the visible residual artifacts in SENSE, Variational-Net and U-Net-256. The quantitative evaluation results as well as the residual artifacts in the error maps also highlight these observations. In addition, compared with coronal PD-FS images, the reconstruction quality of coronal PD images improves overall, except for the Variational-Net method.

In addition to qualitative comparisons of visual quality, quantitative comparisons are also highlighted. Fig. 7 show box plots displaying the median and interquartile range (25th-75th percentile) of the quantitative metrics for PSNR and SSIM with acceleration rate 3, across all test datasets for each knee sequence. For all sequences, PARCEL and Supervised-MoDL achieve similar quantitative performance for both PSNR and SSIM, significantly outperforming SENSE, Variational-Net and U-Net-256. In particularly, for the SSIM metric, the results of the values taken are relatively concentrated, which indicates the better performance of our model. Furthermore, compared to pure neural network methods, model-driven based methods have better reconstruction performances.

4.2.2 Two-Dimensional Random Sampling Results

To further evaluate model reconstruction quality, we continued our experiments using a two-dimensional random
mask for the in-house brain dataset. Figs. 8 and 9 show the quantitative results about PSNR and SSIM at different acceleration rates. Compared with the knee reconstruction results with one-dimensional random sampling masks, similar trends can be obtained for the brain reconstruction with two-dimensional random sampling masks. For example, the reconstruction performance of SSDU, Supervised-MoDL and PARCEL proposed in this paper is better than SENSE, Variational-Net and U-Net-256. In addition, the results generated by PARCEL are close to the fully supervised method Supervised-MoDL. The details are shown in Figs. 7, 8 and 9.

In addition to the quantitative comparison with PSNR and SSIM, the visual quality is also presented. Fig. 10 shows qualitative reconstruction results of the brain testing set using various reconstruction methods at the acceleration rates of 4 and 8 along with their error maps using a jet color map (blue: low, red: high error). These experimental results show that PARCEL can reconstruct the images with improved PSNR and SSIM. The error map clearly shows that SSDU, Supervised-MoDL and PARCEL maintain detailed information better than the other methods. The pure U-Net network fails to achieve good results. However, it is effective in realizing the reconstruction ability of the MRI image domain. Similar to the experimental results with the Knee, SENSE, Variational-Net and U-Net-256 all produced smoothed results in the brain experiment, resulting in blurred local details, as shown in the detail enlargement in Fig. 10. In addition, at higher acceleration rates, our method PARCEL outperforms SSDU in terms of detail recovery and is close to the result of Supervised-MoDL.

4.3 Validation of Contrastive Representation Learning

To evaluate the effects of contrastive representation learning, we conducted experiments about contrastive representation loss with the different random masks. First, experiments are performed on the one-dimensional random sampling pattern, and Fig. 11 show the quantitative comparison results of the contrastive representation loss. Among them, Single-Net is an end-to-end self-supervised MoDL model, and the loss function is the mean-squared error loss. Parallel-Net is a PARCEL model with only the undersampled calibration loss. CL is a PARCEL model with the undersampled calibration and contrastive representation loss.

From the quantitative results, it shows that the PSNR and SSIM metrics are improved after adding the contrastive representative loss. For example, the PSNR metric improved from 37.3140 dB to 39.3004 dB and the SSIM metric improved from 0.9302 to 0.9401, as shown in Fig. 11. In addition, Fig. 12 show the qualitative results. Compared with traditional self-supervised learning methods, such as...
Single-Net, the qualitative result of the reconstructed images has been improved by using a contrastive representative learning approach with parallel networks. Also, this phenomenon can be observed in the error map in Fig. 12.

In addition, we have done corresponding experiments on two-dimensional random sampling pattern. Figs. 13 and 14 show a quantitative comparison of the brain dataset with the acceleration rates of 4 and 8. From the experimental results, the quantitative results were greatly improved by adding the contrastive representation loss. For example, the PSNR metric improves from 38.5802 dB to 43.3204 dB at the acceleration rate of 4 in Fig. 13. In addition to the quantitative results, Fig. 15 show the qualitative results of the contrastive representation loss on the brain dataset. The first column indicates the reference image, and the other columns represent the reconstruction results and the corresponding error maps. Similar conclusions to the previous ones can be drawn from the results of the error maps. It is evident that contrastive representation learning improves the quality of the reconstructed image. Moreover, the detailed information of local regions is improved. Overall, the parallel network Parallel-Net achieves more information.
compensation than traditional self-supervised learning methods such as Single-Net, while the parallel network CL using contrastive representation learning achieves better reconstruction by learning the similarity between the high-dimensional embeddings of the upper and lower encoded representations.

