Coil-Agnostic Attention-Based Network for Parallel MRI Reconstruction

Jingshuai Liu[1](✉), Chen Qin[2], and Mehrdad Yaghoobi[1]

[1] IDCOM, School of Engineering, University of Edinburgh, Edinburgh, UK
{J.Liu, m.yaghoobi-vaighan}@ed.ac.uk
[2] Department of Electrical and Electronic Engineering, Imperial College London, London, UK

Abstract. Magnetic resonance imaging (MRI) is widely used in clinical diagnosis. However, as a slow imaging modality, the long scan time hinders its development in time-critical applications. The acquisition process can be accelerated by types of under-sampling strategies in k-space and reconstructing images from a few measurements. To reconstruct the image, many parallel imaging methods use the coil sensitivity maps to fold multiple coil images with model-based or deep learning-based estimation methods. However, they can potentially suffer from the inaccuracy of sensitivity estimation. In this work, we propose a novel coil-agnostic attention-based framework for multi-coil MRI reconstruction which completely avoids the sensitivity estimation and performs data consistency (DC) via a sensitivity-agnostic data aggregation consistency block (DACB). Experiments were performed on the FastMRI knee dataset and show that the proposed DACB and attention module-integrated framework outperforms other deep learning-based algorithms in terms of image quality and reconstruction accuracy. Ablation studies also indicate the superiority of DACB over conventional DC methods.

Keywords: MRI reconstruction · Coil agnostic · Attention network.

1 Introduction

Magnetic resonance imaging (MRI) provides a non-invasive imaging tool and can be applied to visualize different types of tissues. However, the acquisition speed of MRI raw data is fundamentally limited due to both hardware and physiological constraints. One common practice to accelerate the process is to under-sample the measurements in k-space. However, the signals sampled below the Nyquist-Shannon rate suffer from aliasing artifacts in image domain. With the assumption that the target images can be expressed using sparse representations in image domain [9] or in some transformed space [13, 20, 28], compressed sensing (CS) methods solve the ill-posed problems via iterative model-based algorithms. Nevertheless, the sparsity prior can be difficult to hold in real-world scenarios and model complicated features [38], which restrains the growth of CS-based methods in modern MRI. The other method to reduce the acquisition

time is parallel imaging (PI), which employs multiple coils to unfold the targets in image domain and exploits the coil correlations to generate clean reconstructions. GRAPPA is introduced in [11] which predicts the missing k-space data with the acquired multi-coil data, e.g. the auto-calibration signal (ACS) lines. It is still challenging to remove strong aliasing artifacts under low sampling rates, using traditional reconstruction methods.

In recent years, deep neural networks show superior performance in image super-resolution, de-noising, compressed sensing, and in-painting [40, 35, 2, 37]. Many methods using neural networks to reconstruct MR images are proposed. The work in [19] uses magnitude and phase networks to perform residual learning of aliasing artifacts and achieve de-aliased outputs. The method in [33] retrieves promising reconstructions by maximizing the posteriori estimated via a pre-trained variational auto-encoder (VAE). The method in [3] predicts the missing k-space data from the sampled points using a neural network. Many PI methods conventionally incorporate the sensitivity maps in the reconstruction pipelines [1, 8], which can be pre-computed via model-based algorithms, e.g. ESPIRIT [34]. However, the reconstruction performance can be affected when the sensitivity estimation is inaccurate with few ACS lines at a high acceleration factor [39, 32]. A different class of methods have been proposed to jointly predict MR images and sensitivity maps. Deep J-Sense [4] unrolls an alternating optimization to jointly refine the reconstructed images and the estimated sensitivity maps. Joint-ICNet [16] merges deep learning networks with model-based methods and updates the sensitivity maps at each stage. Those methods rely on the sampled multi-coil data to estimate the coil sensitivities, which can be susceptible to coil configuration discrepancies and suffer from estimation errors.

In this paper, we introduce a novel multi-coil MRI reconstruction framework which incorporates spatial attention modules and interleaved data aggregation consistency blocks into a multi-level densely connected network structure. The proposed method is completely coil-agnostic when exhibiting data consistency in reconstruction steps, and therefore invulnerable to coil configurations. We show that it outperforms other state-of-the-art methods qualitatively and quantitatively. In ablations, we evaluate the effectiveness of model components and demonstrate the superior performance over other sensitivity estimation-based approaches. The implementation of our method is available at https://github.com/JLiu-Edinburgh/CoilfreeMRI-ACCV2022. The main contributions are summarized below:

- we propose a coil-agnostic attention-based network for parallel MRI reconstruction;
- we develop a novel data aggregation consistency block (DACB) to explicitly utilize the multi-coil k-space consistency constraints without the estimation of sensitivity maps, where data consistency is performed with complex-valued measurements using a coil-invariant mapping to complex domain;
- we introduce a novel locality-aware spatial attention module (LoA-SAM) to achieve both adaptability and locality-aware spatial contexts and incorporate

multi-level dense connections to facilitate feature transmission at multiple levels, which shows to improve the performance;
– we perform extensive experiments on texture-rich FastMRI knee dataset to validate the efficacy of the proposed coil-agnostic method, showing its superior performance against the competing methods.

2 Related Works

2.1 Deep Neural Networks in MRI Reconstruction

Recently, deep neural networks have widely drawn attention and convolutional neural networks (CNN) achieved great success in computer vision tasks [18, 29]. Many methods leverage the representations of neural networks to obtain gratifying achievements in MRI reconstruction. A deep cascade of CNNs was proposed in [30] to recover the aliased MR images. LPD-Net was introduced in [43] which converts the conventional CS problem into two easy sub-problems that are solved via a primal-dual network. The method in [23] adopts a CNN-based iterative framework to remove noise-like artifacts caused by under-sampling. A deep framework was proposed in [22] which fuses network structures with the conventional iterative optimization to provide more accurate solutions to CS-MRI problems. Generative adversarial networks (GAN) [10] produce photo-realistic images via an adversarial game between a generator and a discriminator. The method in [38] adopts a GAN-based framework to achieve sharp and realistic reconstructions. GANCS [25] leverages the interleaved null-space projections and deep residual blocks to remove the aliasing artifacts. The method in [41] uses a self-attention layer in the generator to capture the global information of high-level features. A GAN-based framework was proposed in [21] to ensure rich and natural textures using multi-level feature refinement and attentive selection. The methods in [26, 6] incorporate the GAN prior in MRI reconstruction by optimizing the latent space of a pre-trained generative network.

