CDNet: Complex-Valued Deep Unrolled Composite Denoising Network for Accelerated Magnetic Resonance Image Reconstruction

David Muigai[1], Elijah Mwangi[2], Henry Kiragu[3]
Department of Electrical Engineering,
Pan African University Institute of Basic Sciences Technology and Innovation, Nairobi, Kenya[1]
Faculty of Engineering, University of Nairobi, Kenya[2]
Department of Electrical and Telecommunication Engineering, Multimedia University of Kenya, Nairobi, Kenya[3]

Abstract*—Magnetic resonance imaging (MRI) is an essential modality in contemporary medical diagnosis, offering high-resolution images with excellent soft-tissue contrast of internal organs in a non-invasive, non-ionizing, and non-carcinogenic manner. However, the primary drawback of this technique is a lengthy data-acquisition process, which raises the possibility of picking up artefacts as well as discomfort in patients. To accelerate it, data is acquired at sub-Nyquist rates, and the image is recovered from undersampled, noisy measurements using the compressed sensing (CS) technique. Composite priors in CS–MRI have demonstrated superior performance compared to single priors, despite having limited image priors, which also leads to slower reconstruction speed. To date, deep unfolded networks (DUNs), which integrate the powerful data-driven prior of traditional deep learning (DL) with the strong interpretability of optimization algorithms, are the most outstanding reconstruction technology in MRI acceleration, yielding superior and interpretable results. This study proposes the iterative three-operator splitting (TOS) method to solve the composite regularization problem and a composite denoising network (CDNet) that maintains consistent arithmetic structures with it. The CDNet is implemented using complex-valued DL strategies for richer representation of MRI. Extensive experiments using brain and knee raw k-space from the FastMRI dataset with acceleration factors (AF) ranging from ×2 to ×10 and assessed using peak signal–to–noise ratio (PSNR), structural similarity index (SSIM), and normalized root–mean–squared error (NRMSE) validate the CDNet. Across AFs, CDNet achieved dataset-averaged PSNR/SSIM/NRMSE of 26.50–33.47dB/0.7787–0.8992/0.2562–0.1303 for brain MRI and 28.08–35.52dB/0.7069–0.8734/0.1929–0.0950 for knee MRI. The proposed network yields superior reconstruction accuracy and faster reconstruction speed compared to the iterative TOS for composite regularization, demonstrating the effectiveness of deep unrolling in accelerating iterative optimization. Furthermore, CDNet outperforms other state-of-the-art DUNs in reconstruction accuracy at comparable inference speed, demonstrating its vast potential for improving patient care, scanner throughput, and healthcare economics.

Keywords*—*Accelerated MRI reconstruction; composite priors; iterative three-operator splitting; compressed sensing; deep unfolding; composite denoising network

I. Introduction

Magnetic resonance imaging (MRI) is a critical and widely used research and diagnostic tool in modern medicine due to its non-invasive and non-ionizing acquisition method, which yields excellent contrast images of soft tissue. These properties make it preferred for visualizing the internal organs compared to other medical imaging modalities such as X–rays, computed tomography, and related techniques.

The MRI acquisition technique is based on the fact that the majority of the human body is composed of water molecules, which are made of two hydrogen atoms and one oxygen atom. Hydrogen-atom nuclei consist of a single proton, the tiniest magnetic unit. When exposed to an external magnetic field, these units align with and spin around the axis of the magnetic field at an angular frequency that depends on the strength of the field. This specific frequency is referred to as the precessional or Larmor frequency [1, 2]. The precessional frequency increases with the strength of the magnetic field. When a radio frequency (RF) pulse is applied at this precessional frequency, the spinning protons can absorb the RF energy, causing them to transition from a lower to a higher energy state, either aligning with or opposing the magnetic field. At this point, the protons are considered to be in resonance. Once the RF pulse is discontinued, the protons return to their original state and the energy difference gives rise to the MR signal. Protons in various tissues release energy at different rates due to the distinct chemical compositions and physical properties of each tissue.

The primary drawback of MRI is its slow data acquisition [3-6]. A magnetic resonance (MR) image is formed from multiple k-space acquisitions, separated by time intervals called repetition times or TR. Each acquisition results from applying an RF excitation. However, due to instrumental and physiological limitations, these acquisitions are performed sequentially for a specific field of view (FOV) [1]. Consequently, it takes a considerable amount of time to fully acquire the entire k-space needed to produce a single image.

This slow imaging speed poses a significant challenge, limiting clinical applications, particularly in real-time MRI, such as dynamic cardiac imaging, where only a limited number of samples can be acquired per cardiac cycle. In traditional MRI, k-space sampling adheres to the Nyquist criterion, which depends on the FOV and the MR image resolution. Accelerating MR data acquisition is possible by employing high-magnitude gradients, as these can reduce the TR time. However, the practical use of such rapidly switching gradients is limited because frequent changes can cause peripheral nerve stimulation in patients [1]. This inherent speed limitation of the MRI system has prompted the exploration of alternative technologies to increase MRI speed by undersampling k-space without sacrificing reconstruction quality.

Compressed sensing (CS) [7-10] enables the reconstruction of high-quality MR images from substantially fewer measurements than required by the Nyquist sampling criterion, at an additional computational cost compared to the inverse Fourier transform required for traditional full sampling. In CS, the reconstruction problem is formulated as an optimization that combines an ℓ_1-norm prior term to promote sparsity with an ℓ_2-norm data-fidelity term that enforces consistency with the acquired measurements. Iterative algorithms such as the iterative shrinkage-thresholding algorithm (ISTA) [11], fast iterative shrinkage-thresholding algorithm (FISTA) [12], and alternating direction method of multipliers (ADMM) [13] are then used to solve the problem, approximating the ground truth image.

Composite prior models in CS–MRI reconstruction have consistently demonstrated superior performance compared to single prior models [14-16]. Regularization terms are frequently employed to produce solutions that exhibit various characteristics based on the composite priors. Such priors include the sparse prior in the temporal Fourier domain [17], the total variation prior [18] that enforces sparsity in the spatial domain, and matrix low-rank priors [19], among others. However, the use of multiple priors leads to a composite regularization problem whose solution is computationally intensive, making it unsuitable for practical CS-MRI implementation [1, 14]. Also, there exist limited image priors in CS. In addition to the lengthy reconstruction times, empirical selection of hyperparameters, which is more complex than a single prior model, is nontrivial and difficult to choose for optimal performance. The composite prior model exhibits limited acceleration, up to ×5, with higher acceleration factors yielding images in which artifacts co-exist with useful visual information.

Deep unfolded networks (DUNs) [20-22] with multiple priors can be used to overcome the aforementioned challenges of solving the composite regularization problem with iterative algorithms. This is achieved by unrolling a fixed number of iterations of an optimization algorithm into a trainable end-to-end deep network, where the computationally expensive regularization terms are learned from data, while the computationally efficient data-consistency term is computed explicitly. Consequently, the unrolling converts the problem-specific hyperparameters of iterative algorithms into learnable parameters, thereby alleviating their empirical selection. DUNs have enabled the integration of the powerful data-driven prior of traditional deep learning (DL) with the strong interpretability of optimization algorithms, yielding superior and interpretable reconstruction results, making it the outstanding reconstruction technology in MRI acceleration to date [21, 23]. However, most DUNs [21, 24-28] employ a single DL method to enforce prior information, which may be limited.

To address the aforementioned challenges, we propose a complex-valued deep unrolled composite denoising network (CDNet) for accelerated magnetic resonance image reconstruction with the following specific contributions.

- We solve the composite regularization problem using the iterative three-operator splitting method and reconstruct MR images.
- We develop a complex-valued learned reconstruction architecture, CDNet, that maintains a consistent arithmetic structure via deep unfolding of the iterative three-operator splitting for composite regularization.
- We employ multiple complex-valued DL strategies to effectively capture both fine-grained spatial details and global structural information, thereby providing a richer representation of MRI.
- We conducted extensive experiments on raw complex brain and knee k-space data from the publicly available fastMRI dataset to evaluate CDNet and compare it with other state-of-the-art DUNs, demonstrating that CDNet outperforms existing methods.

II. Related Works

A. Single Priors

Sriram *et al.* [29] proposed an end–to–end variational network, dubbed E2E–VarNet, that accelerated MR image reconstruction in parallel imaging. Unlike prior variational networks that estimated the coil sensitivity maps before reconstruction, the authors in this work jointly learned the maps and the reconstruction process directly from the undersampled k-space. The work unrolled a variational optimization scheme into a network consisting of 12 cascades, with each cascade comprising a data-consistency and U-Net-based regularization block that refines the k-space. In addition to the reconstruction process, the authors added a sensitivity map estimation (SME) module, thus enabling end-to-end training. The E2E – VarNet was trained on knee and brain data from FastMRI and consistently outperformed other variational networks in reconstruction accuracy, as assessed using structural similarity index (SSIM) and peak signal–to–noise ratio (PSNR), especially at higher acceleration factors. Despite E2E–VarNet superseding other variational networks and demonstrating the significance of learning sensitivity maps, it exhibited some limitations. First, the SME module added half a million more parameters than other variational networks, increasing memory and computational costs. Secondly, the model evaluation relied primarily on quantitative metrics, namely normalized mean squared error (NMSE), SSIM, and PSNR, which may not capture subtle, clinically relevant details. This would necessitate further clinical validation before deployment to avoid a degradation in diagnostic accuracy.

Zhou *et al.* [30] proposed a dual–domain recurrent network, dubbed DuDoRNet, that jointly restores the k-space and image while embedding a deep T1 prior. Each cascade contained two dilated residual dense networks, one for k-space restoration and the other for image restoration, each with a data consistency layer and using a TI-weighted MRI for complementary structural guidance. The authors conducted extensive experiments on the network with three undersampling patterns, namely radial, Cartesian, and spiral trajectories. They used an in-house MRI dataset of 20 patients consisting of 252, 36, and 72 training, validation, and testing 2D images on acceleration of up to 6 using 5 cascades. The results were evaluated using PSNR, SSIM, and mean squared error (MSE) and demonstrated superior performance compared to other DL-based methods and CS-based methods. Despite the good reconstruction performance, the DuDoRNet has several limitations. First, it relies on the availability of a fully sampled T1 scan, which may be absent in protocol-limited clinical settings or time-critical medical examinations. Second, the recurrent dual-domain architecture increases computational complexity and memory requirements, limiting real-time clinical deployment. Thirdly, the network evaluation focused primarily on anatomical data, while other MRI data, such as pathological data, remain unexplored.