Meanwhile, we also compare the difference between the two outputs of the parallel network before and after the inclusion of the contrastive representation loss, and the results are shown in Table 1. According to the experimental results, after the inclusion of the contrastive representation loss, the difference between the two outputs representation of the parallel network is smaller, indicating that two output representations are closer. The reason is that contrastive representation learning mines the similar internal information of the inputs ($y_1$ and $y_2$ in Fig. 2) by learning the similarity between the high-dimensional embedding features ($z_1$ and $z_2$ in Eq. (10)), thus ensuring that the output representation of the two networks is as similar as possible.

4.4 Ablation Study

To evaluate the different loss function proposed, relevant experiments were carried out. Table 2 show the results of ablation study conducted with different loss functions. Among these, the calculation formula of contrastive representation loss is shown in (10), and the calculation formula of reconstructed calibration loss is shown in (9). The contrastive representation loss was previously discussed. Here, we mainly discuss the reconstructed calibration loss and the combination between reconstructed calibration loss and contrastive representation loss. The ablation experimental results are based on knee and brain datasets. For the reconstructed calibration loss, on the knee dataset, the PSNR value is 39.6380/3.3139, the SSIM value is 0.9520/0.0268. On the brain dataset, the PSNR value is 43.3696/2.6607 and 38.1895/2.1416, the SSIM value is 0.9795/0.0083 and 0.9514/0.0161, when the acceleration rates are 4 and 8, respectively. Similarly, for the combination between reconstructed calibration loss and contrastive representation loss, on the knee dataset, the PSNR value is 39.6476/3.3136, the SSIM value is 0.9521/0.0266. On the brain dataset, the PSNR value is 43.4399/2.6598 and 38.2024/2.1305.
2.2844 0.9919
2.1416 0.9514

and With
0.0024
0.0161
2.6598 0.9797
0.0162
þ
þ
3.3136 0.9521
0.0161,
2.6607 0.9795
0.0012
45.2889
þ
2.2453 0.9791
þ
þ
2.8781 0.9962
0.0083 and 0.9519
44.2330
þ
0.0155
þ
þ
2.6119 0.9371
0.0221
2.3688 0.9938
þ
0.0349
þ
0.0229
45.6516
þ
0.0087
2.5057 0.9941
3.3139 0.9520
0.0083
2.1319 0.9487
2.6267 0.9758
0.0160
0.0015
0.0083
STD)
0.0268
þ
2.5712 0.9735
3.3419 0.9513
2.1187 0.9482
0.0087
0.0226
2.8372 0.9401
45.8389
ISCUSSIONS
0.0135
0.0345
Brain
Knee
Dataset Acceleration rate Method PSNR/dB SSIM

The SSIM value is 0.9797 ± 0.0083 and 0.9519 ± 0.0161, respectively.

From the PSNR and SSIM metrics, the reconstruction quality has improved to a certain extent. The results on the knee and brain datasets show that the combination between reconstructed calibration loss and contrastive representation loss achieve the best result. In general, the reconstruction calibration loss between parallel network can further improve the reconstruction performance of the model. In addition, the ability of contrastive representation learning to extract deep information improve the reconstruction effect.

5 DISCUSSIONS

The experimental results show that the MRI reconstruction model constructed by unsupervised contrastive representation learning outperform the classical parallel imaging algorithm and a newly proposed self-supervised method [30] with improved model stability while gradually approaching the performance of the supervised method. Compared with the pure neural network model, the iteratively unfolded network achieves better results per the comparison between the U-Net-256 and PARCEL model shown in Figs. 5 and 6.