Although deep neural network models recorded promising achievements in MRI reconstruction, how to produce high-fidelity reconstructions from highly under-sampled measurements still remains open, which can potentially limit the maximum acceleration factor in MRI scanning. We propose a locality-aware spatial attention module, which learns adaptive position-specific kernels to introduce more spatial diversities, and a GAN-based multi-level densely connected framework to boost the reconstruction performance in a fast MRI diet.

2.2 Sensitivity Encoding-Based Reconstruction

Coil sensitivity maps are widely employed for parallel MRI reconstruction. The method in [12] utilizes the pre-computed sensitivity maps to calculate the sensitivity weighted combination. The methods in [1, 44] use the sensitivity maps in the interleaved data consistency layers. The weighted average block proposed in [8] merges the sensitivity maps concurrently with the preceding outputs in each

reconstruction stage. However, the reconstruction performance can be severely affected by the inaccuracy of sensitivity estimation, particularly when very limited k-space points are sampled under high acceleration factors [32, 16]. The work in [32] attempts to overcome this problem by predicting the sensitivity maps via a trainable network. Joint-ICNet [16] jointly updates the reconstructions and sensitivity maps. However, it can be affected by the estimation inaccuracy, e.g. caused by different machine configurations including coil deployment orders.

To avoid such issues, we propose a novel block to exploit generalized data consistency without using sensitivity maps, which yields superior performance over other approaches depending on coil sensitivity estimation.

3 Method

3.1 Problem Formulation

MRI reconstruction is traditionally modelled as follows:

$$\min_x \|A(x) - y\|^2 + \lambda R(x), \tag{1}$$

where A denotes the encoding operation involving the coil projection, the Fourier transform, and under-sampling in k-space, $y = \{y_i\}_i$ is the collection of multi-coil measurements, and $R(x)$ is a regularization. However, the iterative optimization methods can be computationally onerous. We instead leverage a trained neural network to provide an end-to-end solution. As illustrated in Fig. 1, the devised framework takes as input the under-sampled signals and reconstructs the target clean images via a network cascade. The framework utilizes a novel data aggregation consistency block (DACB) in a coil-agnostic manner and is characterized by locality-aware spatial attention (LoA-SAM) and multi-level dense connections.

3.2 Locality-Aware Spatial Attention Module (LoA-SAM)

Attention modules, e.g. the key constituent of transformers [36], are widely used to model the feature correlations and improve the representation capacity. However, the quadratic complexity hinders their implementations to high-resolution features which are closely linked with dense prediction tasks, e.g. image reconstructions. To avoid the substantial computational increase, we proposed a locality-aware spatial attention module (LoA-SAM) to capture the local contextual cues in an adaptive manner, which shows to achieve reconstruction gains. In this section, we introduce the structure of LoA-SAM.

Relative Feature Aggregation Block (RFA) We introduce a relative feature aggregation block (RFA) in this section, which captures the spatially relative information and potentially adds reconstruction performance gains. The input feature volume h of size $C \times H \times W$, where C is the number of channels and

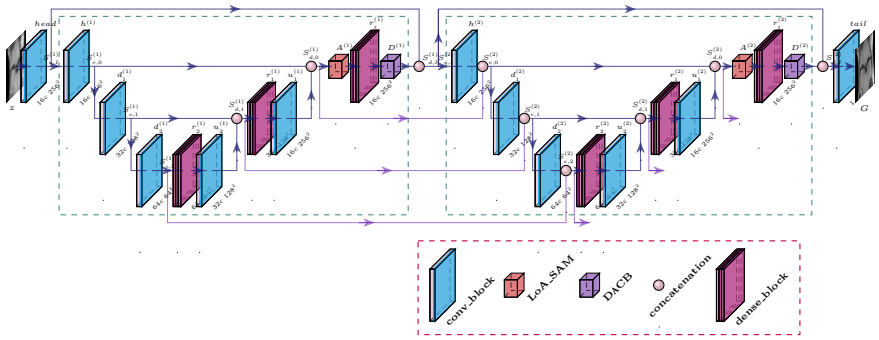

Fig. 1. Overview of the model architecture. The RSS input is first mapped via a head block and fed for a cascade of U-shaped sub-networks with multi-level dense shortcuts. The output features are fused to give the single-channel outcome via a tail block. For simplicity, only two sub-networks with three transitional levels are presented. The final framework has four sub-networks with four encoding-decoding levels

