Compressed Sensing MRI Reconstruction with Co-VeGAN: Complex-Valued Generative Adversarial Network

Bhavya Vasudeva[1]*, Puneesh Deora[1]*, Saumik Bhattacharya[2], Pyari Mohan Pradhan[3]
[1] ISI Kolkata [2] IIT Kharagpur [3] IIT Roorkee

Abstract

Compressed sensing (CS) is extensively used to reduce magnetic resonance imaging (MRI) acquisition time. State-of-the-art deep learning-based methods have proven effective in obtaining fast, high-quality reconstruction of CS-MR images. However, they treat the inherently complex-valued MRI data as real-valued entities by extracting the magnitude content or concatenating the complex-valued data as two real-valued channels for processing. In both cases, the phase content is discarded. To address the fundamental problem of real-valued deep networks, i.e. their inability to process complex-valued data, we propose a complex-valued generative adversarial network (Co-VeGAN) framework, which is the first-of-its-kind generative model exploring the use of complex-valued weights and operations. Further, since real-valued activation functions do not generalize well to the complex-valued space, we propose a novel complex-valued activation function that is sensitive to the input phase and has a learnable profile. Extensive evaluation of the proposed approach[1] on different datasets demonstrates that it significantly outperforms the existing CS-MRI reconstruction techniques.

1. Introduction

Magnetic resonance imaging (MRI) is a frequently used medical imaging modality as it is an excellent non-invasive source of anatomical information. However, it has a considerably long scan time due to sequential acquisition of large volumes of data in the k-space. This can cause significant artifacts due to physiological motion and movement of the patient during the scan and hinder its use in time-critical diagnosis. Compressed sensing (CS) [15] based undersampling can be leveraged to speed up the imaging process. However, it renders the inverse problem ill-posed, making the recovery of high-quality MR images challenging.

Conventional approaches for CS-MRI reconstruction focus extensively on using sparse representations using pre-defined transforms [6, 31, 35], dictionary learning [41, 45], and non-local low-rank regularization [14] to assume prior knowledge on the structure of the MR image. They suffer from long computation time due to the iterative nature of the optimization process. Moreover, the universally applicable sparsifying transforms may not completely capture the fine details present in biological tissues [34].

Deep learning (DL)-based frameworks have enjoyed great success in similar inverse problems like single-image super-resolution [13], denoising [53], etc. These successes and their advantages, such as faster inference (reconstruc-

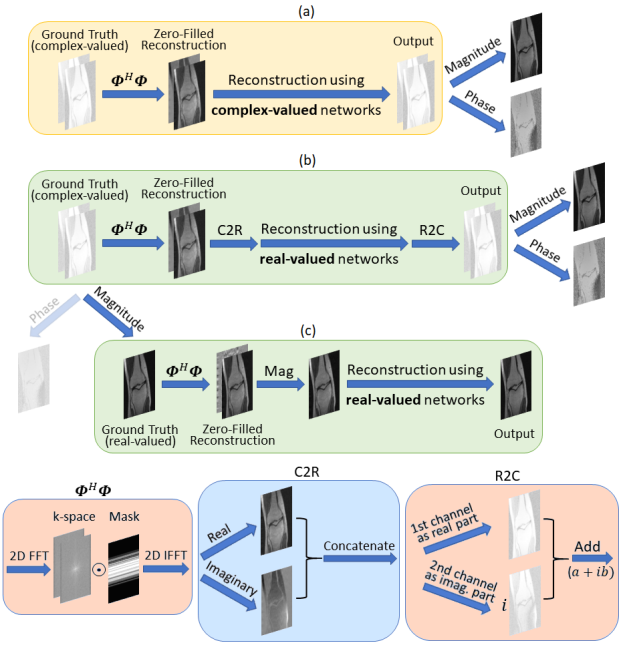

Figure 1. Comparison of the proposed pipeline (a) with the pipelines (b), (c) typically followed by SOTA DL-based methods for CS-MRI reconstruction. Our method uses complex-valued operations at all stages to process the complex-valued MRI data. In contrast, other methods treat it as real-valued data by either concatenating it as two real-valued channels (b) or by using the magnitude images (c). Φ is the undersampling matrix, and $\Phi^H\Phi$ is the matrix to obtain the ZFR. The two channels shown in the ground truth, ZFR, k-space, output, represent the complex-valued data.

*Equal contribution; Work done while at IIT Roorkee.
[1]Code available at `https://github.com/estija/Co-VeGAN`

tion) time and avoidance of an explicit assumption of sparsity, have prompted their use for accelerating MR imaging [1, 4, 7, 29, 36, 54, 55]. Recent advancements of generative adversarial networks (GANs) [19, 28] have inspired their use for CS-MRI reconstruction [7, 12, 36, 40, 54].

Although state-of-the-art (SOTA) DL-based approaches have significantly improved the reconstruction quality, they treat the complex-valued MRI data as real-valued data. This is done either by concatenating the real and imaginary parts as two real-valued channels [1, 21, 36, 40], as shown in Fig. 1(b), or by using the magnitude images [26,29], as shown in Fig. 1(c). Deep neural networks (DNNs) with real-valued weights are then used to process such data. The use of the latter pipeline reconstructs only magnitude images and loses out on phase information. In other works [21, 36, 40], the use of the former pipeline essentially decouples the real and imaginary parts of the complex-valued MRI data, thus damaging the phase information when it is processed and leading to unsatisfactory recovery. Even methods like [1] which use a combination of NNs and model-based approaches suffer from similar limitations. Many applications utilize the phase content, such as phase-contrast imaging, quantifying velocity, and measuring blood and volume flow [18], so it is crucial to recover it faithfully. However, the aforementioned methods fail to do so.

Moreover, when the dataset itself contains only magnitude images, these methods and others [12, 54] use a pipeline similar to the one in Fig. 1(c). However, since the zero-filled reconstruction (ZFR), which is used as the input, is complex-valued, processing it with real-valued networks fails to utilize its complex algebraic structure.

Motivation: As demonstrated in Fig. 1, the inability of real-valued DL-based methods to process complex-valued data causes issues in all the aforementioned cases. Leveraging complex-valued NNs to process the inherently complex-valued MRI data (Fig. 1(a)) naturally appears to be the right direction to proceed.