Xiang *et al.* [31] combined the free tuning and strong generalization of deep neural networks with the interpretability of FISTA to yield an unrolled network, dubbed FISTA-Net, for solving inverse problems in imaging. The FISTA-Net unfolds n FISTA iterations into n blocks, with each block consisting of a convolutional neural network (CNN) – based proximal mapping module, a gradient descent module, and a momentum update module. The FISTA-Net was evaluated quantitatively and visually using X-ray, computational tomography, and electromagnetic tomography. The experimental results show that it performs superiorly to the traditional iterative FISTA, an ISTA-based unrolled network, and other iterative and data-driven baseline algorithms. The network was not validated on MRI datasets with raw k-space measurements, limiting insights into its performance in practical MRI reconstruction scenarios.

Fabian *et al.* [32] introduced a hybrid unrolled multi-scale framework named HUMUS-Net that combines transformers and CNNs to accelerate MR image reconstruction. HUMUS-Net leverages the capability of transformers to identify long-range dependencies via self-attention and the proficiency of CNNs in efficiently extracting local features for MR image reconstruction. This design enables the network to extract multi-scale features by using transformers for low-resolution features and convolutions for high-resolution features. Additionally, HUMUS-Net integrates a sensitivity map estimator for multi-coil MRI reconstruction and employs unrolled iterative reconstruction steps, inspired by iterative optimization algorithms. The unrolled network achieved state-of-the-art performance, as measured by SSIM, PSNR, and NMSE, outperforming other existing networks. Nonetheless, HUMUS-Net is computationally intensive and requires substantial memory, particularly due to the transformer, which limits the depth of the network and patch size. Furthermore, the network requires fixed-size inputs, limiting its adaptability in clinical settings.

Zhang *et al.* [33] introduced an unrolled network, T^2LR-Net, that adaptively learns the optimal transformed domain to leverage tensor low-rank priors in dynamic MRI reconstruction. This method extended the tensor singular value decomposition (t-SVD) to incorporate multidimensional unitary transforms and introduced a novel unitary-transformed tensor nuclear norm (UTNN). By employing CNN modules, the approach surpassed the constraints of manually crafted transforms, learning appropriate transformations driven by reconstruction precision. T^2LR-Net unfolds N iterations of the ADMM algorithm into N modules, each comprising a reconstruction block, a multiplier update block, and a transformed-tensor low-rank prior block integrated with a CNN. Experimental outcomes assessed using signal-to-noise ratio (SNR) and SSIM demonstrated superior performance compared to other unrolling network-based and optimization-based methods. Nevertheless, the repeated SVD operations heighten the network's computational complexity, necessitating significant GPU memory and training time. The singular value thresholding (SVT) operators in each iteration stage may become nondifferentiable under certain conditions, such as when the matrix being solved has repeated singular values, potentially halting training. Additionally, performance slightly diminishes in multi-coil and prospective acquisition scenarios.

Vornehm *et al.* [34] introduced a variational network-based reconstruction framework dubbed cineVN, which unrolled an iterative optimization scheme into a deep network while incorporating the spatio–temporal structure of dynamic cardiac data. The cineVN network architecture consisted of 15 cascades, each containing a refinement term and a data consistency term, with a conjugate gradient block after every third cascade. The coil sensitivity maps were pre-estimated, and the complex data were split into real and imaginary parts. The proposed cineVN was evaluated on in vivo cardiac MRI acquisitions and demonstrated improvements over a spatiotemporal VN and a temporal total variation regularization – CS based approach (TTV-CS) in retrospective evaluation and TTV-CS and ℓ_1-wavelet regularization CS based approach in prospective evaluation. However, there is a potential breath-hold inconsistency between accelerated cineVN and GRAPPA reference acquisitions, as residual variations in respiratory state may persist despite exhalation-only imaging. The proposed network was also evaluated on a specific set of acquisitions and protocols, and a broader validation across different acquisition settings, scanners, and patient populations is essential before clinical deployment.

B. Multiple Priors

Ke *et al.* [20] presented a network dubbed SLR-Net which unfolded the ISTA procedure for dynamic MR imaging. Unlike other networks that focus on sparse priors of MR images alone, the authors in this work added a learned singular value thresholding to explore low-rank priors. The authors unrolled 10 iterations of ISTA into 10 cascades of SLR-Net, with each cascade comprising a sparse prior module, low-rank prior module, reconstruction module, and a multiplier update layer while maintaining the arithmetic structure of ISTA. The learned low–rank prior improved the reconstruction accuracy assessed using PSNR, MSE, and SSIM compared to CS and sparsity-driven DL methods. However, these SVD operations added more computation time, which was limited to 0.5s, a range achieved using the supercomputing power of a GPU. In the absence of such GPUs in clinical deployment, real-time imaging may not be achieved. The module also added the memory requirement of the GPU, increasing hardware requirements.

Huang *et al.* [35] proposed a network dubbed the low-rank plus sparse (L+S) network to overcome two common failings of CS-based reconstruction: limited acceleration and the empirical selection of L+S parameters. The authors unrolled the alternating linearized minimization method to learn the regularization parameters. The L+S-Net contains 10 states, corresponding to 10 iterations of the iterative algorithm, with each stage comprising a data consistency layer, a sparse prior layer, and a low–rank prior layer. Experimental results on cardiac cine datasets demonstrate superior results compared to other existing CS and DL methods, especially at high acceleration factors. Despite the superior reconstruction results, L+S-Net exhibited several limitations. First, the network has SVD computations, which lead to extended reconstruction times. Second, the effect of the low-rank prior may degrade when the number of frames is reduced, since reconstruction is performed for each batch of frames, rendering it unsuitable for low-latency imaging. Third, the network was tested only on heart images, and further testing on other body parts is essential for its validation.

Zhang *et al.* [21] presented an unrolled network, JotlasNet, that incorporated attention-based sparse priors and a tensor low-rank prior for reconstructing dynamic MR images. The tensor low-rank prior exploited structural correlations in the data. The authors used a CNN to learn the priors. They unfolded the network using the composite splitting algorithm and set the number of cascades to 15. JotlasNet yielded better reconstruction results, as evaluated using PSNR and SSIM, on single- and multicoil data compared to other DUNS. However, JotlasNet is prone to numerical instability, which hinders the training process. The gradient of SVD is unstable, especially when singular values are repeated, leading to NaN gradients that hinder training. Also, the SVD significantly increases reconstruction time due to its computational complexity.

From the aforementioned limitations of existing DUNs, we propose the CDNet that allows effective integration of multiple priors in the reconstruction model. The CDNet is not prone to numerical instability and effectively combines multiple deep complex-valued DL strategies to capture different attributes of MR images and offer richer representation in MRI.

III. METHODOLOGY

A. Problem Formulation for Accelerated MRI

Magnetic resonance imaging (MRI) acquisition captures spatially dependent resonance-frequency information from the region under investigation. The MR images in the spatial domain are obtained by taking an inverse Fourier transform of the captured data. Accelerating MRI via undersampling results in an underdetermined system with an infinite set of solutions. The compressive sensing (CS) technique enables recovery of the MR image of interest by solving the basis pursuit denoising (BPDN) problem. The BPDN problem solution is obtained by solving the unconstrained optimization problem given in Eq. (1)

$$\hat{x} = \underset{x}{argmin} \underbrace{\|Fx - s\|_2^2}_{f(x)} + \lambda \underbrace{\|x\|_1}_{\Re(x)} \tag{1}$$

where, F denotes the forward model that maps the MR image x to the scanned data s. The ℓ_1 norm promotes sparsity in a suitable transform such as wavelets, and the λ is a regularization parameter that balances the sparsity $\Re(x)$ and data consistency term $f(x)$.

The solution to Eq. (1) closely approximates the original image. To improve the solution of Eq. (1) and consequently the reconstruction results, we can add additional sparsifying norms to increase overall sparsity. One such norm is the total variation (TV) norm [18], which promotes sparsity of image gradients. Therefore, the addition of the TV norm in Eq. (1) improves the solution by promoting sparsity in the spatial domain alongside the ℓ_1 norm. The general optimization problem can be given as in Eq. (2).

$$\hat{x} = \underset{x}{\text{minimize}}\ g(x) = f(x) + \sum_{i=1}^{k} \lambda_i \Re_i(x) \tag{2}$$

The data consistency term $f(x)$ is a continuously differentiable function with a Lipschitz constant, while the $\Re_i(x)$ terms are nonsmooth convex functions. When $k = 1$, Eq. (2) reduces to Eq. (1), which can be easily solved using two-operator splitting algorithms such as ISTA, FISTA, and ADMM. When $k > 1$, Eq. (2) becomes computationally expensive to solve using the two-operator splitting algorithms due to multiple regularization terms.

B. Iterative Three Operator Splitting Algorithm for Composite Regularization

The iterative three-operator splitting (TOS) algorithm, first proposed by [36] and whose convergence was analyzed in [37], can be used to solve Eq. (2) for $k > 1$ in a computationally efficient manner. For example, we consider the case when $k = 2$, Eq. (2) becomes;

$$\hat{x} = \underset{x}{\text{minimize}}\ g(x) = f(x) + \lambda_1 \Re_1 + \lambda_2 \Re_2 \tag{3}$$

Following the works of [10, 14, 15], we can consider the total variation (TV) norm for spatial sparsity and an ℓ_1 norm in a suitable transform domain for transform sparsity. Eq. (3) can then be rewritten as;

$$\hat{x} = \underset{x}{\text{minimize}}\ g(x) = \underbrace{\|Fx - s\|_2^2}_{f(x)} + \lambda_1 \underbrace{\|\psi x\|_1}_{\Re_1(x)} + \lambda_2 \underbrace{\|x\|_{TV_{iso}}}_{\Re_2(x)} \tag{4}$$

where, ψ is a suitable transform such as wavelets and $\|x\|_{TV_{iso}}$ is the isotropic TV given by;

$$\|x\|_{TV_{iso}} = \sum_{i,j} \sqrt{\left|x_{i+1,j} - x_{i,j}\right|^2 + \left|x_{i,j+1} - x_{i,j}\right|^2} \tag{5}$$

Eq. (4) can be solved iteratively using the TOS strategy. At each iteration, n, a gradient step, a proximal operator of TV, and a proximal operator of ℓ_1 are combined to produce a relaxed output for the next iteration. The TOS for $k = 2$ is given in Algorithm 1.