In this study, we used a 5-layer deep neural network to realize the denoising module, as shown in Fig. 3. This part of the neural network has no fixed requirements, but can utilize the classical U-Net. The reconstruction results might be improved by using more complex network structures. In addition, the sharing of network parameters allows the network to perform a greater number of iterations without increasing the number of parameters. Other optimization methods, such as ISTA [44], can be employed to solve the optimization problem in (2). In the future work, we can investigate new optimization methods to solve the problems in (2). In addition, there are many effective solutions for dealing with complex-valued data [19], [20], [22]. Our framework is flexible regarding the networks to be utilized. Therefore, these solutions can be added to further enhance the performance of our proposed method. In addition, in this study, reconstructed multi-coil MR images and corresponding error maps were plotted to qualitatively evaluate the performance. Quantitative evaluations were conducted by calculating objective metrics, including PSNR and SSIM. In our following study, we may consider some subjective indicators provided by radiologists, such as SaMDs [45], to investigate the clinical acceptance of the reconstruction images.

Furthermore, regarding the generalization of the algorithm, the main concern is whether it works well for data that has not been seen before. Compared with existing methods that require supervised data, our unsupervised approach can be more easily adapted to new data distributions in real-word applications, and can be fine-tuned online for undersampled data acquired to ensure the performance of the model. In addition, the algorithm proposed in this paper is better in terms of explainability compared to the pure deep neural network approach. The proposed model is solved by using the conjugate gradient-based optimization algorithm and unrolled into neural networks, and thus there exists mathematical explanations. From the MR imaging point of view, the data fidelity term is a prerequisite to reconstruct faithfully the images, which can better help the radiologists visualize the interested regions and detect possible diseases.

## TABLE 1

Quantitative Comparison Results of Parallel Network Output Without \( \ell_{cl}(o) \) and With \( \ell_{cl}(u) \) About Contrastive Representation Loss

| Dataset | Acceleration rate | Method       | PSNR/dB       | SSIM   |
|---------|------------------|--------------|---------------|--------|
| Knee    | R=3              | \( \ell_{cl}(o) \) | 44.2330±2.7818 | 0.9764±0.0155 |
|         | R=3              | \( \ell_{cl}(u) \) | 45.2889±2.2453 | 0.9791±0.0135 |
|         | R=4              | \( \ell_{cl}(o) \) | 50.4809±2.5057 | 0.9941±0.0015 |
|         | R=4              | \( \ell_{cl}(u) \) | 51.4517±2.8781 | 0.9962±0.0012 |
|         | R=8              | \( \ell_{cl}(o) \) | 45.6516±2.2844 | 0.9919±0.0024 |
|         | R=8              | \( \ell_{cl}(u) \) | 45.8389±2.3688 | 0.9938±0.0024 |

## TABLE 2

Quantitative Comparison of Reconstructed Magnetic Resonance Imaging Using PARCEL With the Different Losses and Two Baseline Models With Two Acceleration Rates (MEAN ± STD)

| Dataset | Acceleration rate | Method       | PSNR/dB       | SSIM   |
|---------|------------------|--------------|---------------|--------|
| Knee    | R=3              | Single-Net   | 37.3140±3.3727 | 0.9302±0.0349 |
|         | R=3              | PARCEL(\( \ell_{uc} \)) | 39.0562±2.6119 | 0.9371±0.0221 |
|         | R=3              | PARCEL(\( \ell_{uc} + \ell_{cl} \)) | 39.3004±2.8372 | 0.9401±0.0226 |
|         | R=3              | PARCEL(\( \ell_{uc} + \ell_{rc} \)) | 39.6380±3.3139 | 0.9520±0.0268 |
|         | R=3              | PARCEL(\( \ell_{uc} + \ell_{cl} + \ell_{rc} \)) | 39.6476±3.3136 | 0.9521±0.0266 |
|         | R=4              | Single-Net   | 38.5802±3.3419 | 0.9513±0.0229 |
|         | R=4              | PARCEL(\( \ell_{uc} \)) | 43.1302±2.5712 | 0.9725±0.0087 |
|         | R=4              | PARCEL(\( \ell_{uc} + \ell_{cl} \)) | 43.3204±2.6267 | 0.9758±0.0087 |
|         | R=4              | PARCEL(\( \ell_{uc} + \ell_{rc} \)) | 43.3696±2.6607 | 0.9795±0.0083 |
|         | R=4              | PARCEL(\( \ell_{uc} + \ell_{cl} + \ell_{rc} \)) | 43.4399±2.6598 | 0.9797±0.0083 |
|         | R=8              | Single-Net   | 33.4977±2.2391 | 0.8908±0.0345 |
|         | R=8              | PARCEL(\( \ell_{uc} \)) | 38.0694±2.1187 | 0.9482±0.0160 |
|         | R=8              | PARCEL(\( \ell_{uc} + \ell_{cl} \)) | 38.0946±2.1319 | 0.9487±0.0162 |
|         | R=8              | PARCEL(\( \ell_{uc} + \ell_{rc} \)) | 38.1895±2.1416 | 0.9514±0.0161 |
|         | R=8              | PARCEL(\( \ell_{uc} + \ell_{cl} + \ell_{rc} \)) | 38.2024±2.1305 | 0.9519±0.0161 |
6 CONCLUSION