Other motivations for using complex-valued networks stem from their benefits over real-valued networks. Apart from having biological inspiration and significance [42], they offer increased representational capacity and more stability [9,48]. They exhibit faster learning with better robustness to noise [3, 9, 52] and preserve the phase information, which encodes fine structural details of an image [39, 48]. Complex-valued operations can be performed with simple optimization techniques [38] without sacrificing the generalization ability of the network [23].

We seek to exploit the benefits of complex-valued NNs, and considering the capabilities of adversarial learning approaches, we propose a novel complex-valued GAN (Co-VeGAN), which, to the best of our knowledge, is the *first-of-its-kind generative model* leveraging complex-valued weights and operations.

Activation function: Since activation functions in DNNs allow them to learn highly complex functions due to their nonlinear nature, we study their behaviour in the complex domain. Real-valued activations have been studied extensively, with sigmoid, rectified linear unit (ReLU), leaky ReLU, parametric ReLU [22], etc. being widely used. However, they do not prove to be as effective when their complex-valued equivalents are considered, thus producing a requirement for complex-valued activations, only a few of which have been proposed [3, 49]. Their poor transferability to the complex domain can be attributed to their weak or limited sensitivity to the changes in the input phase. To address this, we formulate a novel complex-valued activation by leveraging weighted sinusoid functions.

Loss function: Another limitation of existing frameworks is the widespread use of pixelwise l_2 or l_1 loss functions. This efficiently reconstructs the low-frequency components of the image but often fails to generate the mid and high-frequency information. We introduce a novel Gaussian-weighted wavelet loss function to improve the reconstruction of mid-frequency components of the image.

Contributions: We believe that this work will provide the basis for future works deploying complex-valued GAN-based frameworks for CS-MRI reconstruction as well as other applications, which use complex-valued data. It will also inspire further research on the behaviour of complex-valued activation functions for DL-based applications in general. The major contributions of this work are:

1. We propose a novel Co-VeGAN framework for CS-MRI reconstruction, introducing complex-valued weights and operations in a generative model for the first time. In contrast to the SOTA DL-based methods, which use real-valued operations, our framework exploits the complex algebraic structure of MRI, leading to accurate recovery of both magnitude and phase content.

2. We introduce a novel activation function designed specifically for complex-valued entities. It overcomes the limitations of the existing activation functions, and our study on various activations for CS-MRI reconstruction and their analyses in the complex domain is first-of-its-kind.

3. To overcome the limitations posed by l_2 or l_1 norm-based losses in reconstruction problems, we introduce a novel variant of the wavelet loss function.

4. We also propose a novel dense U-net architecture leveraging dense connections within the contracting and expanding paths to allow feature reuse between layers.

5. We perform comprehensive ablation studies and evaluate our approach on three different datasets. We also analyze the performance of equivalent real-valued networks and show that our framework provides superior reconstruction using significantly fewer trainable parameters.

1.1. Related Work

DL-based Methods for CS-MRI Reconstruction: Yang *et al.* [55] used the alternating direction method of multipliers (ADMM) algorithm [8] to train their deep network. Sun *et al.* [47] proposed DL models for multi-contrast CS-MRI reconstruction using parallel networks and feature sharing strategies. Liu *et al.* [33] introduced a framework combining the learning strategy of data-driven NNs with the iterative solver providing domain knowledge. Bora *et al.* [7] have shown that pre-trained generative models like variational autoencoders and GANs [19] can be used to recover CS signals without assuming sparsity. Yang *et al.* [54] proposed a U-net [43] based generator, following a refinement learning-based approach, with mean squared error (MSE) and perceptual loss to reconstruct the images. Quan *et al.* [40] proposed a fully residual NN using addition-based skip connections. They used cyclic loss for data consistency constraints in training to achieve better reconstruction quality. Deora *et al.* [12] proposed a U-net based generator with a patch-based discriminator [28] to perform the reconstruction. Along with mean absolute error (MAE) and structural similarity (SSIM) [56], they used Wasserstein loss [2] to improve adversarial learning. Mardani *et al.* [36] introduced an affine projection operator in-between the generator and the discriminator to improve the data consistency in the reconstructed images. Hammernick *et al.* [21] split the complex-valued data, and train a variational network, where they learn separate filters for real and imaginary data.

Complex-valued NNs: Trabelsi *et al.* [48] first incorporated complex-valued operations in DNNs. Although complex-valued NNs have several striking benefits, only a few works have explored their use in applications like vision [48, 58], parallel MR imaging [50], MRI fingerprinting [10, 49], and audio-related tasks [48]. Sun *et al.* [46] used complex-valued residual NNs to reconstruct undersampled MR images in both image and frequency domain. Dedmari *et al.* [11] proposed a complex dense fully convolutional NN for CS-MRI reconstruction. El-Rewaidy *et al.* [17] proposed radial batch normalization (BN) to avoid distortion in phase and magnitude for reconstructing 3D MR images.

Although operations like convolution, backpropagation, and BN [27] have been extended to the complex domain, only a few complex-valued activations have been proposed [3, 10, 49]. Most of these address the poor transferability of real-valued activations but fail to account for the variations in the input phase. Virtue *et al.* [49] propose cardioid, which is sensitive to the input phase and preserves it at the output. However, such phase preservation can compromise the flexibility of the activation.

2. Methodology

The acquisition model of CS-MRI can be described as:

$$\mathbf{y} = \mathbf{\Phi x} + \boldsymbol{\zeta} = \mathbf{UFx} + \boldsymbol{\zeta}, \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^{K^2}$ is the desired image in vector form, $\mathbf{y} \in \mathbb{C}^M$ denotes the observed data vector, $\boldsymbol{\zeta} \in \mathbb{C}^M$ captures the noise. $\mathbf{F} \in \mathbb{C}^{K^2 \times K^2}$ is the matrix to compute the 2D Fourier transform, $\mathbf{U} \in \mathbb{R}^{M \times K^2}$ is the matrix for undersampling. Given an observation $\mathbf{y}$, the aim of reconstruction is to recover $\mathbf{x}$ in the presence of a non-zero $\boldsymbol{\zeta}$.