Algorithm 1: Three Operator Splitting Algorithm for $k = 2$

Inputs: Complex-valued zero-filled image y_0^0, Binary Mask *M*, number of iterations *N*, step size α and relaxation λ
For $n = 0, 1, 2, 3 \ldots N$
$x_1^n = prox_{\alpha\Re_1}(y_0^n)$
$\nabla f(x_1^n) = F^{-1}(M \odot (Fx_1^n - s))$
$y_1^n = 2x_1^n - y_0^n - \alpha \nabla f(x_1^n)$
$x_2^n = prox_{\alpha\Re_2}(y_1^n)$
$x^{n+1} = y_0^n + \lambda(x_2^n - x_1^n)$
End
Return x_1^N

where, *prox* is the proximal operator defined as in Eq. (6) for a function φ [35].

$$prox_{\gamma\varphi}(x) = \underset{y \in \mathbb{C}^d}{argmin}\ \varphi(y) - \frac{1}{2\gamma}\|x - y\|_2^2 \qquad (6)$$

For $k > 2$ the algorithm 1 can be reformulated as iterates of *prox* of $\Re_i$ and the gradient step to give Algorithm 2.

Algorithm 2: Three Operator Splitting Algorithm for $k > 2$

Inputs: Complex-valued zero-filled image y_0^0, Binary Mask *M*, number of iterations *N*, step size α and relaxation λ
For $n = 0, 1, 2, 3 \ldots N$
$x_1^n = prox_{\alpha\Re_1}(y_0^n)$
$\nabla f(x_1^n) = F^{-1}(M \odot (Fx_1^n - s))$
$y_1^n = 2x_1^n - y_0^n - \alpha \nabla f(x_1^n)$
$x_2^n = prox_{\alpha\Re_2}(y_1^n)$
$\nabla f(x_2^n) = F^{-1}(M \odot (Fx_2^n - s))$
$y_2^n = 2x_2^n - y_1^n - \alpha \nabla f(x_2^n)$
$x_3^n = prox_{\alpha\Re_3}(y_2^n)$
•
•
•
$x_{K-1}^n = prox_{\alpha\Re_{K-1}}(y_{K-2}^n)$
$\nabla f(x_{K-1}^n) = F^{-1}(M \odot (Fx_{K-1}^n - s))$
$y_{K-1}^n = 2x_{K-1}^n - y_{K-2}^n - \alpha \nabla f(x_{K-1}^n)$
$x_K^n = prox_{\alpha\Re_K}(y_{K-1}^n)$
$x^{n+1} = y_0^n + \lambda(x_K^n - x_1^n)$
End
Return x_1^N

C. Composite Denoising Network

Solving the $\Re_i(x)$ terms in the multiple-regulation optimization problem for MR image reconstruction in Eq. (2) efficiently and effectively is non-trivial. This may require creating subproblems and solving them using iterative algorithms, either simultaneously or sequentially, to achieve the desired solution. This leads to iterations within the main iterations of the TOS algorithm, which can significantly increase reconstruction time. Also, choosing the values of λ and α for effective MR image reconstruction is challenging. Motivated by these challenges of the iterative TOS algorithm for multiple regularization, we derive a learned reconstruction scheme dubbed composite denoising network (CDNet) by deep unrolling its iterations. We replace $prox_{\alpha\Re_i}$ with parametrized complex-valued DL strategies $\xi_{k\theta_i}$. The use of multiple complex-valued DL methods in the reconstruction network effectively combines their advantages, enabling the model to capture both fine-grained spatial details and global structural information. Complex-valued reconstruction networks also offer a richer representation of MR images as they are inherently complex. Algorithm 2 can be unrolled to Algorithm 3 given below.

Algorithm 3: Composite denoising network (CDNet)

Inputs: Complex-valued zero-filled image y_0^0, Binary Mask *M*, number of iterations *N*, step size α and relaxation λ
Learnable parameters: step size α, relaxation λ and $\xi_{k\theta_i}$
For $n = 0, 1, 2, 3 \ldots N$
$x_1^n = \xi_{1\theta_i}(y_0^n)$
$\nabla f(x_1^n) = F^{-1}(M \odot (Fx_1^n - s))$
$y_1^n = 2x_1^n - y_0^n - \alpha \nabla f(x_1^n)$
$x_2^n = \xi_{2\theta_i}(y_1^n)$
$\nabla f(x_2^n) = F^{-1}(M \odot (Fx_2^n - s))$
$y_2^n = 2x_2^n - y_1^n - \alpha \nabla f(x_2^n)$
$x_3^n = \xi_{3\theta_i}(y_2^n)$
•
•
•
$x_{K-1}^n = \xi_{K-1\theta_i}(y_{K-2}^n)$
$\nabla f(x_{K-1}^n) = F^{-1}(M \odot (Fx_{K-1}^n - s))$
$y_{K-1}^n = 2x_{K-1}^n - y_{K-2}^n - \alpha \nabla f(x_{K-1}^n)$
$x_K^n = \xi_{K\theta_i}(y_{K-1}^n)$
$x^{n+1} = y_0^n + \lambda(x_K^n - x_1^n)$
End
Return x_1^N

The block diagram of CDNet is given in Fig. 1. $\xi_{k\theta_i}$ are learnable parameters for composite denoising while D_k are data consistency terms computed numerically.

1) Discriminative learning for CDNet: We demonstrate the effectiveness of CDNet in accelerating MRI reconstruction by choosing K = 2 achieving a favorable trade-off between computational complexity, reconstruction time, and reconstruction accuracy. We propose a complex-valued CNN (CvCNN) to capture the overall anatomical structure and a complex-valued Unet (CvUnet) to capture the spatially local features and also enhance global features for $\xi_{1\theta_i}$ and $\xi_{2\theta_i}$ respectively. These DL methods offer lower computational complexity and hardware requirements compared to transformers and generative adversarial networks (GANs), thus preferred for CDNet. CvCNN and CvUNet differ in architectural design, but they are implemented in the same building blocks, namely, complex-valued convolution (CVC), complex-valued rectified linear unit ($\mathbb{C}$ Re L U) and radial batch normalization (RBN). While both architectures employ identical complex-valued building blocks, CvCNN uses a sequential feed-forward design, whereas CvUNet adopts an encoder–decoder structure with skip connections.

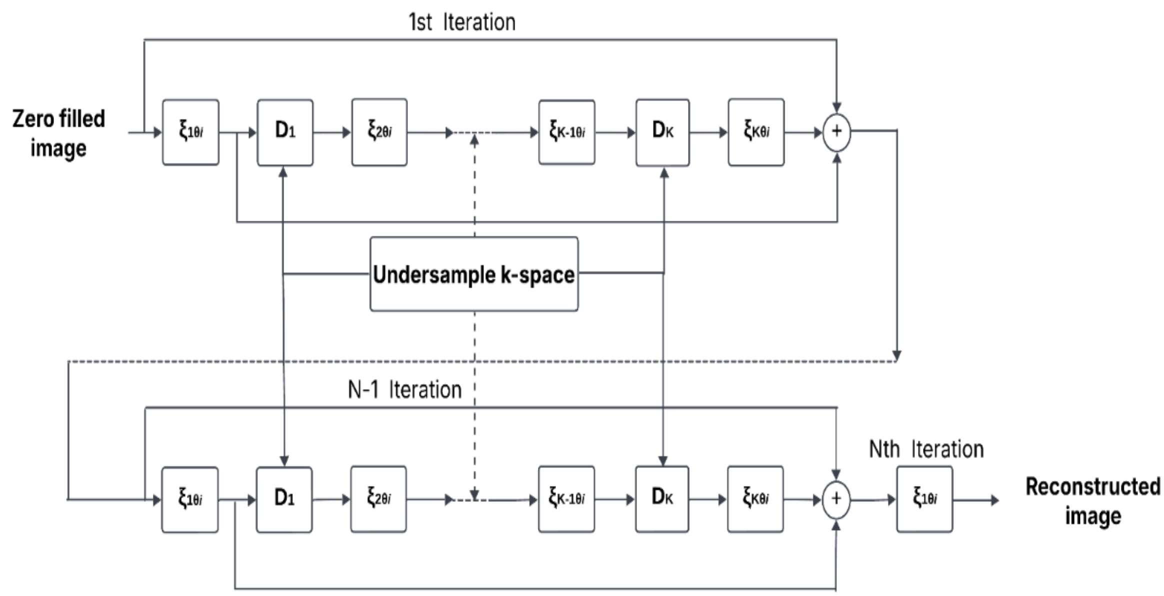

Fig. 1. Block diagram of the proposed CDNet.

The complex-valued convolution between the input zero-filled image $y_n^k = a + jb$ and kernel $k = c + jd$ can be realized using real-valued arithmetic by convolving the kernels c or d with each a and b as given in Eq. (7).

$$y_k^n * k = (c * a - d * b) + j(d * a + c * b) \tag{7}$$

where, $*$ represent convolution.