In this paper, we propose a physics-based unsupervised contrastive representation learning method to speed up parallel MR imaging. It has a parallel framework to contrastively learn two branches of model-based unrolling networks directly from augmented undersampled multi-coil k-space data. To guide the two networks in capturing the inherent features and representations for MR images, a sophisticated co-training loss with three essential components has been designed. Finally, the final MR image is reconstructed with the trained contrastive networks. PARCEL was evaluated on two vivo datasets and compared to five state-of-the-art methods. The results show that PARCEL can learn useful representations for more accurate MR reconstructions without relying on fully sampled datasets. In the future we will investigate contrastive representation learning for dynamic or multi-contrast MR imaging.

ACKNOWLEDGMENTS

This work was supported in part by Scientific and Technical Innovation 2030—“New Generation Artificial Intelligence” Project under Grants 2020AAA0104100 and 2020AAA0104105, in part by the National Natural Science Foundation of China under Grant 61871371, in part by the Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application under Grant 2022B1212010011, in part by the Basic Research Program of Shenzhen under Grant JCYJ20180507182400762, in part by Shenzhen Science and Technology Program under Grant RCYX20210706092104034, in part by the Youth Innovation Promotion Association Program of Chinese Academy of Sciences under Grant 2019351, and in part by the Key Technology and Equipment R&D Program of Major Science and Technology Infrastructure of Shenzhen under Grants 202100102 and 202100104. Shanshan Wang and Ruoyu Wu are co-first authors.