We attempt to recover $\mathbf{x}$ by using a GAN model. A GAN comprises two networks, namely a generator and a discriminator. To generate images that are similar to the samples of the distribution of true data $\mathbf{y}_t$, the generator G attempts to map an input vector $\mathbf{z}$ to the output $G(\mathbf{z})$. On the other hand, the aim of the discriminator D is to classify the generated samples $G(\mathbf{z})$ and the samples from the distribution of $\mathbf{y}_t$. As mentioned earlier, for the CS-MRI reconstruction problem, both the observation $\mathbf{y}$ and the desired image $\mathbf{x}$ are complex-valued, whereas the parameters of a GAN are real-valued in existing methods. Thus, we propose a Co-VeGAN framework, which is described below.

2.1. Complex-valued GAN

We propose a complex-valued GAN which consists of a complex-valued $G : \mathbb{C}^{K \times K} \rightarrow \mathbb{C}^{K \times K}$, and a real-valued $D : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^{1 \times 1}$. Although the parameters of both G and D can be complex-valued, we opt to use a real-valued discriminator, mainly because doing so enables it to take the visual soundness of the magnitude images into account while discriminating the generated and actual images. The use of a real-valued discriminator also makes our framework suitable for real-valued datasets, where the ground truth (GT) is real-valued. To constrain the phase content of the generated output to be similar to that of the GT, we use a pixelwise loss described in Section 2.4.

2.2. Complex-valued operations

Trabelsi *et al.* [48] proposed complex-valued operations for DL. We discuss activation in detail and the rest in brief here. Further details can be found in the supplementary.

Convolution: The complex-valued convolution of kernel $\mathbf{W}$ with feature maps $\mathbf{F}$ is given by $\mathbf{A} = (\mathbf{W_R} * \mathbf{F_R} - \mathbf{W_I} * \mathbf{F_I}) + i(\mathbf{W_R} * \mathbf{F_I} + \mathbf{W_I} * \mathbf{F_R})$, where $*$ denotes convolution, i denotes the imaginary unit, and subscripts $\mathbf{R}$ and $\mathbf{I}$ denote the real and imaginary parts of the complex-valued entities, respectively.

Backpropagation: For a real-valued loss function f, which is non-holomorphic and differentiable with respect to its real and imaginary parts [30], the gradient descent update of an l-dimensional weight vector $\mathbf{w}$ is given by $\mathbf{w}_{t+1} = \mathbf{w}_t - \rho \nabla_{\bar{\mathbf{w}}} f(\mathbf{w})$, where ρ is the learning rate (LR), $\bar{(\cdot)}$ denotes the complex conjugate, and the gradient of f is:

$$\nabla_{\bar{\mathbf{w}}} f(\mathbf{w}) = \Big[\frac{\partial f}{\partial \bar{w}_1} \cdots \frac{\partial f}{\partial \bar{w}_l}\Big]^T; \frac{\partial f}{\partial \bar{w}} = \frac{1}{2}\Big(\frac{\partial f}{\partial w_R} + i\frac{\partial f}{\partial w_I}\Big). \tag{2}$$

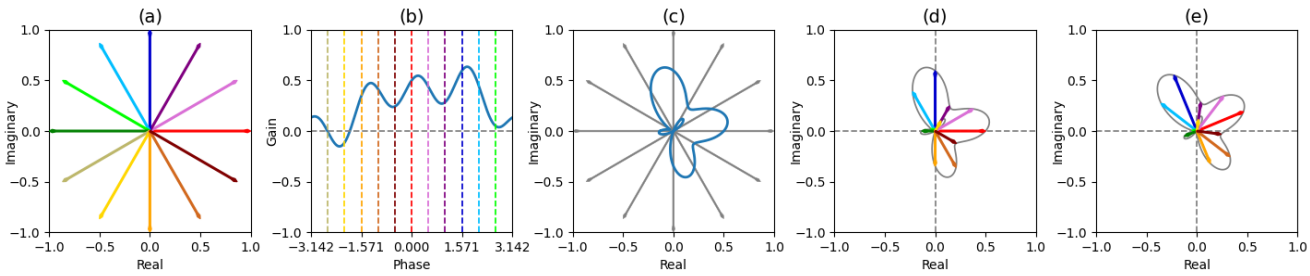

Figure 2. An example of the proposed activation PC-WSS, with parameters $[w_0, w_1, w_2, \theta_0, \theta_1, \theta_2] = [0.08, -0.04, 0.06, 0.6, 0.4, 0.2]$. (a) A set of complex-valued inputs, (b) plot of gain vs input phase, (c) plot of gain in the complex space, (d) outputs of PC-WSS with ϕ = 0, and (e) outputs with $\phi = \frac{\pi}{8}$. We see that the gain varies with the phase (b,c) depending on the parameters (w_i, θ_i) and impacts how the inputs (a) are scaled to get the corresponding outputs (d). Further, ϕ causes phase changes in the outputs (e).

Batch Normalization: In complex BN (CBN), the input is first whitened: $\mathbf{x}_{std} = \mathbf{B}^{-\frac{1}{2}}(\mathbf{x} - \mathbb{E}[\mathbf{x}])$, where $\mathbb{E}$ denotes expectation operator and $\mathbf{B}$ is the covariance matrix:

$$\mathbf{B} = \begin{bmatrix} \text{Cov}(\mathbf{x_R}, \mathbf{x_R}) & \text{Cov}(\mathbf{x_R}, \mathbf{x_I}) \\ \text{Cov}(\mathbf{x_I}, \mathbf{x_R}) & \text{Cov}(\mathbf{x_I}, \mathbf{x_I}) \end{bmatrix}, \tag{3}$$

and learnable parameters $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ are then used to scale and shift it: $\mathbf{x}_{BN} = \boldsymbol{\gamma}\mathbf{x}_{std} + \boldsymbol{\beta}$.

Activation: One common class of complex-valued activation functions [3] is obtained by applying same nonlinearity separately on the real and imaginary parts. The complex equivalent of ReLU ($\mathbb{C}$ReLU) is given by:

$$\mathbb{C}ReLU(a) = ReLU(a_R) + iReLU(a_I), \tag{4}$$

where a is a complex-valued input. Similarly, we can also obtain the complex equivalent of parametric ReLU [22] ($\mathbb{C}$PReLU). There is another class of complex activations where the nonlinearity is applied on the magnitude $|a|$ and phase $\angle a$ of a. One such activation is zReLU [20], which allows the input to pass only if it lies in the first quadrant.