The RBN module stabilizes convergence and accelerates network training. The magnitude is scaled while preserving the phase, ensuring relative differences in complex quantities are preserved. We express the complex data y_{RBN} in polar form as given in Eq. (8) to obtain its magnitude f and phase θ. We normalize f as in Eq. (9)

$$y_{RBN} = f e^{j\theta} \tag{8}$$

$$f_{RBN} = \left(\frac{f - \mu_f}{\sqrt{\sigma_f^2}} \right) \kappa_1 + \kappa_2 + \varepsilon \tag{9}$$

where, σ_f^2 and μ_f are the variance and mean of f, κ_1 and κ_2 are trainable parameters for scaling and shifting the data distribution, and ε is a constant for ensuring f_{RBN} remains positive.

The network was activated with a $\mathbb{C}ReLU$ obtained by applying a real-valued ReLU to the real g_R and imaginary part g_I of the complex data g separately [38].

$$\mathbb{C}ReLU(g) = ReLU(g_R) + j\,ReLU(g_I) \tag{10}$$

2) Loss function: The CDNet was trained with a compound loss L^{comp} of ℓ_1 and multi-scale SSIM (MS-SSIM) loss given in Eq. (11) and its auxiliary parameters given in Eq. (12) to (15). The balancing weight χ was determined empirically in [39] to be 0.84, while [40] demonstrated the superiority of L^{comp} compared to the MSE, Huber, SSIM, perceptual loss, MSSIM, and ℓ_1 loss in training medical image reconstruction models.

$$L^{comp} = \chi L^{MS-SSIM}(|rec|, |gt|) + (1 - \chi) L^{\ell_1}(|rec|, |gt|) \tag{11}$$

where, rec, gt, $L^{MS-SSIM}$, L^{ℓ_1} is the reconstructed image, ground truth image, MS-SSIM loss given in Eq. (11), and ℓ_1 loss given in Eq. (16), respectively, and $|.|$ denotes magnitude.

$$L^{MS-SSIM} = l_N\left(|rec|, |gt|\right)^{\upsilon_N} \prod_{j=1}^{N} \left[c_j\left(|rec|, |gt|\right) \right]^{\delta_j} \left[s_i\left(|rec|, |gt|\right) \right]^{\tau_j} \tag{12}$$

where, l, c and s are luminance, contrast, and structure given in Eq. (13), (14), and (15), respectively, and υ, δ_j and τ_j are weights of the distortion factors at different levels.

$$l(|rec|, |gt|) = \frac{2\mu_{|rec|}\mu_{|gt|} + C_1}{\mu_{|rec|}^2 + \mu_{|gt|} + C_1} \tag{13}$$

where, $\mu_{|rec|}$ and $\mu_{|gt|}$ are averages of reconstructed and ground truth images, respectively, and C_1 is a constant.

$$c(|rec|, |gt|) = \frac{2\sigma_{|rec|}\sigma_{|gt|} + C_2}{\sigma_{|rec|}^2 + \sigma_{|gt|}^2 + C_2} \tag{14}$$

where, $\sigma_{|rec|}$ and $\sigma_{|gt|}$ are variances of the reconstructed and ground truth images, respectively, and C_2 is a constant.

$$s(|rec|, |gt|) = \frac{\sigma_{|rec||gt|} + C_3}{\sigma_{|rec|}\sigma_{|gt|} + C_3} \tag{15}$$

where, $\sigma_{|rec||gt|}$ is the covariance between the reconstructed and ground truth images.

$$L^{\ell_1} = \left\| |rec| - |gt| \right\|_1 \tag{16}$$

D. Implementation Details

1) Iterative details: We implemented the iterative TOS for $k = 2$ in the open-source Python-based PyTorch library on a PC with an Intel CPU of 2.30 GHz and 64 GB RAM. $\mathfrak{R}_1$ was solved using the Fast Iterative Shrinkage-Thresholding Algorithm FISTA [12] with ψ chosen to be the discrete wavelet transform for promoting sparsity in the transform domain, while $\Re_2$ was solved using the anisotropic TV norm given in Eq. (5) for promoting sparsity in the spatial domain. The algorithm was set to 200 iterations, with each consisting of one thresholding step, one data consistency step, one anisotropic TV step, and a relaxation update. The relaxation λ was set to 0.05, while the step size α was adaptive.

2) Unrolled details: The CDNet was implemented using the open - source Python-based PyTorch library in a GPU–enabled Kaggle environment utilizing an NVIDIA Tesla P100. The number of cascades was set to 4, a number validated empirically through experiments and offering a good tradeoff between reconstruction time, hardware requirements, and reconstruction accuracy. The CvCNN complex convolutions use a kernel size of 3×3, 5 convolutional layers, and 32 hidden channels. The CvUnet uses 3 levels with a bottleneck, employing 3×3 convolutions, 2×2 max-pooling and transposed convolutions for down- and up-sampling, 32 base channels, and symmetric skip connections. An Adam optimizer [41] was employed to train the proposed network with a learning rate of 0.002 and momentum parameters set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The training was conducted for ×2, ×3, ×4, ×5, ×7, and ×10 accelerations using a batch size of 3 and limited to 55 and 35 epochs for the brain and knee datasets, respectively, to avoid overfitting and preserve generalization.

E. Performance metrics

The proposed CDNet was evaluated quantitatively using PSNR, SSIM, and NRMSE computed against the ground truth images, while sample reconstructed images are given for qualitative assessment. The reconstruction time was measured in seconds. The PSNR between the reconstructed rec and ground truth gt image, whose magnitude is normalized to [0, 1], is given by Eq. (17).

$$PSNR(|rec|, |gt|) = -10\, log_{10} MSE \tag{17}$$

where, MSE is computed as in Eq. (18).

$$MSE(|rec|, |gt|) = \frac{\sum_{i=1}^{R}\sum_{j=1}^{C}(|rec|-|gt|)^2}{RC} \tag{18}$$

where, R and C is the number of rows and columns, respectively. The PSNR ranges from 0 to +∞ , with higher value.

values indicating better reconstruction results.

The SSIM between rec and gt is given in Eq. (19).

$$SSIM\left(|rec|,|gt|\right) = l\left(|rec|,|gt|\right)^{\upsilon} c\left(|rec|,|gt|\right)^{\delta} s\left(|rec|,|gt|\right)^{\tau} \tag{19}$$

where, l, c and s are luminance, contrast, and structure defined in Eq. (13), (14), and (15), respectively, and υ, δ and τ are weights of the distortion factors at different levels. SSIM ranges from 0 to 1, with 1 indicating complete similarity and vice versa.

The NRMSE between rec and gt is given in Eq. (20).

$$NRMSE(|rec|, |gt|) = \sqrt{\frac{MSE(|rec|,|gt|)}{MSE(|gt|,0)}} \tag{20}$$

A low NRMSE indicates good reconstruction results and vice versa.

IV. Results and Discussion

A. Dataset

The brain and knee raw complex k-space data from the FastMRI dataset [42] were used for this research. This is the largest public dataset for MR slices created in collaboration between Facebook AI Research and NYU Langone Health. We selected the single-coil knee dataset, which comprises 973, 199, and 108 volumes for training, validation, and testing, respectively, each containing 34742, 7135, and 3903 slices, respectively. However, the proposed network was evaluated using MR volumes with high-quality slices without fat suppression. This enabled the proposed network to learn anatomical details and the general structure from high-SNR slices with strong overall contrast. A total of 170, 50, and 60 volumes, each containing 6150, 1808, and 2229 slices, respectively, were selected to evaluate the proposed network.

Brain raw complex k-space was also used to evaluate the proposed network. Since the FastMRI dataset provides multicoil k-space for brain slices and the proposed network was evaluated using single-coil data, we enumerated the single-coil data using an adaptive root–sum–of–squares with phase. The magnitude, phase, and complex image were obtained as given in Eq. (21), (22), and (23), respectively.

$$|gt(i,j)| = \sqrt{\sum_{n=1}^{N}|gt_n(i,j)|^2} \tag{21}$$

$$e^{j\theta_{ave(i,j)}} = \frac{\sum_{n=1}^{N} gt_n(i,j)}{|\sum_{n=1}^{N} gt_n(i,j)|} \tag{22}$$

$$gt(i,j) = |gt(i,j)|.\, e^{j\theta_{ave(i,j)}} \tag{23}$$

where, N is the number of coils and $gt_n(i,j)$ is the complex-valued ground truth image for the nth coil. A total of 500, 100, and 150 volumes containing 7910, 1584, and 2392 slices, respectively, were enumerated and their k-spaces and images stored in hierarchical data format version 5 (HDF5) for efficient organization and retrieval during training. A sample of enumerated brain magnitude and phase MR images of size 768 × 396 is shown in Fig. 2.

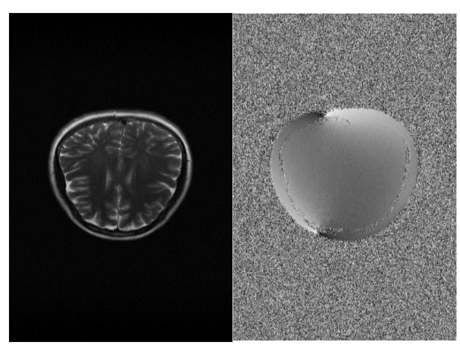

Fig. 2. From left to right, a sample magnitude image and its corresponding phase image.

B. Data Acquisition and Preprocessing

The acquisition was accelerated by undersampling fewer samples with a variable-density pattern that mimics realistic MRI scanners. The phase-encoding lines from the center region corresponding to a percentage equal to the acceleration factor (AF) were retained, as this region contains information about the contrast and gross structure of the MR image. The other lines were sampled probabilistically, ensuring denser sampling near the center and sparser sampling towards the periphery, reducing acquisition artifacts.

The input images, namely the zero-filled image obtained as an inverse fast Fourier transform (IFFT) of the undersampled k-space filled with zeros at the missing k-space positions, and the ground truth images obtained as an IFFT of the fully sampled k-space, were normalized to [0, 1] while preserving the phase according to Eq. (24)

$$gt_N(i,j) = \frac{gt(i,j)}{max_{i,j}|gt(i,j)|+\varepsilon} \tag{24}$$

where, $gt_N(i,j)$ is the normalized ground truth image and ε is a small constant for numerical stability set to 10^{-8}.