$H \times W$ refers to the spatial resolution, is respectively mapped to h_p and h_q by two branches, as illustrated in Fig. 2 (d). We shift h_p as follows,

$$ps(h_p; \Delta_i)(x) = h_p(x + \Delta_i), \tag{2}$$

where x denotes the spatial position, i.e. $h(x)$ is a vector formed with the elements across the channel axis at position x, and ps refers to the pixel-shift operator which spatially shifts the feature volume by $\Delta_i \in \{-L, .., 0, .., L\}^2$. Zero-padding is adopted for boundary pixels. L is selected to be 2 in our experiments. The relative direction vector $d(x; \Delta_i)$ at position x is given as follows,

$$d(x; \Delta_i) = ps(h_p; \Delta_i)(x) - h_q(x). \tag{3}$$

Each direction vector is combined with the corresponding relative similarity $s(x; \Delta_i)$ which is computed as follows,

$$\begin{aligned} s^*(x; \Delta_i) &= 1 - sim(ps(h_p; \Delta_i)(x), h_q(x)) \\ s(x; \Delta_i) &= 1 - \frac{s^*(x; \Delta_i)}{\tau(\max_j(s^*(x; \Delta_j)) + \epsilon)}, \end{aligned} \tag{4}$$

where sim denotes the Cosine similarity. The hyperparameter τ is practically set to 0.25, and ϵ is 10^{-6}. The feature aggregation $\hat{h}(x)$ is given by:

$$\begin{aligned} g(x; \Delta_i) &= conv([d(x; \Delta_i), s(x; \Delta_i)]) \\ \hat{h}(x) &= conv([g(x; \Delta_i)]_i), \end{aligned} \tag{5}$$

where $conv$ denotes a convolutional layer and [] refers to the concatenation of feature maps along the channel axis.

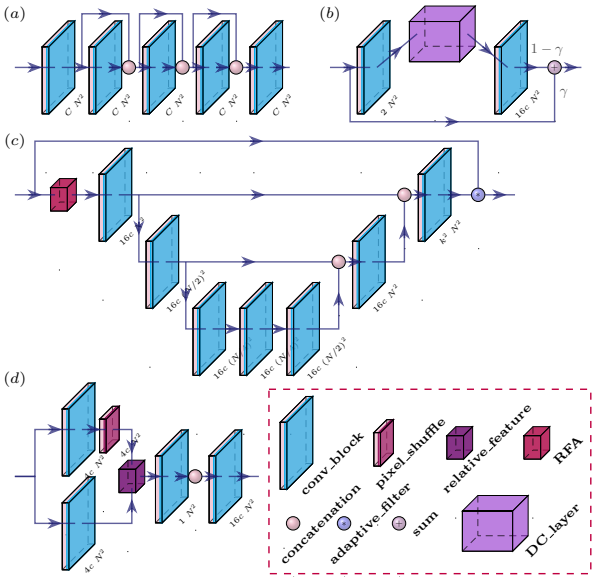

Fig. 2. Illustration of basic modules. C and N^2 indicate the channel and spatial sizes of output features. a) densely connected block, b) DACB, c) LoA-SAM, and d) RFA

Adaptive Spatial Attention Selection The feature aggregation from RFA is passed to a tiny U-net, as shown in Fig. 2 (c), to predict the spatially varying kernels of size $k^2 \times H \times W$. The spatial attention selection is implemented by computing the weighted average of the neighboring positions:

$$\tilde{h}(x) = h(x) + \sum_{\delta \in \Delta} \omega(x, \delta) h(x + \delta), \tag{6}$$

where $\tilde{h}(x)$ is the output of LoA-SAM, Δ is the kernel grid, e.g. $\{-1, 0, 1\}^2$ for $k = 3$ in our experiments, δ is the moving footprint, and $\omega(\delta, x)$ denotes the predicted preference of position $x + \delta$. Different from [46, 5] which predict the kernels of size $Ck^2 \times H \times W$ for adaptive filtering and introduce a substantial computational burden, LoA-SAM captures both the adaptive property and spatial diversity of attention patterns with reducing computational overheads.

3.3 Data Aggregation Consistency Block (DACB)

Data Consistency for MRI Reconstruction We define two data consistency (DC) operators: the soft DC operator Ω^* and the hard DC operator Ω as follows,

$$\begin{aligned} \Omega^*(x; y, m) &= F^{-1}(m \odot \frac{F(x) + \upsilon y}{1 + \upsilon} + (1 - m) \odot F(x)) \\ \Omega(x; y, m) &= F^{-1}(m \odot y + (1 - m) \odot F(x)), \end{aligned} \tag{7}$$

where $x = f(h)$ denotes the prediction given by reducing the channels of h, υ is a parameter, m refers to the sampling mask, F and F^{-1} are the Fourier transform and its inverse, and $\odot$ denotes the Hadamard multiplication. Ω^* is adopted in [1, 15] and [27] updates the prediction by:

$$x \leftarrow x - \eta F^{-1}(m \odot F(x) - y). \tag{8}$$

The two update rules can be converted into a general form as below,

$$x \leftarrow \gamma x + (1-\gamma)\Omega(x; y, m). \tag{9}$$

In parallel imaging scenarios, multi-coil signals can be combined and projected via sensitivity maps as follows,

$$\begin{aligned} s &= \Phi(\{S_i\}; \{C_i\}) = \frac{\sum_i \bar{C}_i S_i}{\sum_j |C_j|^2} \\ S_i &= \Psi(s; C_i) = C_i s, \end{aligned} \tag{10}$$

where S_i denotes the i-th coil image, and C_i and $\bar{C}_i$ refer to the i-th sensitivity map and its conjugate. The output x is projected into the coil sensitivity via (10) and updated, e.g. use (9). The multiple images are then combined into a single view. The data consistency in (9) is eventually extended to be:

$$x \leftarrow \gamma x + (1-\gamma)\Phi(\{\Omega(\Psi(x; C_i); y_i, m)\}; \{C_i\}), \tag{11}$$

where y_i denotes the measurement acquired by the i-th coil.

Data Consistency Block Data consistency is widely used for multi-coil MRI reconstructions. The sensitivity maps were either estimated by model-based algorithms [1, 8, 12], or via trainable neural networks [32, 16]. The former method requires extra computational cost and the reconstruction quality can be severely affected when the sampled ACS lines are not enough to guarantee the estimation accuracy. The latter approach transfers the estimation into end-to-end reconstruction pipelines, whereas it introduces extra parameter overheads and can be sensitive to machine configurations, e.g. coil number and deployment order, imposing restrictions on the generalization ability of reconstruction systems.