However, all the aforementioned activation functions are weakly sensitive to the changes in the input phase, which is not only important from the biological perspective but also crucial for the complex representation. Virtue *et al.* [49] proposed the cardioid activation, where the input phase is preserved at the output, and the output magnitude is sensitive to the input phase. It is given by:

$$Cardioid(a) = g(\angle a)a = \frac{1}{2}(1 + cos(\angle a))a, \tag{5}$$

where $g(\cdot)$ denotes the gain, which is positive-valued for phase preservation. However, phase preservation reduces the flexibility of the activation because phase change from the input to the output can be viewed as the delay of a neuron. Moreover, since the phase is modified at the CBN layer, its preservation at the activation layer might not prove useful. Also, the gain of cardioid is fixed, and its shape is not learnable, due to which it favours positive real-valued inputs over negative ones, as $g(0) = 1$ and $g(\pi) = 0$, and real-valued inputs over purely imaginary inputs, as $g(\frac{\pi}{2}) = 0.5$.

Proposed activation: We propose a novel activation function, where each neuron can introduce a phase change (PC) ϕ in the input. This can also be viewed as allowing coupling between a_R and a_I. To allow the output magnitude to be sensitive to the input phase, any continuous function of the input phase can be used as the gain. This requires the gain to be periodic with period 2π. We propose the use of a weighted sum of sinusoids (WSS) in this work to provide flexibility and maintain periodicity. The proposed PC-WSS is formulated as:

$$PC\text{-}WSS(a) = \frac{\sum_{p=0}^{P_S-1} w_p\{1 + cos\,(2^p(\angle a - \theta_p))\}}{2\sum_{p=0}^{P_S-1}|w_p| + \epsilon} ae^{i\phi}, \tag{6}$$

where w_p, θ_p and ϕ are trainable parameters, P_S denotes the number of sinusoids, and ϵ is a small constant to avoid division by 0. The use of trainable parameters w_p and θ_p allow the framework to learn suitable gain functions. The number of extra parameters introduced by PC-WSS is very small since, for a particular channel of a layer, the weights are shared. The numerator is normalized by the sum of the absolute value of the weights to bound the gain in $[-1, 1]$ and avoid exploding gradients. To control the trade-off between the flexibility and the increase in the number of trainable parameters, P_S is set as 3 in this study. Fig. 2 illustrates an example of PC-WSS[2] to demonstrate its flexible learnable shape and that it can cause changes in the input phase.

To demonstrate the importance of allowing negative gain and using ϕ in PC-WSS, we consider two of its variants. In the first case, ϕ is set as 0, and the gain is bounded in $[0, 1]$. This phase preserving (PP) equivalent of PC-WSS is named PP-WSS. When $w_0 = 1$ and the rest of the parameters are set as 0, PP-WSS is reduced to cardioid. To highlight the importance of ϕ, we consider another variant, where negative gain is allowed, as in PC-WSS, but ϕ is set as 0. This

[2]To further analyze how the trainable parameters influence the activation function learned, the following parametric graph can be explored: `https://www.desmos.com/calculator/erhfrn3gkv`.

tangent inverse ($tan^{-1}(\frac{a_I}{a_R})$) preserving (TIP) equivalent of PC-WSS is named TIP-WSS.

2.3. Network Architecture

The generator architecture used in the proposed framework is based on a U-net architecture. The left side is a contracting path, where each step involves creating downsampled feature maps using a convolutional layer with stride 2, followed by CBN and activation. The right side is an expanding path, where each step consists of upsampling (by a factor of two), a convolutional layer to create new feature maps, followed by CBN and activation. The features from the contracting path are concatenated to the features in the expanding path to provide richer context about low-level features for superior reconstruction.

We propose the use of dense connections between the steps (layers) within the contracting and expanding paths. These dense connections [24] improve the information flow between the layers and encourage feature reuse. Since the feature maps at various layers are not of the same size, average pooling and upsampling (with bilinear interpolation) operations have been introduced in the dense connections between the layers of the contracting and expanding paths, respectively. However, changing the size of the feature maps by a factor greater than r (less than size $\frac{K}{2^r} \times \frac{K}{2^r}$) increases the computational and memory requirement and reduces the quality of information available to the subsequent layers. r is set as 3 in this work.

Further, residual-in-residual dense blocks (RRDBs) [51] are incorporated at the lowest layer of the generator, where feature maps of size $\frac{K}{32} \times \frac{K}{32}$ are present. Each block uses residual learning across each dense block and a group of three dense blocks. At both levels, the residuals are scaled by α before being added to the identity mapping. The RRDBs make identity mappings easier to learn through residual connections and rich information accessible to deeper layers through dense connections. At the output of the generator, a hyperbolic tangent activation is applied[3]. A diagram can be found in the supplementary.

The discriminator architecture is based on a standard CNN with 11 convolutional layers, each of which is followed by BN and leaky ReLU activation. We use a patch-based discriminator to increase the focus on the reconstruction of high-frequency content. It scores each patch of the input image separately and uses their mean as the output.

2.4. Training Losses

Adversarial Loss: To constrain the generator to produce the MR image corresponding to the samples acquired in the k-space, it is conditioned [37] over the ZFR given by:

$$\mathbf{x_u} = \mathbf{\Phi}^H \mathbf{y} = \mathbf{F}^H \mathbf{U}^H \mathbf{y}, \tag{7}$$

where $\mathbf{x_u} \in \mathbb{C}^{K^2}$, H is the Hermitian operator. A Wasserstein distance-based loss [2] is used to obtain the solution to the minimax game between G and D, as it leads to more stable training than vanilla GAN. This is formulated as:

$$\begin{aligned} \min_G \max_D L_{GAN} = \mathbb{E}_{\mathbf{x}\sim p_x(\mathbf{x})}[D(|\mathbf{x}|)] \\ - \mathbb{E}_{\mathbf{x_u}\sim p_{x_u}(\mathbf{x_u})}[D(|G(\mathbf{x_u})|)], \end{aligned} \tag{8}$$

where $p_x(\mathbf{x})$ is the distribution of the GT images, and $p_{x_u}(\mathbf{x_u})$ is the distribution of the aliased ZFR images.

To solve this, an alternating process of updating G once and D n_D times is followed. To enforce the Lipschitz constraint on D, weight clipping is applied [2].