REFERENCES

[1] M. A. Griswold et al., “Generalized autocalibrating partially parallel acquisitions (GRAPPA),” Magn. Reson. Med., vol. 47, no. 6, pp. 1202–1210, 2002.
[2] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: Sensitivity encoding for fast MRI,” Magn. Reson. Med., vol. 42, no. 5, pp. 952–962, 1999.
[3] D. K. Sodickson and W. J. Manning, “Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays,” Magn. Reson. Med., vol. 38, no. 4, pp. 591–603, 1997.
[4] P. M. Robson, A. K. Grant, A. J. Madhuranthakan, R. Lattanzi, D. K. Sodickson, and C. A. McKenzie, “Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions,” Magn. Reson. Med., vol. 60, no. 4, pp. 893–907, 2008.
[5] S. Wang et al., “Learning joint-sparse codes for calibration-free parallel MR imaging,” IEEE Trans. Med. Imaging, vol. 37, no. 1, pp. 251–261, Jan. 2018.
[6] M. Lustig and J. M. Pauly, “SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space,” Magn. Reson. Med., vol. 64, no. 2, pp. 457–471, 2010.
[7] K. H. Jin, D. Lee, and J. C. Ye, “A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank hankel matrix,” IEEE Trans. Comput. Imag., vol. 2, no. 4, pp. 480–495, Dec. 2016.
[8] K. Hammerink et al., “Learning a variational network for reconstruction of accelerated MRI data,” Magn. Reson. Med., vol. 79, no. 6, pp. 3055–3071, 2018.
[9] Y. Yang, J. Sun, H. Li, and Z. Xu, “ADMM-CSNet: A deep learning approach for image compressive sensing,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 3, pp. 521–538, Mar. 2020.
[10] S. Wang, T. Xiao, Q. Liu, and H. Zheng, “Deep learning for fast MR imaging: A review for learning reconstruction from incomplete k-space data,” Biomed. Signal Process. Control, vol. 68, 2021, Art. no. 102579.
[11] K. Kwon, D. Kim, and H. Park, “A parallel MR imaging method using multilayer perceptron,” Med. Phys., vol. 44, no. 12, pp. 6209–6224, 2017.
[12] Y. Jun, T. Eo, H. Shin, T. Kim, H.-J. Lee, and D. Hwang, “Parallel imaging in time-of-flight magnetic resonance angiography using deep multistream convolutional neural networks,” Magn. Reson. Med., vol. 81, no. 6, pp. 3840–3853, 2019.
[13] M. Akçakaya, S. Möller, S. Weingärtner, and K. Ugurbul, “Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging,” Magn. Reson. Med., vol. 81, no. 1, pp. 439–453, 2019.
[14] S. Wang et al., “Accelerating magnetic resonance imaging via deep learning,” in Proc. IEEE 13th Int. Symp. Biomed. Imag., 2016, pp. 514–517.
[15] D. Lee, J. Yoo, S. Tak, and J. C. Ye, “Deep residual learning for accelerated MRI using magnitude and phase networks,” IEEE Trans. Biomed. Eng., vol. 65, no. 9, pp. 1985–1995, Sep. 2018.
[16] M. Mardani et al., “Deep generative adversarial neural networks for compressive sensing MRI,” IEEE Trans. Med. Imaging, vol. 38, no. 1, pp. 167–179, Jan. 2019.
[17] A. Sriram, J. Zbontar, T. Murrell, C. L. Zitnick, A. Defazio, and D. K. Sodickson, “GrappaNet: Combining parallel imaging with deep learning for multi-coil MRI reconstruction,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020, pp. 14315–14322.
[18] K. Pawar, G. F. Egan, and Z. Chen, “Domain knowledge augmentation of parallel MR image reconstruction using deep learning,” Computerized Med. Imag. Graph., vol. 92, 2021, Art. no. 101968.
[19] C.-M. Feng, Z. Yang, G. Chen, Y. Xu, and L. Shao, “Dual-octave convolution for accelerated parallel MR image reconstruction,” in Proc. AAAI Conf. Artif. Intell., 2021, pp. 116–124.
[20] C.-M. Feng, Z. Yang, H. Fu, Y. Xu, J. Yang, and L. Shao, “DONet: Dual-octave network for fast MRI image reconstruction,” 2021, arXiv:2105.05980.
[21] J. Chiampi, J. Callavero, J. Hainal, A. Price, and D. Rueckert, “A deep cascade of convolutional neural networks for dynamic MRI image reconstruction,” IEEE Trans. Med. Imag., vol. 37, no. 2, pp. 491–503, Feb. 2018.
[22] S. Wang et al., “DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution,” Magn. Reson. Imag., vol. 68, pp. 136–147, 2020.
[23] J. Zhang and B. Ghanem, “STA-Net: Interpretable optimization-inspired deep network for image compressive sensing,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 1828–1837.
[24] H. K. Aggarwal, M. P. Mani, and M. Jacob, “MoDL: Model-based deep learning architecture for inverse problems,” IEEE Trans. Med. Imag., vol. 38, no. 2, pp. 394–405, Feb. 2019.
[25] F. Knoll et al., “Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues,” IEEE Signal Process. Mag., vol. 37, no. 1, pp. 128–140, Jan. 2020.
[26] I. Y. Chun, Z. Huang, H. Lim, and J. Fessler, “Momentum-Net: Fast and convergent iterative neural network for inverse problems,” 2019, arXiv:1907.11818.
[27] H. K. Aggarwal and M. Jacob, “J-MoDL: Joint model-based deep learning for optimized sampling and reconstruction,” IEEE J. Sel. Topics Signal Process., vol. 14, no. 6, pp. 1151–1162, Oct. 2020.
[28] C.-M. Feng, H. Fu, S. Yuan, and Y. Xu, “Multi-contrast MRI super-resolution via a multi-stage integration network,” in Proc. Int. Conf. Med. Imag. Comput. Comput.- Assist. Intervention, 2021, pp. 140–149.
[29] C.-M. Feng et al., “Multi-modal transformer for accelerated MR imaging,” 2021, arXiv:2106.14248.
[30] H. Mardani, S. A. Hoxsini, S. Möller, J. Ellermann, K. Ugurbul, and M. Akçakaya, “Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data,” Magn. Reson. Med., vol. 84, no. 6, pp. 3172–3191, 2020.
[31] F. Liu, R. Kijowski, G. El Fakhri, and L. Feng, “Magnetic resonance parameter mapping using model-guided self-supervised deep learning,” Magn. Reson. Med., vol. 85, no. 6, pp. 3211–3226, 2021.
B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellerrmann, K. Ugurbil, and M. Akakaya, “Ground-truth free multi-mask self-supervised physics-guided deep learning in highly accelerated MRI,” in Proc. IEEE 18th Int. Symp. Biomed. Imag., 2021, pp. 1850–1854.