The focal point of consistency enforcement is to "correct" the intermediate reconstructions with the observed measurements. Instead of using sensitivity maps in data consistency operations, e.g. (11), we propose a simple yet effective block, named data aggregation consistency block (DACB), to tactfully perform approximated data consistency without sensitivity maps. Similar to the update rules in (9) and (11), we map the feature volume h into a 2-channel output, as illustrated in Fig. 2 (b), representing the complex-valued signal aggregation, and correct it using the summation of the measurements. Different from individually enforcing data consistency on each coil view and combining the outputs together using sensitivity maps, which is the conventional deep multi-coil MRI

reconstruction method, we propose to leverage the linearity of the Fourier transform, i.e. the aggregation of the transforms equals the transformed aggregation, and adopt the summation operator instead of coil combination. To eliminate the impact of the bottleneck design, a shortcut is adopted at the feature level to take the advantage of residual learning. The data consistency is executed by,

$$h \leftarrow \gamma h + (1-\gamma) f^*(\Omega(f(h); \sum y_i, m)), \tag{12}$$

where f and f^* are two convolutional layers used to fuse feature channels. DACB incorporates no sensitivity maps and can avoid defects caused by the inaccurate sensitivity estimation. This method is inherently robust to coil configurations due to the permutation-invariance property of the linear aggregation operation.

3.4 Framework Design

Multi-Level Dense Connections We proposed a multi-level densely connected architecture which “reuses" the learned features at multiple levels via intra- and inter-connections. The illustrative diagram is displayed in Fig. 1 where two U-shaped sub-networks with three transitional levels are presented for simplicity. In order to enable information propagation between sub-networks and consecutively reuse the features, all preceding feature volumes at the same level are collected via skip-connections and combined with the current features, which forms the multi-level densely connected structure.

Network Architecture In our experiments, each sub-network, shown in Fig. 1, has four transitional layers where the bottom level has $64c$ feature channels with $c = 2$. Densely connected layers [14], illustrated in Fig. 2 (a), are adopted at decoding levels. The LoA-SAM and DACB blocks are embedded at the top decoding levels. Two convolutional layers are used as the head and tail blocks to expand and reduce feature channels. Four sub-networks are used in our pipeline. The input zero-filled z is computed using Root-Sum-of-Squares (RSS):

$$z = \sqrt{\sum |F^{-1}(y_i)|^2}. \tag{13}$$

The output reconstruction G is a single-channel image.

3.5 Objective Function

The L_1 loss and structural similarity index (SSIM) [45] are used to measure the reconstruction errors. Moreover, we adopt an adversarial loss L_{adv} [24] to encourage rich and sharp details. Let G and s be the reconstruction and the RSS map of the fully sampled reference. The total loss is given by:

$$L = E_{\{(G,s)\}}[\lambda_{rec}((1-\alpha)L_1(G,s) + \alpha L^{SSIM}(G,s)) + \lambda_{adv} L_{adv}], \tag{14}$$

where $\alpha = 0.64$, $\lambda_{rec} = 10$ and $\lambda_{adv} = 0.05$. We found that the least squares GAN [24] provides stable training. The discriminator in CycleGAN [47] is adopted.

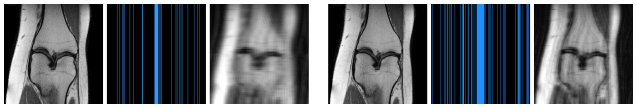

Fig. 3. First) fully sampled, second) mask, last) zero-filled at 8× (*left*) and 4× (*right*)

4 Experiments

4.1 Multi-Coil Reconstruction Comparisons

Implementations. We use the FastMRI multi-coil knee database [42] to conduct experiments, which contains rich textural and structural details. k-space raw data are acquired via 15 coils and converted to image domain via the inverse Fourier transform. Fixed random masks are used in under-sampling as shown in Fig. 3, where the reduction factors are respectively 8 and 4. The model was trained for 35 epochs with a batch size of 2, using an Adam optimizer [17] with β_1=0.5, β_2=0.999, and a learning rate of 10^{-5}, in PyTorch on a NVIDIA GTX 1080Ti. We compare our method with other deep learning approaches: VS-Net [8], ISTA-Net+ [44], DeepCascade [30, 16], DC-CNN [31], FastMRI U-net [42, 16], and Joint-ICNet [16]. The sensitivity maps are estimated via ESPIRIT [34] and used in VS-Net, ISTA-Net+, and DeepCascade. DC-CNN simultaneously reconstructs multiple coil views, FastMRI U-net uses image magnitude maps, and Joint-ICNet estimates coil sensitivities in the reconstruction pipeline. DeepCascade and U-net were also used in [16] as competing methods. The evaluation metrics are PSNR and SSIM, where higher values are better, and FID and KID [7], where lower values are preferred. More qualitative results, including 4× accelerated reconstructions, are shown in Supplementary Material.

Results. The reconstructions using different methods are displayed in Fig. 4, where it shows that the proposed reconstruction framework produces superior results to other comparison methods. It generates more fine-grained textures and better preserves anatomic structures, which leads to more faithful and visually appealing reconstructions. From the last sample in Fig. 4, we observed that the reconstructions using VS-Net, ISTA-Net+, and DeepCascade suffer from artifacts, see the zoomed views. It is caused by the irregular artifacts in the estimated sensitivity maps, see Section 4.3 for more details. The quantitative comparisons presented in Table 1 show that our method consistently outperforms other competing approaches in terms of all evaluation metrics (p-value$\ll$0.05). Our model has fewer parameters than Joint-ICNet with faster inference speed, which potentially enables real-time reconstruction. Ablation studies are conducted to demonstrate that the superior performance is attributable to the proposed model components and structure, and not simply owing to the size of model.