Content Loss: To bring the reconstructed output closer to the corresponding GT image, other losses are required. First, we incorporate an MAE-based loss, $L_{\ell_1} = \mathbb{E}[\|G(\mathbf{x_u}) - \mathbf{x}\|_1]$, where $\|\cdot\|_1$ denotes the ℓ_1 norm. The ℓ_1 norm is preferred to the ℓ_2 norm which can lead to overly smooth and blurry reconstruction. Also, employing MAE in the complex domain helps in accurate phase as well as magnitude reconstruction.

Structural Similarity Loss: To improve the perceptual quality of the reconstructed MR image and preserve the structural details, a mean SSIM (mSSIM) based loss is incorporated. It maximizes the patch-wise SSIM between the magnitude images of reconstructed output and the corresponding GT by minimizing the following expression:

$$L_{mSSIM} = 1 - \mathbb{E}\left[\frac{1}{P}\sum_{p=1}^{P} SSIM(|G_p(\mathbf{x_u})|, |\mathbf{x}_p|)\right],$$

where P denotes the number of patches in the image, and SSIM is calculated as follows:

$$SSIM(\mathbf{M}, \mathbf{N}) = \frac{2\mu_M\mu_N + \epsilon_1}{\mu_M^2 + \mu_N^2 + \epsilon_1} \frac{2\sigma_{MN} + \epsilon_2}{\sigma_M^2 + \sigma_N^2 + \epsilon_2},$$

where $\mathbf{M}$, $\mathbf{N}$ are two image patches, μ_M, μ_N denote their means, σ_M^2, σ_N^2 denote their variances, σ_{MN} denotes their covariance, ϵ_1, ϵ_2 are slack values to avoid division by 0.

Wavelet Loss: To further enhance the textural details in the generated image, a weighted version of MAE in the wavelet domain is incorporated. This is inspired by another inverse problem (super-resolution) [25]. To decompose the image into sets of wavelet coefficients $\mathbf{C}$, which are equal in size, and correspond to even division of bands in the frequency domain, the wavelet packet decomposition is applied. For an r_w level decomposition which produces $P_w = 4^{r_w}$ sets of wavelet coefficients of size $K_w \times K_w$ with $K_w = \frac{K}{\sqrt{P_w}}$, the wavelet loss is formulated as:

[3]For real-valued data, we apply three more operations. First, the output range (of real and imaginary parts) is changed to [0, 1]. Then, the absolute value is obtained (which lies in [0,$\sqrt{2}$]), and scaled back to [-1, 1].

Table 1. Quantitative results for ablation study of the proposed model

Network Settings	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}
Complex-valued GAN	✗	✓	✓	✓	✓	✗
RRDBs	✗	✗	✓	✓	✓	✓
Dense U-net	✗	✗	✗	✓	✓	✓
Wavelet loss	✗	✗	✗	✗	✓	✓
Generator parameters	2M	1.2M	1.5M	1.7M	1.7M	3.5M
PSNR (dB) / mSSIM	39.640 / 0.9823	40.048 / 0.9866	41.418 / 0.9879	43.798 / 0.9902	**45.044 / 0.9919**	42.864 / 0.9858

Table 2. Quantitative comparison of various activation functions

Activation	zReLU	$\mathbb{C}$ReLU	$\mathbb{C}$PReLU	Cardioid	PP-WSS	TIP-WSS	PC-WSS
PSNR (dB) / mSSIM	35.991 / 0.9690	45.044 / 0.9919	45.165 / 0.9920	45.128 / 0.9919	45.066 / 0.9919	45.429 / 0.9925	**45.678 / 0.9927**

$$L_{wvt} = \frac{1}{P_w K_w^2} \sum_{p=1}^{P_w} \gamma_p \left[\sum_{i,j=1}^{K_w} |\mathbf{C}^p_{|G(\mathbf{x_u})|}(i,j) - \mathbf{C}^p_{|\mathbf{x}|}(i,j)| \right],$$

where γ_p denotes the weight of the p^{th} set of coefficients. Since MAE and mSSIM focus more on preserving low and high-frequency content, respectively, higher weights are assigned to the wavelet coefficients corresponding to the band-pass components to improve their reconstruction. To do so, they are set according to the normalized probability density function of a Gaussian distribution with mean $\frac{(P_w-1)}{2}$ and variance σ_w^2. r_w is set as 3 and σ_w^2 as 12.5.

Overall Loss: The overall loss L, used to train G, is formulated as a weighted sum of the losses presented above:

$$L = \lambda_1 L_{GAN} + \lambda_2 L_{\ell_1} + \lambda_3 L_{mSSIM} + \lambda_4 L_{wvt}. \quad (9)$$

In this work, $\lambda_1 = 0.01$, $\lambda_2 = 20$, $\lambda_3 = 1$, and $\lambda_4 = 100$.

Details of training settings can be found in the supplementary. After training, a single forward pass through the trained generator is used to obtain the reconstructed image.

3. Results and Discussion

3.1. Datasets

We evaluate our models on three datasets, namely MICCAI 2013 grand challenge dataset [32] from which we use T_1 weighted MR images of the brain, MRNet dataset [5] from which we use coronal images of the knee, and fastMRI dataset [57] from which we use coronal proton density-weighted single-coil images of the knee. The fastMRI dataset contains complex-valued images, which are cropped to size 320×320. The other two datasets provide only real-valued images of size 256×256. To improve the robustness to noise, we apply data augmentation by adding synthetic 10% and 20% complex Gaussian noise in the Fourier space [12], before undersampling. The proportion of noise-free and noisy images is kept the same as in [12]. For the MICCAI 2013 dataset, 20 787 images taken from the train set are used for training, and 2000 images from the test set are used for testing. For the other two datasets, 12 500 images are randomly chosen for training, and 2000 non-overlapping images are chosen for testing. The testing is carried out for three sets: noise-free images (*i.e.* images without synthetic noise), images with 10% noise, and images with 20% noise. To obtain the undersampled data, we use 30% 1D Gaussian (1D-G) or Cartesian sampling[4].