The knee MR images were center-cropped to 320 × 320, while the brain MR images were center-cropped to 192 × 192. The 192 × 192 was achieved by resizing in the spatial domain from the original sizes to 320 × 320, followed by a resizing in the Fourier domain to 192 × 192. This ensured a reduction in the background, which is a region of non-interest in the medical domain, while preserving the region of interest. This allows the proposed network to learn useful anatomical details with less computational overhead without compromising the MR image. A sample brain MR image of size 640 × 320 resized to 192 × 192 is shown in Fig. 3.

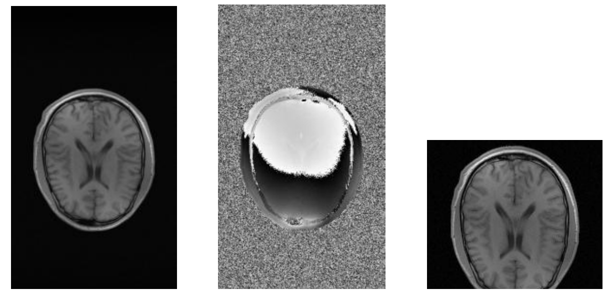

Fig. 3. From left to right: sample magnitude brain MR image and its corresponding phase image of size 640 × 320 resized to 192 × 192.

C. Iterative TOS for composite regularization

Fig. 4 shows sample reconstructed MR images, with the first row showing the brain image and the second row showing the knee image at ×2 to ×5 AF. Ground-truth images are also provided as baselines for comparison. The performance metrics are given in Table I. The average reconstruction time for brain MR is 1.4340s, while that of knee MR is 2.6721s. This difference is due to unequal resizing of the two images, with the knee being resized to 320 × 320 while the brain is resized to 192 × 192. The performance metrics decreased as AF increased, due to fewer measurements acquired for reconstruction. At AFs ×4 and ×5, residual acquisition artifacts coexist with the anatomical structure of the MR image.

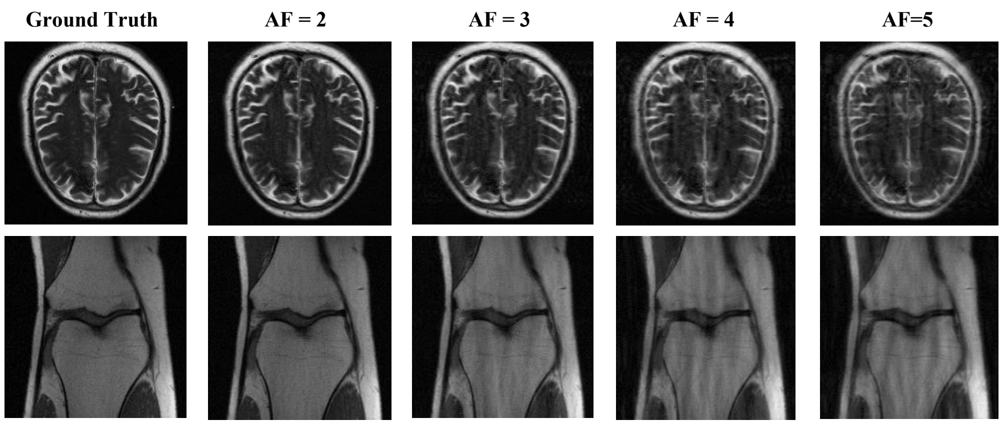

Fig. 4. Sample iterative TOS reconstruction results.

TABLE I. PERFORMANCE METRICS FOR SAMPLE RECONSTRUCTED MR IMAGES IN FIG. 4

	Brain				Knee			
AF	**PSNR (dB)**	**SSIM**	**NRMSE**	**Time (s)**	**PSNR (dB)**	**SSIM**	**NRMSE**	**Time (s)**
×2	31.95	0.9000	0.0861	1.4465	32.21	0.8280	0.0775	2.6039
×3	27.01	0.7920	0.1548	1.4759	27.33	0.7838	0.1263	2.7217
×4	24.49	0.7390	0.2073	1.4131	25.29	0.7166	0.1586	2.6298
×5	23.06	0.6982	0.2573	1.4005	24.28	0.6845	0.1747	2.7331

D. CDNet

1) The effect of the number of cascades n: We investigated the effect of increasing the number of cascades in the CDNet using the brain k-space with a ×5 acceleration and the statistical quantitative results presented in Table II. The PSNR, SSIM, NRMSE, and time report the average values computed over the entire brain test images. The reconstruction time increases with an average of 0.02875s per added cascade. The quantitative metrics improved greatly when learning started, with PSNR, SSIM, and NRMSE reporting improvements of 4.86 dB, 0.2040, and 0.1506, respectively. These results continued to improve when using two cascades, with PSNR, SSIM, and NRMSE reporting improvements of 0.86 dB, 0.0201, and 0.0183 compared to using a single cascade. The third cascade also showed improvements of 0.49 dB, 0.0225 and 0.0052 in PSNR, SSIM and NRMSE, respectively, compared to learning with two cascades. Since the results were still improving significantly, we also experimented with the CDNet with four cascades. The results showed improvements of 0.24 dB, 0.0046, and 0.0083 for PSNR, SSIM, and NRMSE, respectively. The fifth cascade did not significantly improve the quantitative results, indicating saturation. However, the reconstruction time continues to increase, and hence the choice of $N = 4$ which offers a good trade-off between reconstruction time and accuracy for the proposed CDNet.

A sample reconstructed image and its corresponding error map computed as the difference between the ground truth and the reconstructed image are shown in Fig. 5 for qualitative assessment. The zero-filled and ground-truth images are provided as a baseline for visual comparison. The reconstructed image using a single cascade shows significant improvement over the zero-filled image. However, noticeable differences between the reconstructed and ground truth image are visible. The image reconstructed using two cascades resembles the ground truth image better, with fine anatomical details clearer than in the single-cascade reconstruction. Still, noticeable differences between the reconstructed and ground truth images are visible. The reconstructed image using three cascades is visually better than that using two cascades, with edges showing fine anatomical details. The images reconstructed with four and five cascades are imperceptible to the human eye compared to the ground truth, despite their error maps showing differences. Table III gives the performance metrics for the sample MR image shown in Fig. 5.

TABLE II. DETERMINING THE NUMBER OF CASCADES *N* USING BRAIN DATA ON ×5 ACCELERATION

No of cascades *n*	Image	PSNR (dB)	SSIM	NRMSE	Time (s)
0	Zero filled	22.60	0.5806	0.3822	-
1	Reconstructed	27.46	0.7846	0.2316	0.1020
2	Reconstructed	28.32	0.8047	0.2133	0.1317
3	Reconstructed	28.81	0.8272	0.2081	0.1600
4	**Reconstructed**	**29.05**	**0.8318**	**0.1998**	**0.1885**
5	Reconstructed	29.09	0.8320	0.1996	0.2170

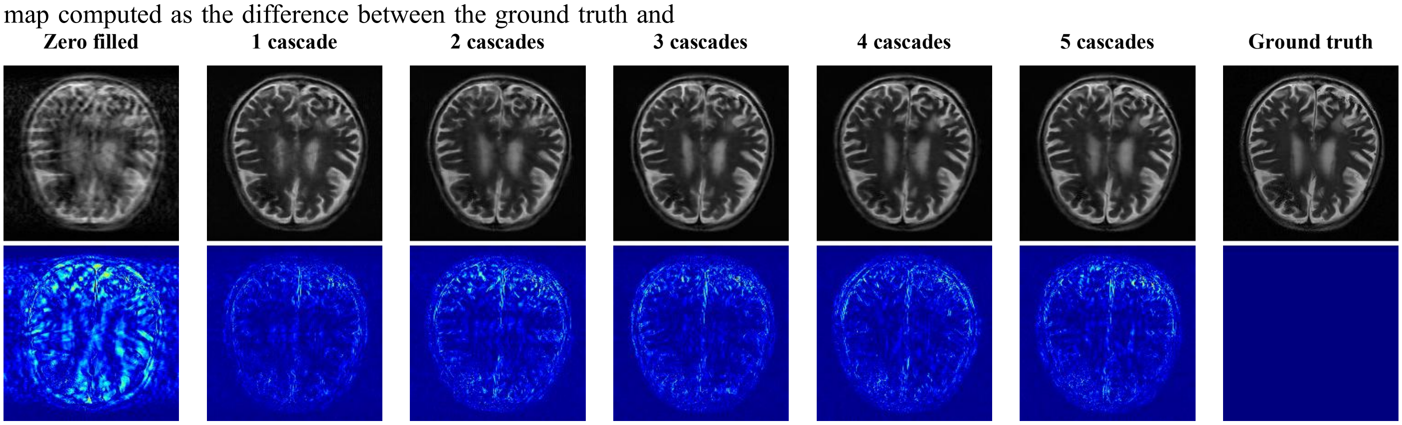

Fig. 5. Sample brain MR images reconstructed using different numbers of cascades.

TABLE III. PERFORMANCE METRICS FOR THE BRAIN MR IMAGE SHOWN IN FIG. 5

No of cascades *n*	Image	PSNR (dB)	SSIM	NRMSE
0	Zero filled	20.86	0.6025	0.6222
1	Reconstructed	27.64	0.8293	0.4212
2	Reconstructed	27.89	0.8457	0.4151
3	Reconstructed	28.50	0.8844	0.4007
4	**Reconstructed**	**29.13**	**0.8926**	**0.3865**
5	Reconstructed	28.89	0.8899	0.3918

2) CDNet learning characteristics: Fig. 6 shows the train and validation loss curves over epochs under different AFs for brain and knee datasets, respectively.

The losses showed a steeper decline for the first 30 and 20 epochs for the brain and knee datasets, respectively, indicating faster initial learning. The losses later decreased slowly up to the 40th epoch and 25th epoch for the brain and knee datasets, respectively, after which they oscillated with small fluctuations, indicating convergence. The losses increased with an increase in AF as less data was used for reconstruction, making it a more challenging task. The train and validation curves for each AF exhibit a small gap, indicating that overfitting is effectively controlled.