W. Gan et al., “SS-HRCS: Self-supervised joint image reconstruction and coil sensitivity calibration in parallel MRI without ground truth,” in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2021, pp. 4048–4056.

C. Hu, C. Li, H. Wang, Q. Liu, H. Zheng, and S. Wang, “Self-supervised learning for MRI reconstruction with a parallel network training framework,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention, 2021, pp. 382–391.

T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in Proc. Int. Conf. Mach. Learn., 2020, pp. 1597–1607.

J.-B. Grill et al., “Bootstrap your own latent-a new approach to self-supervised learning,” in Proc. Adv. Neural Inf. Process. Syst., 2020, pp. 21271–21284.

A. Bardes, J. Ponce, and Y. LeCun, “VICReg: Variance-invariance-covariance regularization for self-supervised learning,” 2021, arXiv:2105.04906.

M. Uecker et al., “ESPIRiT–An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA,” Magn. Reson. Med., vol. 71, no. 3, pp. 990–1001, 2014.

J. Zbontar et al., “fastMRI: An open dataset and benchmarks for accelerated MRI,” 2018, arXiv:1811.08839.

S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in Proc. Int. Conf. Mach. Learn., 2015, pp. 448–456.

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014, arXiv:1412.6980.

C. Li, W. Li, C. Liu, H. Zheng, J. Cai, and S. Wang, “Artificial intelligence in multiparametric magnetic resonance imaging: A review,” Med. Phys., vol. 49, no. 10, pp. 1024–1054, 2022.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. Imag. Sci., vol. 2, no. 1, pp. 183–202, 2009.

M. L. Machal, “Impact of Class I. Software as Medical Devices (SaMDs) on Public Health,” in Advances in Informatics, Management and Technology in Healthcare, Amsterdam, Netherlands: IOS Press, 2022, pp. 33–36.

Shanshan Wang (Senior Member, IEEE) received the dual PhD degrees in information technologies and biomedical engineering from the University of Sydney and Shanghai Jiao Tong University. She is a professor with the Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. Her research interests include machine learning, fast medical imaging, and radiomics. She has published more than 100 journal and conference papers in these areas.

Cheng Li received the PhD degree from Nanyang Technological University, Singapore. Currently, she is an assistant professor in Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. Her research interests include medical image processing, signal processing, multi-modal image fusion, and machine learning.

Juan Zou is currently working toward the PhD degree with the School of Physics and Optoelectronics, XianJian University. Her research interests include medical imaging processing, machine learning and deep learning, and federated learning.

Ziyao Zhang is currently working toward the PhD degree with the Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China and the University of Chinese Academy of Sciences, Beijing, China. His research interests include computer vision, three-dimensional imaging, deep neural network, medical image analysis.

Qiegen Liu received the PhD degree in biomedical engineering from Shanghai Jiao Tong University. Since 2012, he has been with the School of Information Engineering, Nanchang University, Nanchang, China. His current research interests include sparse representations, deep learning and their applications in image processing, computer vision and MRI reconstruction.

Yan Xi received the PhD degree in biomedical engineering from Shanghai Jiao Tong University. He is the general manager of Jiang Su First-Imaging Medical Equipment Co., Ltd. His research interests include CT imaging method, machine learning. He has published more than 20 journal papers in these areas.

Hairong Zheng received the PhD degree from the University of Colorado. He is a professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His main research areas are medical imaging technology and systems, information technology and theoretical methods in medical imaging.

Ruoyou Wu is currently working toward the PhD degree with the Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China and the University of Chinese Academy of Sciences, Beijing, China. His research interests include magnetic resonance imaging, deep learning and image processing.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.