3.2. Ablation Studies

Table 1 shows the results for the ablation study of the proposed model[5]. The first case demonstrates a real-valued GAN model, which consists of a U-net based generator without RRDBs, without the dense connections, with $\mathbb{C}$ReLU activation and with $\lambda_4 = 0$ in (9). In the following cases, the effect of adding complex-valued operations, RRDBs, dense connections, and L_{wvt} (by setting $\lambda_4 = 100$) are observed. Evidently, each step results in significant improvement in peak signal-to-noise ratio (PSNR) and mSSIM. It is noteworthy that despite a 40% decrease in the number of generator parameters, the complex-valued equivalent (2^{nd} setting) outperforms the real-valued GAN model presented as the 1^{st} setting. To further highlight the usefulness of complex-valued representations, the real-valued version of the fifth case is considered by doubling the number of output channels (details in supplementary). Despite having more than twice the number of trainable parameters, this model significantly underperforms as compared to its complex-valued counterpart (5^{th} setting). For the rest of the results, we use the 5^{th} network settings.

3.3. Activation Functions

Table 2 shows the quantitative results for comparing various activation functions[6]. It is observed that zReLU, which only allows the inputs lying in the first quadrant to pass, has the worst performance. $\mathbb{C}$PReLU outperforms $\mathbb{C}$ReLU and zReLU. This may be because, unlike these activations, $\mathbb{C}$PReLU does not suppress input information strongly for non-zero values of the slope parameter. Moreover, it has an adaptable shape, allowing it to bypass issues like dead neurons, commonly observed in $\mathbb{C}$ReLU. It is observed that

[4]The results for other sampling ratios (10%, 20%, 30%) and patterns (1D-G, radial, spiral) can be found in the supplementary.

[5]Qualitative results are shown in the supplementary.

[6]Qualitative results are shown in the supplementary.

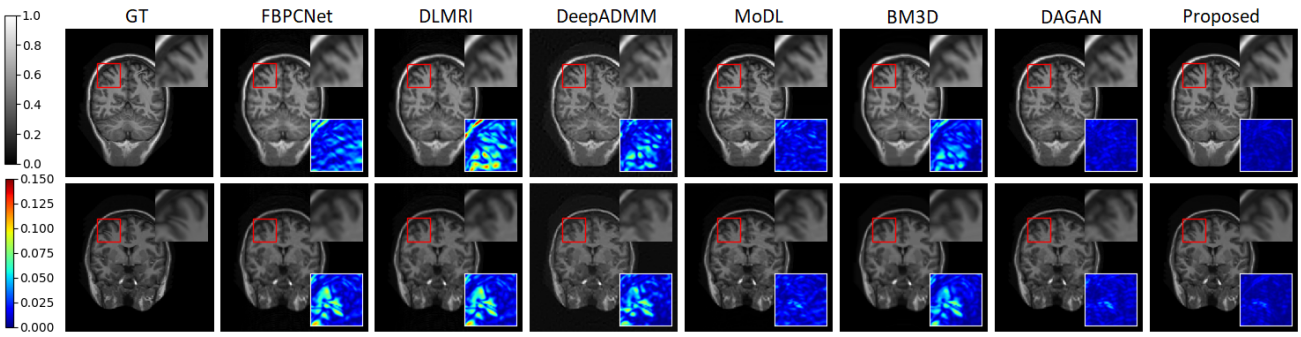

Figure 3. Qualitative results and comparison of the proposed method for two images taken from the MICCAI 2013 dataset. Inset- top right: the ZiR enclosed by the red box, bottom right: absolute difference between the ZiR and its corresponding GT. Best viewed in 200% zoom.

cardioid performs better than these activations, as it is sensitive to phase, and PP-WSS obtains a similar performance as cardioid. However, these phase preserving activation functions slightly underperform as compared to $\mathbb{C}$PReLU, which does not have phase preservation capabilities. Further, when the possibility of a negative gain is introduced in PP-WSS, the resulting TIP-WSS obtains a significantly better performance. This indicates that if phase change is allowed, the flexibility thus produced leads to a boost in performance. The same is true for the proposed PC-WSS after the parameter ϕ is introduced. This experiment not only highlights the superiority of PC-WSS but also shows that phase preserving activations prove less effective in a complex GAN framework. The superior performance of TIP-WSS and PC-WSS over $\mathbb{C}$PReLU can be attributed to higher flexibility and sensitivity to the input phase. The rest of the results are reported by using PC-WSS activation[7].

3.4. Comparison with State-of-the-art

MICCAI 2013 Dataset: Fig. 3 demonstrates the qualitative results[8] of our final model and comparison with SOTA methods. We used the publicly available official implementations of the SOTA methods to analyze their performances on the test set. As seen in the reconstructed results for the first image, methods like FBPCNet, DLMRI, and BM3D are unable to properly reconstruct the high-frequency content (edges) in the zoomed-in region (ZiR). Also, most of the SOTA methods obtain blurry reconstruction results, which might be due to the use of l_2 loss. This can be clearly seen by the quality of reconstruction of the fine details in the ZiR of the second image. FBPCNet, DLMRI, DeepADMM, and BM3D obtain an overly smooth low-quality reconstruction, which appears to be slightly out of focus. Moreover, we can also observe subtle artifacts in the background of the ZiR of the first image. On close inspection of the lower half of the first image, we can see that FBPCNet, DLMRI, DeepADMM produce outputs with aliasing artifacts. The reconstructed images produced by DAGAN, although significantly better than other methods, do not capture subtle details. The visual representations of the proposed scheme show that it preserves the fine structural details and most closely aligns with the GT.

Table 3. Quantitative results and comparison with previous methods using MICCAI 2013 dataset

Method	Noise-free images	10% noise level	20% noise level
	PSNR (dB) / mSSIM		
FBPCNet [29]	35.996 / 0.8655	34.025 / 0.6011	31.421 / 0.4226
DLMRI [41]	37.405 / 0.8732	34.144 / 0.6140	31.564 / 0.4346
DeepADMM [55]	41.545 / 0.8946	39.078 / 0.8105	35.373 / 0.6000
MoDL [1]	42.383 / 0.9760	40.204 / 0.9481	36.844 / 0.8572
BM3D [16]	42.521 / 0.9764	37.836 / 0.7317	33.657 / 0.4947
DAGAN [54]	43.329 / 0.9860	**42.006** / 0.9814	**39.160** / 0.9619
Proposed	**45.678** / **0.9927**	41.809 / **0.9838**	39.083 / **0.9718**

Table 3 illustrates the quantitative comparison of the proposed method with the six SOTA approaches. It can be observed that there is an appreciable boost in both PSNR and mSSIM values obtained by our approach. On comparing the results for images with 10% and 20% noise, it is observed that some methods like FBPCNet, DLMRI, and BM3D experience a steep decline in the reconstruction quality, indicating a lack of robustness. Augmenting the training data with noisy images for our method increases its robustness towards noise. The values obtained by our method for noisy images are significantly better than the other methods, but they are comparable to DAGAN. This might be due to the high complexity of the proposed model, which allows superior reconstruction quality in the noise-free case at the cost of an increase in the sensitivity to the noise-level [44].