4.2 Ablation Studies on Model Components

Framework Configurations. To perform ablation studies on model components: multi-level dense connections, LoA-SAM, and DACB, we implement the

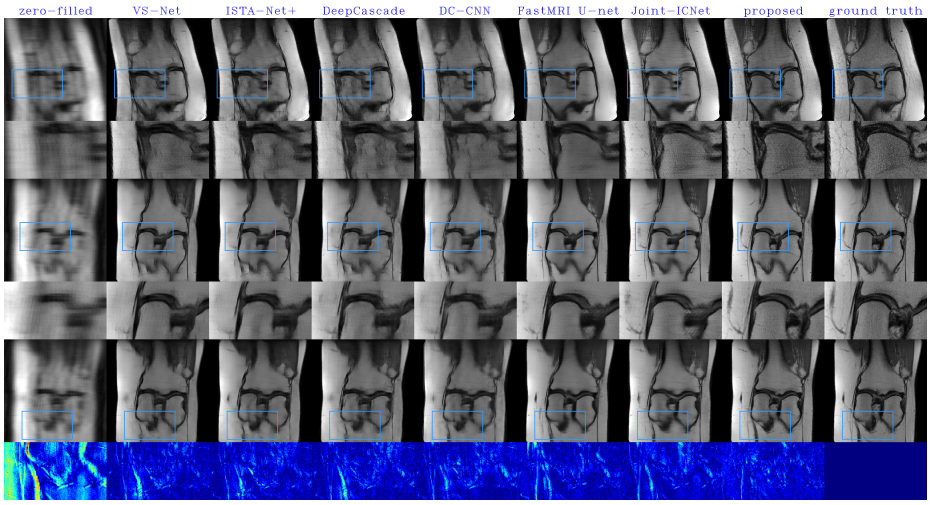

Fig. 4. Comparison results of 8× accelerated MRI reconstruction.

Table 1. Quantitative results of accelerated multi-coil knee MRI reconstruction. 4× accelerated reconstructions have the same runtime (s) and model size (MB) as 8×

	method	PSNR↑	SSIM↑	FID↓	KID↓	runtime↓	size↓
8×	zero-filled	28.75	0.817	379.58	0.459	-	-
	VS-Net [8]	36.18	0.912	141.15	0.078	0.034	4.32
	ISTA-Net+ [44]	36.32	0.911	170.11	0.108	0.189	1.47
	DeepCascade [30, 16]	35.82	0.909	146.17	0.081	0.083	1.73
	DC-CNN [31]	34.96	0.892	180.47	0.124	0.035	**1.11**
	FastMRI U-net [42, 16]	36.19	0.922	145.58	0.095	**0.011**	10.59
	Joint-ICNet [16]	35.59	0.921	137.57	0.090	0.283	263.43
	proposed	**37.74**	**0.930**	**76.08**	**0.011**	0.112	159.11
4×	zero-filled	33.16	0.883	201.20	0.173	-	-
	VS-Net [8]	40.56	0.956	111.34	0.052	-	-
	ISTA-Net+ [44]	40.81	0.956	112.45	0.052	-	-
	DeepCascade [30, 16]	40.57	0.956	109.46	0.050	-	-
	DC-CNN [31]	39.32	0.945	111.32	0.049	-	-
	FastMRI U-net [42, 16]	39.37	0.951	108.52	0.055	-	-
	Joint-ICNet [16]	39.27	0.955	88.69	0.035	-	-
	proposed	**41.10**	**0.963**	**64.37**	**0.009**	-	-

final pipeline, represented in (E), and its different variants in (A)-(D) and (F). For fair comparisons, feature volumes of model (A) and (F) are repeated to approximate the parameter overheads of their counterparts with multi-level dense connections. For the same reason, the convolutional layers in DACB are maintained for (A)-(C). The U-net in LoA-SAM is only used for kernel prediction with limited parameters, while the reconstruction branch with a vast majority

Table 2. Ablations on model components at 8× acceleration

method	ML-dense	LoA-SAM	DACB	PSNR↑	SSIM↑	FID↓	KID↓
(A)	×	×	×	36.53	0.921	89.83	0.025
(B)	√	×	×	37.06	0.926	85.39	0.024
(C)	√	√	×	37.33	0.926	84.16	0.022
(D)	√	×	√	37.53	0.929	80.14	0.017
(E) ours	√	√	√	**37.74**	**0.930**	**76.08**	**0.011**
(F)	×	√	√	37.11	0.927	82.13	0.018

Table 3. Ablations on data consistency blocks and data type at 8× acceleration. The proposed method performs no sensitivity estimation and is therefore free of ME-Arti

method	input	DC	output	PSNR↑	SSIM↑	FID↓	KID↓	ME-Arti
(E)-CSM	complex	CSDC	complex	36.58	0.916	94.56	0.029	strong
(E)-CSM-a	complex	CSDC	single	36.83	0.920	87.46	0.022	weak
(E)-CSM-b	RSS	CSDC	single	36.99	0.925	83.95	0.020	trivial
(E) proposed	RSS	DACB	single	**37.74**	**0.930**	**76.08**	**0.011**	free

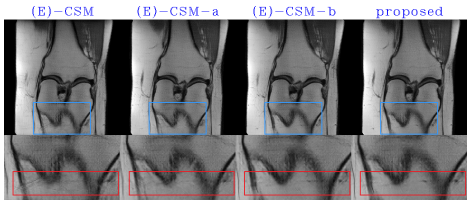

Fig. 5. Ablation results on data consistency blocks and data type at 8× acceleration.

of parameters remains unchanged for a fair comparison. The model configurations and final performance are shown in Table 2.

Results. The results in Table 2 demonstrate that the proposed model components consistently improve the reconstruction performance in terms of evaluation metrics. Comparing model (B) and (E) with model (A) and (F), we found that the multi-level dense connections improve the reconstructions in all metrics. It can be observed that compared with model (B), both LoA-SAM and DACB which are respectively used in model (C) and (D) deliver performance gains in all evaluations. The complete model (E) consistently gives the best results, which indicates the efficacy of the proposed method and implies a positive synergy effect between LoA-SAM and DACB in improving reconstruction quality.