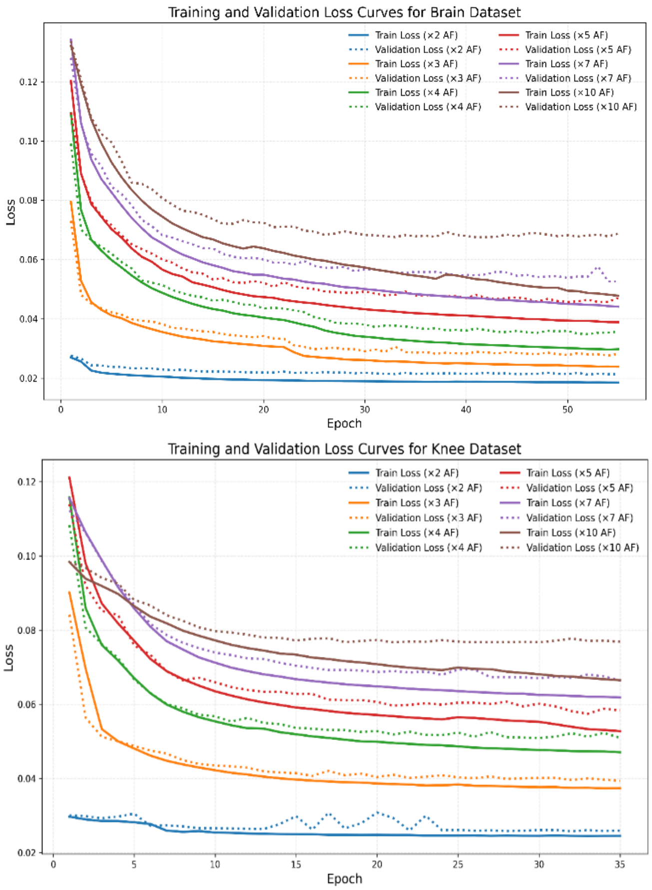

Fig. 6. From top to bottom, plots of training and validation loss over epochs for the brain and knee datasets, respectively.

3) Quantitative results

a) Brain data results: Table IV presents the quantitative statistical results for brain data, reporting average PSNR, SSIM, and NRMSE values computed over the entire test brain images. The zero-filled results were used as a baseline for comparison. The table also provides the improvements from the zero-filled images. The zero-filled results decreased as the AF was increased from ×2 to ×7, as fewer samples were acquired. The zero-filled results for ×10 are slightly higher than those of ×7 due to a higher percentage of retained center lines without a significant decrease in the number of acquired phase-encoding lines. The reconstructed result decreases when the AF is increased from ×2 to ×10 due to fewer samples being acquired, which increases acquisition artifacts and makes it challenging for the CDNet to denoise. The CDNet showed great improvement in reconstruction accuracy, highlighted in bold for all AFs.

b) Knee data results: Table V presents the statistical results for knee data, reporting the mean values of PSNR, SSIM, and NRMSE computed over the entire knee test data. The average reconstruction time for the knee MR image is 0.2963 seconds per slice. This is higher than the reconstruction time for a brain MR image of 0.1885 seconds per slice. The higher reconstruction time for knee MR images is due to their larger size (320 × 320) compared to brain MR images (192 × 192). The background of brain MR images was large and the region of interest could be carried comfortably by cropping it in a smaller size compared to knee images. The zero-filled performance metrics decreased with increasing AF from ×2 to ×5, with ×7 and ×10 showing slight improvements. This is due to fewer samples being acquired for reconstruction, while the slight improvement is due to a higher percentage of retained center lines.

TABLE IV. STATISTICAL QUANTITATIVE RESULTS FOR BRAIN DATASET

	Zero filled			Reconstructed			Improvements		
AF	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE
×2	30.71	0.8199	0.1652	33.47	0.8992	0.1303	**2.76**	**0.0793**	**0.0349**
×3	25.29	0.6703	0.2806	31.69	0.8776	0.1532	**6.40**	**0.2073**	**0.1274**
×4	23.25	0.6066	0.3535	30.48	0.8563	0.1754	**7.23**	**0.2497**	**0.1781**
×5	22.60	0.5806	0.3822	29.05	0.8318	0.1998	**6.45**	**0.2512**	**0.1824**
×7	22.27	0.5635	0.4021	28.06	0.8128	0.2198	**5.79**	**0.2493**	**0.1823**
×10	22.53	0.5727	0.3945	26.50	0.7787	0.2562	**3.97**	**0.2060**	**0.1383**

TABLE V. STATISTICAL QUANTITATIVE RESULTS FOR KNEE DATASET

	Zero filled			Reconstructed			Improvements		
AF	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE
×2	32.08	0.8261	0.1277	35.42	0.8734	0.0950	**3.34**	**0.0473**	**0.0327**
×3	25.65	0.7137	0.2341	31.88	0.8234	0.1273	**6.23**	**0.1097**	**0.1068**
×4	23.72	0.6555	0.2958	29.76	0.7966	0.1610	**6.04**	**0.1411**	**0.1348**
×5	23.54	0.6302	0.3088	29.26	0.7798	0.1775	**5.72**	**0.1496**	**0.1313**
×7	24.19	0.6099	0.2937	28.84	0.7580	0.1859	**4.65**	**0.1481**	**0.1078**
×10	25.41	0.5978	0.2571	28.08	0.7069	0.1929	**2.67**	**0.1091**	**0.0642**

The reconstruction accuracy decreases significantly as the number of acquired samples decreases. For all the AFs, the CDNet showed significant improvement from the zero-filled results highlighted in bold.

4) Qualitative results

a) Brain data results: A sample reconstructed image for each AF is shown in Fig. 6 for visual assessment, and their performance metrics are given in Table VI. Each reconstructed image is shown alongside its binary mask used to undersample the k-space, its zero-filled image, and its ground truth image, which serves as a baseline for comparison and an error map for visualizing the difference between the reconstructed and ground truth images, as they are imperceptible. The 1s corresponding to white in the binary mask show the position of acquired phase-encoding lines, while the 0's corresponding to black show the position of unacquired phase-encoding lines. The binary masks show some phase-encoding lines retained at the center, with denser sampling near the center and progressively sparser towards the periphery. This prevents severe acquisition artifacts, especially at higher AFs. The acquisition artifacts in zero-filled images increase with increasing AFs, since fewer data are available for reconstruction. The reconstructed and ground-truth images are visually similar. The bright regions indicating discrepancies between the reconstructed and ground-truth images in error maps are more numerous at higher AFs than at lower AFs. This indicates that CDNet's reconstruction accuracy decreases with an increase in AF. However, the acquisition artifacts are absent in all AF sample reconstructed images, demonstrating high-quality CDNet reconstruction results.

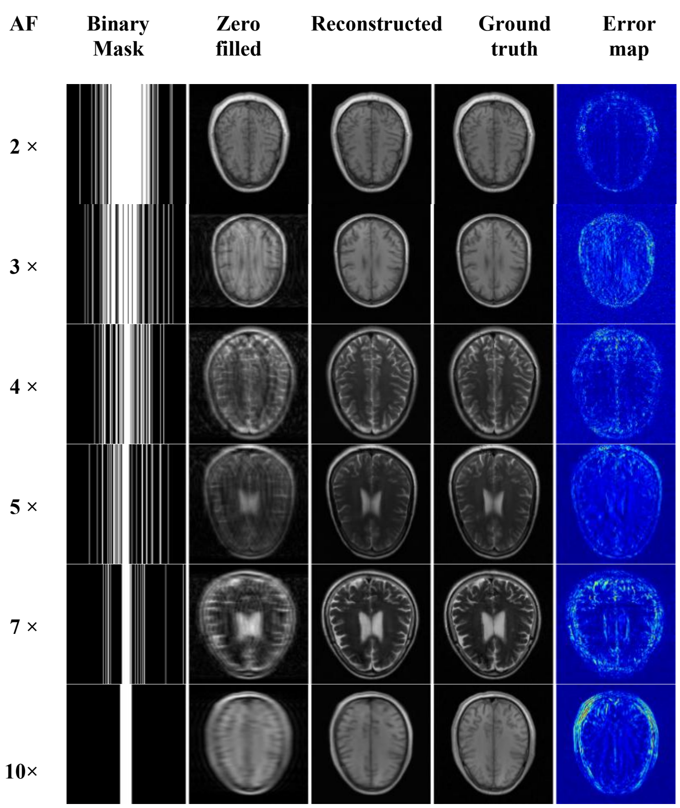

Fig. 7. Sample brain MR images reconstructed at different AFs.

a) Knee data results: A sample reconstructed knee MR image for each AF is shown in Fig. 7 for qualitative evaluation, and their performance metrics are given in Table VII. The reconstructed images are shown alongside their binary mask, their zero-filled and ground truth images, which serve as a baseline for visual comparison, and their error maps, which show differences between the reconstructed images and the ground truth images, as they are visually indistinguishable from the ground truth.