MRNet Dataset: To further analyze the effectiveness of our method compared to SOTA methods, we perform another set of experiments using the MRNet dataset, shown in Table 4[9]. We can see that the proposed method significantly outperforms both MoDL and DAGAN for noise-free and noisy images. The proposed method obtains better results for images with 20% noise as compared to those obtained by MoDL and DAGAN for noise-free images.

[7] Plots of the training losses, comparison of runtime with other methods, and ablation study of the loss function can be found in the supplementary.

[8] A high-resolution version of this figure is shown in the supplementary.

[9] Qualitative results are shown in the supplementary.

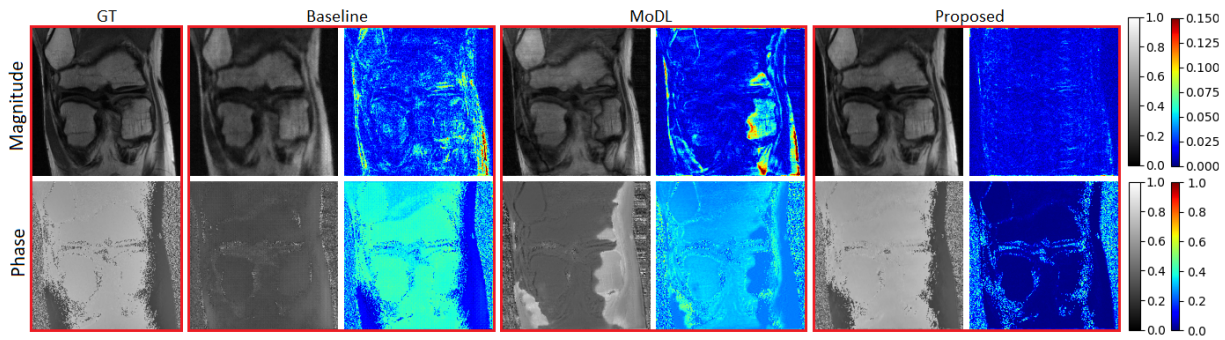

Figure 4. Qualitative results and comparison of the proposed method for complex-valued images from the fastMRI dataset. For each method, the two columns show the reconstructed outputs and their absolute difference with the GT.

Table 4. Quantitative results and comparison using MRNet dataset

Method	Noise-free images	10% noise level	20% noise level
	PSNR (dB) / mSSIM		
ZFR	19.916 / 0.6753	19.889 / 0.6385	19.806 / 0.5681
MoDL [1]	28.471 / 0.8643	27.659 / 0.7740	25.903 / 0.6303
DAGAN [54]	31.529 / 0.8754	30.452 / 0.8182	28.267 / 0.7098
Proposed	**34.010 / 0.9306**	**33.031 / 0.9097**	**31.670 / 0.8802**

Table 5. Quantitative results & comparison using fastMRI dataset

Method	Parms	Noise-free images	10% noise level	20% noise level
		PSNR (dB) / mSSIM		
ZFR	-	19.813 / 0.788	19.795 / **0.788**	19.744 / **0.773**
Baseline	133M	22.957 / 0.593	22.928 / 0.588	22.841 / 0.574
MoDL [1]	1.1M	32.201 / 0.585	32.140 / 0.564	31.970 / 0.513
Proposed	1.7M	**34.538 / 0.789**	**34.231** / 0.787	**33.468** / 0.767

fastMRI dataset: To evaluate the performance of our approach on complex-valued images, we perform another set of experiments using the fastMRI dataset. For comparison, we report the results for ZFR, MoDL, and a baseline based on DAGAN. Instead of taking the magnitude images as input, the baseline takes the complex-valued images concatenated as two real-valued channels. The output is also modified in the same way. Magnitude images are used for the adversarial and VGG loss, while complex-valued images are used for the MSE and frequency loss [54].

Table 5 and Fig. 4 show the quantitative and qualitative results[10], respectively. Despite having significantly fewer trainable parameters (Table 5) in the generator as compared to the baseline, our approach obtains a superior reconstruction. This is highlighted by the similarity of the recovered magnitude and phase images to the GT. The trade-off between the PSNR and mSSIM for ZFR and baseline indicates that the latter gives an improved phase reconstruction at the cost of low-quality magnitude reconstruction. The baseline output demonstrates a significant loss in both magnitude and phase content, with the former being visibly blurry and the latter showing alarming differences. Although MoDL obtains a better PSNR than the baseline, it visibly fails to reconstruct some portions of the magnitude image, hallucinates some details, and gives an inaccurate phase recovery. Our approach also obtains robust reconstructions in the presence of noise in the undersampled data. These observations reiterate the importance of complex-valued operations for processing the CS-MRI data.

[10]Qualitative results of ZFR and a high-resolution version of this figure can be found in the supplementary.

3.5. Zero-shot Inference

In this experiment, the model trained on brain images from the MICCAI 2013 dataset is tested for the reconstruction of images of canine legs from the MICCAI 2013 challenge. This model achieves an average PSNR of 42.949 dB and mSSIM of 0.9864 when inferred for 2000 test images[11]. Although the test set images are of completely different anatomy compared to those used for training, our approach can obtain high-quality reconstruction.

Potential Hallucination by GANs: The proposed scheme using the L_{ℓ_1}, L_{mSSIM}, L_{wvt} loss functions during training tries to align the reconstructed MR image closely to the GT, in effect trying to reduce the possibility of hallucination by GAN. Experiments using different sampling masks, datasets, and zero-shot inference help demonstrate that the proposed model displayed no signs of hallucination and generalizes well. An expert radiologist confirmed the same after examining the GT and reconstructed images.