4.3 Ablation Studies on Data Consistency Methods

Implementations. We compare the proposed method with three variants to elucidate how the proposed DACB and RSS-based pipeline benefit the reconstruction performance. The first variant, dubbed (E)-CSM, has 2-channel inputs

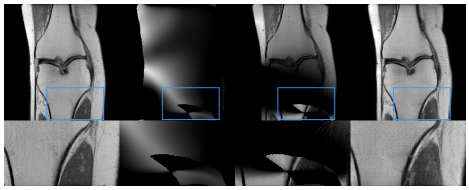

Fig. 6. Map estimation artifacts. First) RSS of fully sampled signal, second) positive component of the imaginary part of a sensitivity map, third) positive component of the imaginary part of the combined image, and last) reconstruction using (E)-CSM

Table 4. Ablation analysis on residual feature connection in DACB at 8× acceleration

method	DC	PSNR↑	SSIM↑	FID↓	KID↓
(E) proposed	DACB	**37.74**	**0.930**	**76.08**	**0.011**
(E)-ResFree	ResFree	37.40	0.928	79.37	0.014

and outputs representing complex values. The pre-computed sensitivity maps are used to combine multi-coil frames via (10). Data consistency is implemented using (11), which we refer to as CSDC. It is extended to be (E)-CSM-a and (E)-CSM-b by progressively adopting the single-channel magnitudes as output and input. The evaluation and visual results are presented in Table 3 and Fig. 5.

Results. From Table 3, we observed that the data type has considerable effects on the reconstruction results and DACB further contributes a performance gain. Comparing to (E)-CSM-b which uses the conventional sensitivity-based DC approach, the proposed method shows superior results, substantiating the efficacy of DACB. In Fig. 5, we found that (E)-CSM suffers from strong artifacts which can be alleviated by adopting (E)-CSM-a and -b. The proposed model avoids the sensitivity estimation and yields clean reconstructions. We postulate that the artifacts emanate from the sensitivity maps, when they are not appropriately estimated and the irregularities attend to the complex-valued inputs. Such artifacts can be propagated to the final reconstructions when the model fails to overcome them, which is demonstrated in the next subsection.

Map Estimation Artifacts (ME-Arti). We demonstrate how the artifacts in Fig. 5 are caused. The RSS map in Fig. 6 shows no artifacts. We found that such irregular shapes stealthily attend to separate image channels, representing complex-valued signals, and can be manifested by solely visualizing the values of the same sign, e.g. negative components. We display a sensitivity map and an image combined via (10) in Fig. 6, where the irregular shapes coincide with the artifacts in the reconstruction using (E)-CSM. It shows the impact of inaccurate sensitivity estimation and implies the usefulness of the proposed method.

Table 5. Ablation analysis on relative feature aggregation at 8× acceleration

method	RFA	PSNR↑	SSIM↑	FID↓	KID↓
(E) proposed	√	**37.74**	**0.930**	**76.08**	**0.011**
(E)-no-RFA	×	37.69	0.929	78.43	0.013

4.4 Ablation Analysis on Data Consistency Bottleneck Design

In Section 3.3, we postulate the negative impact of the bottleneck structure in conventional DC operations, i.e. channel reduction and expansion. To verify this, we remove the skip-connection in DACB and implement the update rule below,

$$h \leftarrow f^*(\gamma f(h) + (1-\gamma)\Omega(f(h); \sum y_i, m)). \tag{15}$$

We denote by ResFree the update in (15), refer to the resultant model as (E)-ResFree, and present the comparisons in Table 4. We found that ResFree reduces PSNR and SSIM with a concomitant increase in FID and KID, which indicates the impact of the bottleneck structure in normal DC operations. It is shown that DACB is able to alleviate this issue and improve the model performance.

4.5 Ablation Analysis on Attention Feature

In this section, we testify the effectiveness of the relative feature aggregation (RFA) block in LoA-SAM. We implement a variant, dubbed (E)-no-RFA, which removes RFA and instead predicts the adaptive kernels directly from the input features. The comparison results are shown in Table 5, where the removal of RFA decreases PSNR and SSIM scores by a small margin and introduces moderately higher FID and KID. It shows that the relative information provided by RFA is preferable over the original features and enhances the capacity of LoA-SAM. Note that RFA uses simple mathematical operations. We leave it a future work to explore the potential benefits of more complex aggregation methods.

4.6 Robustness to Coil Configurations

As an alternative to model-based sensitivity estimation algorithms, neural networks can be used to predict the sensitivity maps, e.g. Joint-ICNet [16]. A different method is introduced in DC-CNN [31] which respectively performs data consistency for parallel coil images. However, such methods, as substantiated in the following, can be sensitive to machine configurations, which restricts models from generalizing to real-world scenarios. We randomly permute the coil orders in inference to demonstrate the robustness of the proposed pipeline to coil configurations. We compare it with Joint-ICNet and DC-CNN which take multi-coil images as input respectively to the sensitivity estimation network and the reconstruction model. Coil frame sequences are rolled by randomly sampled shift numbers. The FastMRI database provides limited diversities of coil orders.

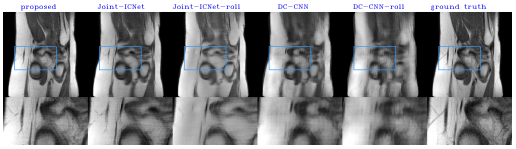

Fig. 7. Comparisons of 8× accelerated MRI reconstruction using coil permutation.