TABLE VI. PERFORMANCE METRICS FOR BRAIN MR IMAGES IN FIG. 6

	Zero filled			Reconstructed		
AF	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE
×2	32.85	0.8984	0.2835	37.56	0.9483	0.2162
×3	23.80	0.6501	0.5022	36.03	0.9328	0.2484
×4	22.23	0.6480	0.5757	32.92	0.9107	0.3111
×5	22.40	0.5888	0.5937	28.84	0.8889	0.4098
×7	20.70	0.5602	0.6228	28.44	0.8686	0.3987
×10	22.24	0.6368	0.5157	27.69	0.8409	0.3769

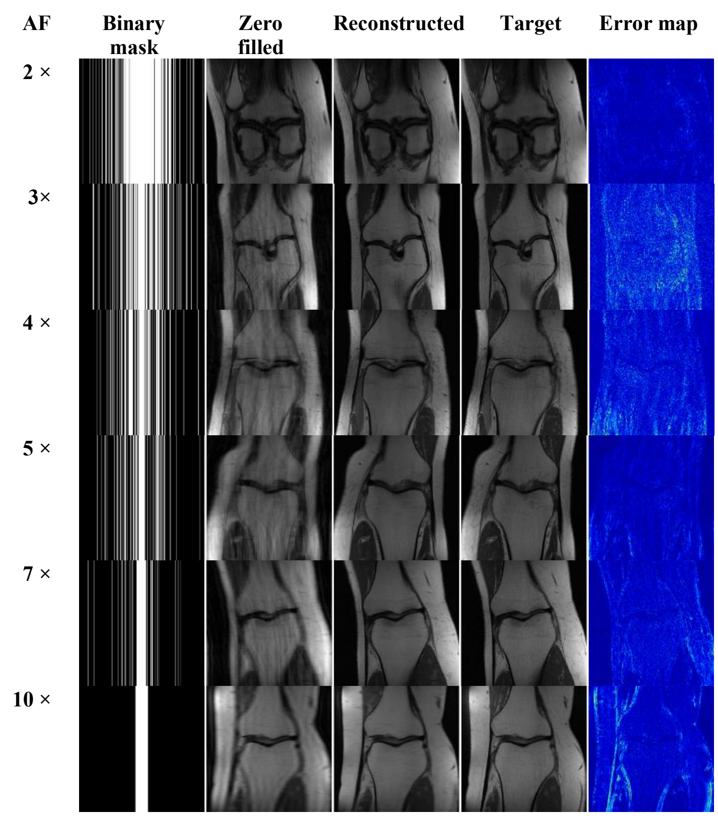

Fig. 8. Sample knee MR images reconstructed at different AFs.

Similar to binary masks in brain data quantitative results, the 1s correspond to white and show positions of acquired phase encoding lines, while the 0s correspond to black and show positions of unacquired measurements.

Some lines are retained at the center to preserve the anatomical structure, while the remaining lines are sampled densely near the center and progressively sparser towards the periphery. This avoids severe acquisition artifacts that may render the reconstructed image unsuitable for diagnosis. The zero-filled images show that acquisition artifacts increase with higher AF as fewer samples are collected for reconstruction. The reconstructed and ground-truth images in each AF are visually similar. However, the error maps show the differences between the reconstructed and ground-truth images, with brighter regions indicating discrepancies between the two images being more at higher AFs and blue regions indicating identical regions dominating in lower AFs. The reconstructed images are of good quality and artifact-free at all AFs, demonstrating good image reconstruction of the proposed CDNet.

TABLE VII. PERFORMANCE METRICS FOR KNEE MR IMAGES IN FIG. 7

	Zero filled			Reconstructed		
AF	PSNR (dB)	SSIM	NRMSE	PSNR (dB)	SSIM	NRMSE
×2	33.10	0.8747	0.2789	37.24	0.9155	0.2198
×3	26.39	0.7339	0.4184	34.28	0.8598	0.2655
×4	23.16	0.7024	0.4950	32.31	0.8253	0.2923
×5	24.02	0.6732	0.4713	32.20	0.8252	0.2943
×7	25.15	0.6511	0.4093	30.61	0.7778	0.2990
×10	25.41	0.6531	0.3854	28.42	0.7373	0.3240

E. Model Comparison

The proposed CDNet was compared to zero-filled results obtained by filling the unacquired phase-encoding lines with zeros and performing an inverse FFT on the zero-filled k-space, and with the results of four other state-of-the-art unrolled architectures, namely a deep cascade of convolutional neural networks (DC-CNN) [43], variational network (VarNet) [44], interpretable optimization-inspired deep network ISTAnet [45], and learned primal–dual reconstruction LPDNet [46]. Their architectures were obtained from the authors' websites and retrained using the same network settings and data as in CDNet for fair comparison. The quantitative results for an AF of ×5 are given in Table VIII, with PSNR, SSIM, and NRMSE reporting average values computed over the entire test brain MR images. The proposed CDNet yielded better statistical results evaluated using the performance metrics than other unrolled networks at a comparable reconstruction time. This demonstrates the superiority of the CDNet over other state-of-the-art unrolled architectures for MR image reconstruction.

TABLE VIII. QUANTITATIVE RESULTS FOR MODEL COMPARISON USING AN AF OF ×5

Method	PSNR (dB)	SSIM	NRMSE	Time (s)
Zero-filled	22.60	0.5806	0.3822	-
DC-CNN	25.64	0.7342	0.2721	0.1184
VarNet	26.25	0.7484	0.2532	0.1257
ISTANet	26.46	0.7575	0.2558	0.1156
LPDNet	28.12	0.7981	0.2213	0.1158
CDNet	**29.05**	**0.8318**	**0.1998**	**0.1885**

Fig. 9 shows a sample brain MR image and its corresponding error map reconstructed using iterative TOS for composite regulation, DC-CNN, VarNet, ISTAnet, LPDNet, and the proposed CDNet for qualitative comparison. The performance metrics of this sample image are given in Table IX. The zero-filled and ground truth images are also shown to act as baselines for visual comparison. When the sample image is reconstructed using the iterative TOS for composite regulation, it results in an image with visible residual acquisition artifacts coexisting with the anatomical structure of the image. The image was also reconstructed in 1.4510s, which is approximately × 7.7 higher than that of CDNet. This high reconstruction time is due to the high number of iterations required for convergence of iterative CS algorithms. The proposed CDNet yielded the best image assessed visually, with edges of fine details seen with ease compared to the other unrolled networks and the iterative TOS. This shows that the proposed CDNet yields superior reconstruction accuracy compared to other deep unrolled networks in MR image reconstruction.

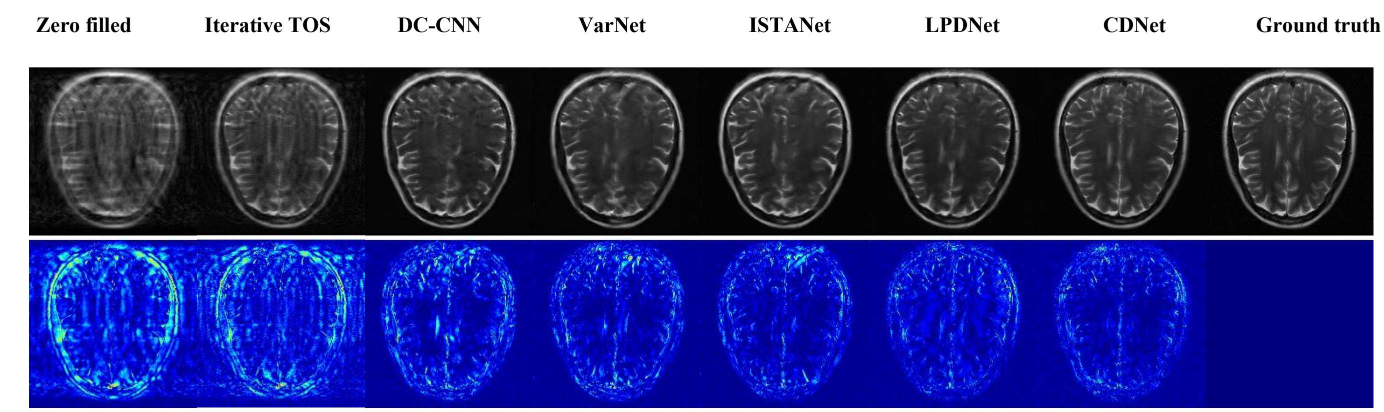

Fig. 9. Sample brain MR image and its corresponding error map reconstructed using comparative methods.

TABLE IX. PERFORMANCE METRICS OF THE SAMPLE BRAIN MR IMAGE IN FIG. 8

Method	PSNR (dB)	SSIM	NRMSE
Zero-filled	21.12	0.5463	0.6238
Iterative TOS	25.61	0.7141	0.4759
DC-CNN	26.48	0.7978	0.4585
VarNet	27.68	0.8313	0.4278
ISTANet	27.98	0.8457	0.4205
LPDNet	29.33	0.8845	0.3891
CDNet	**30.63**	**0.8900**	**0.3609**

V. CONCLUSION

In this work, we introduce a composite denoising network for reconstructing high-quality MR images using fewer measurements acquired at sub-Nyquist rates. In particular, we solve the composite regularization problem in MRI using the iterative three-operator splitting strategy and reconstruct MR images using two priors that promote sparsity in the transform and spatial domains. We then unfolded the iterative algorithm into a composite denoising network that maintains consistent arithmetic architectures with it and employs multiple complex-valued deep learning strategies for capturing different attributes of the MR images. We implemented the CDNet with two complex-valued DL strategies, namely complex-valued CNN and U-Net, for capturing the overall anatomical structure and spatially local features with enhanced global features, respectively. The hyperparameters of the iterative algorithm are learned from data, eliminating the need for empirical selection. Extensive experiments using raw brain and knee k-space from FastMRI, with acceleration factors ranging from ×2 to ×10 was conducted to validate the proposed network. The proposed network is also compared with other state-of-the-art unrolled networks, demonstrating its superiority in reconstruction accuracy at a comparable speed by achieving an average PSNR of 29.05 dB, SSIM of 0.8318, NRMSE of 0.1998, and reconstruction time of 0.1885s for ×5 AF. Despite the promising performance of CDNet, increasing the number of complex-valued DL strategies K to capture different features of MR images increases memory and computational requirements, which may limit its deployment in resource-constrained medical imaging environments. Future work may explore incorporating additional lightweight complex-valued deep learning strategies into the CDNet, including transformers and generative adversarial networks (GANs), among others, to capture more diverse image features and further enhance reconstruction quality. The CDNet addresses the challenges of solving the composite regularization problem using iterative algorithms in an interpretable manner, showcasing significant potential for real-time MRI.

ACKNOWLEDGMENT

The authors are grateful to Pan African University under the African Union Commission for the support provided throughout the research.

REFERENCES

[1] M. I. Doneva, M. Akcakaya and C. Prieto, Magnetic resonance image reconstruction: Theory, Methods and Applications. Academic Press, 2022.

[2] H. Kiragu, E. Mwangi and G. Kamucha, "An Efficacious MRI Sparse Recovery Method Based on Differential Under-Sampling and k-Space Interpolation," *2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON)*, Palermo, Italy, 2020, pp. 382-387, doi: 10.1109/MELECON48756.2020.9140563.

[3] W. Bian and Y. K. Tamilselvam, "A review of Optimization-Based Deep Learning Models for MRI Reconstruction," *AppliedMath,* vol. 4, no. 3, pp. 1098–1127, Sep. 2024, doi: 10.3390/appliedmath4030059.

[4] M. Safari, Z. Eidex, C.-W. Chang, R. L. J. Qiu, and X. Yang, "Advancing MRI reconstruction: A systematic review of deep learning and compressed sensing integration," *Biomedical Signal Processing and Control,* vol. 111, p. 108291, Jan. 2026, doi: 10.1016/j.bspc.2025.108291.

[5] J. Huang *et al*., "Data- and Physics-Driven Deep Learning Based Reconstruction for FAST MRI: Fundamentals and Methodologies," *IEEE Reviews in Biomedical Engineering*, vol. 18, pp. 152–171, Oct. 2024, doi: 10.1109/rbme.2024.3485022.

[6] X. Kui *et al*., "A comprehensive survey on magnetic resonance image reconstruction," *Image and Vision Computing*, vol. 165, p. 105832, Nov. 2025, doi: 10.1016/j.imavis.2025.105832.

[7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006, doi: 10.1109/tit.2006.871582.

[8] S. Villota and E. Inga, "Sparse transform and compressed sensing methods to improve efficiency and quality in magnetic resonance medical imaging," *Sensors,* vol. 25, no. 16, p. 5137, Aug. 2025, doi: 10.3390/s25165137.

[9] J. C. Ye, "Compressed sensing MRI: a review from signal processing perspective," *BMC Biomedical Engineering*, vol. 1, no. 1, p. 8, Mar. 2019, doi: 10.1186/s42490-019-0006-z.

[10] T. Hong, L. Hernandez-Garcia and J. A. Fessler, "A Complex Quasi-Newton Proximal Method for Image Reconstruction in Compressed Sensing MRI," in *IEEE Transactions on Computational Imaging*, vol. 10, pp. 372-384, 2024, doi: 10.1109/TCI.2024.3369404.

[11] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, Aug. 2004, doi: 10.1002/cpa.20042.

[12] A. Beck and M. Teboulle, "A fast iterative Shrinkage-Thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences,* vol. 2, no. 1, pp. 183–202, Jan. 2009, doi: 10.1137/080716542.

[13] J. Yang and Y. Zhang, "Alternating Direction Algorithms for L1-Problems in Compressive Sensing," *SIAM Journal on Scientific Computing,* vol. 33, no. 1, pp. 250–278, Jan. 2011, doi: 10.1137/090777761.

[14] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, Oct. 2007, doi: 10.1002/mrm.21391.

[15] J. Huang, S. Zhang, and D. Metaxas, "Efficient MR image reconstruction for compressed MR imaging," *Medical Image Analysis*, vol. 15, no. 5, pp. 670–679, Jun. 2011, doi: 10.1016/j.media.2011.06.001.

[16] J. Huang, S. Zhang, H. Li, and D. Metaxas, "Composite splitting algorithms for convex optimization," *Computer Vision and Image Understanding,* vol. 115, no. 12, pp. 1610–1622, Sep. 2011, doi: 10.1016/j.cviu.2011.06.011.

[17] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 103–116, Dec. 2008, doi: 10.1002/mrm.21757.

[18] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992, doi: 10.1016/0167-2789(92)90242-f.

[19] J. He, Q. Liu, A. G. Christodoulou, C. Ma, F. Lam, and Z.-P. Liang, "Accelerated High-Dimensional MR imaging with sparse sampling using Low-Rank tensors," *IEEE Transactions on Medical Imaging*, vol. 35, no. 9, pp. 2119–2129, Apr. 2016, doi: 10.1109/tmi.2016.2550204.

[20] Z. Ke, W. Huang, Z. Cui et al., "Learned Low-Rank priors in dynamic

MR Imaging," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3698–3710, Jul. 2021, doi: 10.1109/tmi.2021.3096218.

[21] Y. Zhang, H. Gui, N. Yang, and Y. Hu, "JotlasNet: Joint tensor low-rank and attention-based sparse unrolling network for accelerating dynamic MRI," *Magnetic Resonance Imaging*, vol. 118, p. 110337, Jan. 2025, doi: 10.1016/j.mri.2025.110337.

[22] X. Qiao, W. Li, G. Wang, and Y. Huang, "MCU-Net: A multi-prior collaborative deep unfolding network with gates-controlled spatial attention for accelerated MRI reconstruction," *Neurocomputing*, vol. 633, p. 129771, Feb. 2025, doi: 10.1016/j.neucom.2025.129771.

[23] M. J. Muckley, B. Riemenschneider, A. Radmanesh et al., "Results of the 2020 FastMRI Challenge for Machine Learning MR Image Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2306–2317, Apr. 2021, doi: 10.1109/tmi.2021.3075856.

[24] B. Xin, TS Phan, L Axel , et al. Learned half-quadratic splitting network for magnetic resonance image reconstruction. *arXiv* 2021. preprint arXiv:2112.09760

[25] Y. Zhang, X. Li, W. Li, and Y. Hu, "Deep unrolling shrinkage network for dynamic MR imaging," *2023 IEEE International Conference on Image Processing (ICIP),* pp. 1145–1149, Sep. 2023, doi: 10.1109/icip49359.2023.10223077.

[26] P. Huang, C. Zhang, X. Zhang, X. Li, L. Dong, and L. Ying, "Self-Supervised deep unrolled reconstruction using regularization by denoising," *IEEE Transactions on Medical Imaging*, vol. 43, no. 3, pp. 1203–1213, Nov. 2023, doi: 10.1109/tmi.2023.3332614.

[27] T. Li, Q. Yan, Q. Zou, and Q. Dai, "Gates-Controlled Deep Unfolding network for image compressed sensing," *IEEE Transactions on Computational Imaging*, vol. 10, pp. 103–114, Jan. 2024, doi: 10.1109/tci.2024.3354423.

[28] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional recurrent neural networks for dynamic MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 280–290, Aug. 2018, doi: 10.1109/tmi.2018.2863670.

[29] A. Sriram, J. Zbontar, T. Murrell, A. Defazio, C. L. Zitnick, N. Yakubova, F. Knoll, P. Johnson, End-to-end variational networks for accelerated MRI reconstruction, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23, Springer, 2020, pp. 64–73.

[30] B. Zhou and S. K. Zhou, "DuDoRNet: Learning a Dual-Domain Recurrent Network for Fast MRI Reconstruction With Deep T1 Prior," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020, pp. 4272-4281, doi: 10.1109/CVPR42600.2020.00433

[31] J. Xiang, Y. Dong, and Y. Yang, "FISTA-NET: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Transactions on Medical Imaging,* vol. 40, no. 5, pp. 1329–1339, Jan. 2021, doi: 10.1109/tmi.2021.3054167.

[32] Z. Fabian, B Tinaz, M Soltanolkotabi, "HUMUS-net: hybrid unrolled multi-scale network architecture for accelerated MRI reconstruction" *Adv Neural Inf Process Syst*, 35 , 25306–19, 2022.

[33] Y. Zhang, P. Li, and Y. Hu, "T2LR-Net: An unrolling network learning transformed tensor low-rank prior for dynamic MR image reconstruction," *Computers in Biology and Medicine*, vol. 170, p. 108034, Jan. 2024, doi: 10.1016/j.compbiomed.2024.108034.

[34] M. Vornehm, J. Wetzl, D. Giese, et al., "CineVN: Variational network reconstruction for rapid functional cardiac cine MRI," *Magnetic Resonance in Medicine*, vol. 93, no. 1, pp. 138–150, 2025, doi: 10.1002/mrm.30260.

[35] W. Huang, Z. Ke, Z. Cui et al., "Deep low-Rank plus sparse network for dynamic MR imaging," *Medical Image Analysis*, vol. 73, p. 102190, Jul. 2021, doi: 10.1016/j.media.2021.102190.

[36] D. Davis and W. Yin, "A Three-Operator Splitting Scheme and its Optimization Applications," *Set-Valued and Variational Analysis*, vol. 25, no. 4, pp. 829–858, Jun. 2017, doi: 10.1007/s11228-017-0421-z

[37] H. Wang, M. Fazlyab, S. Chen, and V. M. Preciado, Eds. Robust Convergence Analysis of Three-Operator Splitting, *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2019. doi: 10.1109/ALLERTON.2019.8919695

[38] M. Arjovsky, A. Shah, and Y. Bengio, "Unitary evolution recurrent neural networks," *arXiv (Cornell University)*, Nov. 2015, doi: 10.48550/arxiv.1511.06464.

[39] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Transactions on Computational Imaging,* vol. 3, no. 1, pp. 47–57, Dec. 2016, doi: 10.1109/tci.2016.2644865.

[40] N. Pezzotti, *et al*., "An adaptive intelligence algorithm for undersampled knee MRI reconstruction," *IEEE Access*, vol. 8, pp. 204825–204838, Jan. 2020, doi: 10.1109/access.2020.3034287.

[41] D.P. Kingma and J.L. Ba, "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, Dec 2014.

[42] J. Zbontar, F. Knoll, A. Sriram et al., "FastMRI: an open dataset and benchmarks for Accelerated MRI," *arXiv (Cornell University)*, Jan. 2018, doi: 10.48550/arxiv.1811.08839.

[43] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, Oct. 2017, doi: 10.1109/TMI.2017.2760978

[44] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3055–3071, Nov. 2017, doi: 10.1002/mrm.26977

[45] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable Optimization-Inspired Deep Network for Image Compressive Sensing," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 1828-1837, doi: 10.1109/CVPR.2018.00196.

[46] J. Adler and O. Oktem, "Learned Primal-Dual reconstruction," *IEEE Transactions on Medical Imaging,* vol. 37, no. 6, pp. 1322–1332, Jan. 2018, doi: 10.1109/tmi.2018.2799231.