Table 6. Performance results of 8× accelerated multi-coil knee MRI reconstruction using random coil permutation

method	permutation	PSNR↑	SSIM↑	FID↓	KID↓
proposed	None/rolled	**37.74**	**0.930**	**76.08**	**0.011**
Joint-ICNet [16]	None	35.59	0.921	137.57	0.090
Joint-ICNet [16]	rolled	33.18	0.905	143.43	0.092
DC-CNN [31]	None	34.96	0.892	180.47	0.124
DC-CNN [31]	rolled	31.86	0.848	237.91	0.216

Nevertheless, the great majority follow the same configuration. The results are presented in Fig. 7 and Table 6. It is evident that the performance of both Joint-ICNet and DC-CNN can be severely affected by coil permutation, which shows their sensitivity to machine configuration discrepancies. The fact that the number of coils is required to be kept fixed is another problem which can violate the deployment of reconstruction models from training to real-world scanners. Leveraging the permutation-invariance of the RSS and measurement summation operators, the proposed method circumvents those issues and shows a strong invulnerability and generalization capacity.

5 Conclusion

In this study, we introduce a framework to circumvent the coil sensitivity estimation retrospectively required for multi-coil MRI reconstruction frameworks. We propose a coil-agnostic data aggregation consistency block to perform approximated data consistency and a spatial attention module to enhance the model performance. In experiments, it was demonstrated that the proposed framework outperforms other deep learning-based methods qualitatively and quantitatively. In ablation studies, the proposed coil-agnostic pipeline showed superior performance over conventional approaches in accelerated MRI acquisition settings. The future researches include applying our method to other anatomical structures and extending it to dynamic MRI reconstruction.

References

1. Aggarwal, H., Mani, M., Jacob, M.: MoDL: model-based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging **38**(2), 394–405 (2019). https://doi.org/10.1109/TMI.2018.2865356

2. Ahmad, R., Bouman, C., Buzzard, T., Chan, S., Liu, S., Reehorst, E., Schniter, P.: Plug-and-play methods for magnetic resonance imaging: using denoisers for image recovery. IEEE Signal Processing Magazine **37**(1), 105–116 (Jan 2020). https://doi.org/10.1109/MSP.2019.2949470
3. Anuroop, S., Jure, Z., Tullie, M., Lawrence, Z., Aaron, D., K.S., D.: GrappaNet: combining parallel imaging with deep learning for multi-coil MRI reconstruction. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 14303–14310 (2020). https://doi.org/10.1109/CVPR42600.2020.01432
4. Arvinte, M., Vishwanath, S., Tewfik, A., Tamir, J.: Deep J-Sense: accelerated MRI reconstruction via unrolled alternating optimization (2021)
5. Bako, S., Vogels, T., McWilliams, B., Meyer, M., Novák, J., Harvill, A., Sen, P., DeRose, T., Rousselle, F.: Kernel-predicting convolutional networks for denoising Monte Carlo renderings. ACM Transactions on Graphics (TOG) **36**(4) (July 2017)
6. Bhadra, S., Zhou, W., Anastasio, M.: Medical image reconstruction with image-adaptive priors learned by use of generative adversarial networks. Medical Imaging 2020: Physics of Medical Imaging **11312**, 206 – 213 (2020). https://doi.org/10.1117/12.2549750
7. Bińkowski, M., Sutherland, D., Arbel, M., Gretton, A.: Demystifying MMD GANs. International Conference on Learning Representations (2018)
8. Duan, J., Schlemper, J., Qin, C., Ouyang, C., Bai, W., Biffi, C., Bello, G., Statton, B., Regan, D., Rueckert, D.: VS-Net: variable splitting network for accelerated parallel MRI reconstruction. Springer International Publishing pp. 713–722 (2019)
9. Fair, M., Gatehouse, P., DiBella, E., Firmin, D.: A review of 3D first-pass, whole-heart, myocardial perfusion cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance **17** (2015)
10. Goodfellow, I., Pouget, A., Mirza, M., Xu, B., Warde, F., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Advances in Neural Information Processing Systems **27**, 2672–2680 (2014)
11. Griswold, M., Jakob, P., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic resonance in medicine **47**(6), 1202—1210 (June 2002). https://doi.org/10.1002/mrm.10171
12. Hammernik, K., Klatzer, T., Kobler, E., Recht, M., Sodickson, D., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine **79**(6), 3055–3071 (2017)
13. Hong, M., Yu, Y., Wang, H., Liu, F., Crozier, S.: Compressed sensing MRI with singular value decomposition-based sparsity basis. Physics in Medicine and Biology **56**(19), 6311–6325 (Sep 2021)
14. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 2261–2269 (2017)
15. Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., Metaxas, D.: MRI reconstruction via cascaded channel-wise attention network. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) pp. 1622–1626 (2019). https://doi.org/10.1109/ISBI.2019.8759423
16. Jun, Y., Shin, H., Eo, T., Hwang, D.: Joint deep model-based MR image and coil sensitivity reconstruction network (Joint-ICNet) for fast MRI. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 5270–5279 (June 2021)

17. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. 3rd International Conference on Learning Representations (ICLR) (May 2015), http://arxiv.org/abs/1412.6980
18. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems **25** (2012), https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
19. Lee, D., Yoo, J., Tak, S., Ye, J.: Deep residual learning for accelerated MRI using magnitude and phase networks. IEEE Trans. Biomed. Eng **65**(9), 1985–1995 (2018)
20. Lingala, S., Jacob, M.: Blind compressive sensing dynamic MRI. IEEE Transactions on Medical Imaging **32**(6), 1132–1145 (2013)
21. Liu, J., Yaghoobi, M.: Fine-grained MRI reconstruction using attentive selection generative adversarial networks. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) pp. 1155–1159 (2021)
22. Liu, R., Zhang, Y., Cheng, S., Luo, Z., Fan, X.: A deep framework assembling principled modules for CS-MRI: unrolling perspective, convergence behaviors, and practical modeling. IEEE Transactions on Medical Imaging **39**(12), 4150–4163 (2020). https://doi.org/10.1109/TMI.2020.3014193
23. Liu, Y., Liu, Q., Zhang, M., Yang, Q., Wang, S., Liang, D.: IFR-Net: Iterative feature refinement network for compressed sensing MRI. IEEE Trans. Comput. Imag **6** (2020)
24. Mao, X., Li, Q., Xie, H., Lau, R., Wang, Z., Smolley, S.: Least squares generative adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV) pp. 2813–2821 (2017)
25. Mardani, M., Gong, E., Cheng, J.Y., Vasanawala, S.S., Zaharchuk, G., Xing, L., Pauly, J.M.: Deep generative adversarial neural networks for compressive sensing MRI. IEEE Transactions on Medical Imaging **38**(1), 167–179 (2019). https://doi.org/10.1109/TMI.2018.2858752
26. Narnhofer, D., Hammernik, K., Knoll, F., Pock, T.: Inverse GANs for accelerated MRI reconstruction. Wavelets and Sparsity XVIII **11138**, 111381A (Sep 2019). https://doi.org/10.1117/12.2527753
27. Pezzotti, N., Yousefi, S., Elmahdy, M., van Gemert, J., Schülke, C., Doneva, M., Nielsen, T., Kastryulin, S., Lelieveldt, B., van Osch, M., de Weerdt, E., Staring, M.: An adaptive intelligence algorithm for undersampled knee MRI reconstruction. IEEE Access **8**, 204825–204838 (2020)
28. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Transactions on Medical Imaging **30**(5), 1028–1041 (2011)
29. Sajjadi, M., Schölkopf, B., Hirsch, M.: EnhanceNet: single image super-resolution through automated texture synthesis. Computer Vision (ICCV), 2017 IEEE International Conference on pp. 4501–4510 (2017), https://arxiv.org/abs/1612.07919/
30. Schlemper, J., Caballero, J., Hajnal, J., Price, A., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Transactions on Medical Imaging **37**(2), 491–503 (2018). https://doi.org/10.1109/TMI.2017.2760978
31. Schlemper, J., Duan, J., Ouyang, C., Qin, C., Caballero, J., Hajnal, J., Rueckert, D.: Data consistency networks for (calibration-less) accelerated parallel MR image reconstruction. Abstract from ISMRM 27th Annual Meeting and Exhibition (2019)

32. Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C., Yakubova, N., Knoll, F., Johnson, P.: End-to-end variational networks for accelerated MRI reconstruction. International Conference on Medical Image Computing and Computer-Assisted Intervention pp. 64–73 (2020)
33. Tezcan, K., Baumgartner, C., Luechinger, R., Pruessmann, K., Konukoglu, E.: MR image reconstruction using deep density priors. IEEE Transactions on Medical Imaging **38**(7), 1633–1642 (2019). https://doi.org/10.1109/TMI.2018.2887072
34. Uecker, M., Lai, P., Murphy, M., Virtue, P., Elad, M., Pauly, J., Vasanawala, S., Lustig, M.: ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. Magnetic Resonance in Medicine **71**(3), 990–1001 (2014)
35. Valsesia, D., Fracastoro, G., Magli, E.: Deep graph-convolutional image denoising. IEEE Transactions on Image Processing **29**, 8226–8237 (2020)
36. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I.: Attention is all you need. Advances in Neural Information Processing Systems **30** (2017), https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
37. Wang, Y., Tao, X., Qi, X., Shen, X., Jia, J.: Image inpainting via generative multi-column convolutional neural networks. Advances in Neural Information Processing Systems pp. 331–340 (2018)
38. Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., Firmin, D.: DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Transactions on Medical Imaging **37**(6), 1310–1321 (2018)
39. Ying, L., Sheng, J.: Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). Magnetic Resonance in Medicine **57(6)**, 1196–1202 (2007)
40. Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Loy, C., Qiao, Y., Tang, X.: ESRGAN: enhanced super-resolution generative adversarial networks. The European Conference on Computer Vision Workshops (ECCVW) (September 2018)
41. Yuan, Z., Jiang, M., Wang, Y., Wei, B., Li, Y., Wang, P., Menpes-Smith, W., Niu, Z., Yang, G.: SARA-GAN: self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing MRI reconstruction. Frontiers in Neuroinformatics **14**, 1–12 (2020). https://doi.org/10.3389/fninf.2020.611666
42. Zbontar, J., Knoll, F., Sriram, A., Muckley, M., Bruno, M., Defazio, A., Parente, M., Geras, K., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdzal, M., Romero, A., Rabbat, M., Vincent, P., Pinkerton, J., Wang, D., Yakubova, N., Owens, E., Zitnick, C., Recht, M., Sodickson, D., Lui, Y.: FastMRI: an open dataset and benchmarks for accelerated MRI. CoRR **abs/1811.08839** (2018)
43. Zhang, C., Liu, Y., Shang, F., Li, Y., Liu, H.: A novel learned primal-dual network for image cmpressive sensing. IEEE Access **9**, 26041–26050 (2021). https://doi.org/10.1109/ACCESS.2021.3057621
44. Zhang, J., Ghanem, B.: ISTA-Net: iterative shrinkage-thresholding algorithm inspired deep network for image compressive sensing. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 1828–1837 (2018)
45. Zhao, H., Gallo, O., Frosio, I., Kautz, J.: Loss functions for image restoration wth neural networks. IEEE Transactions on Computational Imaging **3**(1), 47–57 (2017)
46. Zhou, S., Zhang, J., Pan, J., Xie, H., Zuo, W., Ren, J.: Spatio-Temporal Filter Adaptive Network for Video Deblurring. Proceedings of the IEEE International Conference on Computer Vision (2019)

47. Zhu, J., Park, T., Isola, P., Efros, A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV) pp. 2242–2251 (2017). https://doi.org/10.1109/ICCV.2017.244