4. Conclusion

We propose a first-of-its-kind Co-VeGAN framework to combat the fundamental weakness of SOTA DL-based frameworks for CS-MRI reconstruction, *i.e.* their inability to process the inherently complex-valued MRI data. We also propose a learnable complex-valued activation, PC-WSS, to overcome the transferability issues in real-valued activations and perform extensive comparisons to demonstrate its suitability to the complex domain. The insights provided by detailed analyses and ablation studies of our approach, along with the SOTA results obtained on both real and complex-valued datasets, would prove beneficial for future applications reliant on complex-valued data.

[11]Qualitative results can be found in the supplementary.

References

[1] H. K. Aggarwal, M. P. Mani, and M. Jacob. MoDL: Model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging*, 38(2):394–405, 2019.

[2] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning (ICML)*, volume 70, pages 214–223, 2017.

[3] M. Arjovsky, A. Shah, and Y. Bengio. Unitary evolution recurrent neural networks. *ArXiv*, abs/1511.06464, 2015.

[4] C. Bahadir, A. Wang, A. Dalca, and M. Sabuncu. Deep-learning-based optimization of the under-sampling pattern in MRI. *IEEE Transactions on Computational Imaging*, PP:1–1, 07 2020.

[5] N. Bien et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. *PLoS Medicine*, 15, 2018.

[6] Kai Block, Martin Uecker, and Jens Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine*, 57:1086–98, 2007.

[7] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. Compressed sensing using generative models. In *International Conference on Machine Learning (ICML)*, volume 70, pages 537–546, 2017.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[9] I. Danihelka, G. Wayne, B. Uria, N. Kalchbrenner, and A. Graves. Associative long short-term memory. *ArXiv*, abs/1602.03032, 2016.

[10] G. Daval-Frerot et al. Complex-valued neural networks with trainable activation functions for magnetic resonance imaging. *International Society for Magnetic Resonance in Medicine (ISMRM) Workshop on Machine Learning*, 2018.

[11] M. Dedmari, S. Conjeti, S. Estrada, P. Ehses, T. Stoecker, and M. Reuter. Complex fully convolutional neural networks for MR image reconstruction. *Medical Image Computing and Computer Assisted Interventions (MICCAI) Workshops*, 2018.

[12] P. Deora, B. Vasudeva, S. Bhattacharya, and P. M. Pradhan. Structure preserving compressive sensing mri reconstruction using generative adversarial networks. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 2211–2219, 2020.

[13] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2016.

[14] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang. Compressive sensing via nonlocal low-rank regularization. *IEEE Transactions on Image Processing*, 23(8):3618–3632, 2014.

[15] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[16] E. M. Eksioglu. Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI. *Journal of Mathematical Imaging and Vision*, 56(3):430–440, 2016.

[17] H. El-Rewaidy et al. Deep complex convolutional network for fast reconstruction of 3D late gadolinium enhancement cardiac MRI. *NMR in Biomedicine*, 33(7):e4312, 2020.

[18] Peter Gatehouse, Jennifer Keegan, Lindsey Crowe, Sharmeen Masood, Karl-Friedrich Kreitner, and David Firmin. Applications of phase-contrast flow and velocity imaging in cardiovascular MRI. *European Radiology*, 15:2172–84, 2005.

[19] I. Goodfellow et al. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680. 2014.

[20] N. Guberman. On complex valued convolutional neural networks. *ArXiv*, abs/1602.09046, 2016.

[21] Kerstin Hammernik et al. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6):3055–3071, 2018.

[22] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1026–1034, 2015.

[23] A. Hirose and S. Yoshida. Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence. *IEEE Transactions on Neural Networks and Learning Systems*, 23(4):541–551, 2012.

[24] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269, 2017.

[25] H. Huang, R. He, Z. Sun, and T. Tan. Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1698–1706, 2017.

[26] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo. Deep learning for undersampled MRI reconstruction. *Physics in Medicine & Biology*, 63(13), 2017.

[27] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning (ICML)*, volume 37, pages 448–456, 2015.

[28] P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, 2017.

[29] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017.

[30] K. Kreutz-Delgado. The complex gradient operator and the CR-calculus. *ArXiv*, abs/0906.4835, 2009.

[31] Zongying Lai et al. Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform. *Medical Image Analysis*, 27:93 – 104, 2016.

[32] B. Landman and S. Warfield (Eds.). 2013 Diencephalon standard challenge. `https://my.vanderbilt.edu/masi/workshops/`.

[33] R. Liu, Y. Zhang, S. Cheng, Z. Luo, and X. Fan. 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.

[34] Y. Liu et al. Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. *PLOS ONE*, 10(4):1–19, 2015.

[35] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[36] M. Mardani et al. Deep generative adversarial neural networks for compressive sensing MRI. *IEEE Transactions on Medical Imaging*, 38(1):167–179, 2019.

[37] M. Mirza and S. Osindero. Conditional generative adversarial nets. *ArXiv*, abs/1411.1784, 2014.

[38] T. Nitta. On the critical points of the complex-valued neural network. In *International Conference on Neural Information Processing (ICONIP)*, volume 3, pages 1099–1103, 2002.

[39] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, 1981.

[40] T. M. Quan, T. Nguyen-Duc, and W. Jeong. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Transactions on Medical Imaging*, 37(6):1488–1497, 2018.

[41] S. Ravishankar and Y. Bresler. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Transactions on Medical Imaging*, 30(5):1028–1041, 2011.

[42] D. P. Reichert and T. Serre. Neuronal synchrony in complex-valued deep networks. *ArXiv*, abs/1312.6115, 2013.

[43] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 234–241. Springer, 2015.

[44] P. Roy, S. Ghosh, S. Bhattacharya, and U. Pal. Effects of degradations on deep neural network architectures. *ArXiv*, abs/1807.10108, 2018.

[45] P. Song, L. Weizman, J. F. C. Mota, Y. C. Eldar, and M. R. D. Rodrigues. Coupled dictionary learning for multi-contrast MRI reconstruction. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 2880–2884, 2018.

[46] L. Sun et al. A dual-domain deep lattice network for rapid MRI reconstruction. *Neurocomputing*, 397:94–107, 2020.

[47] L. Sun, Z. Fan, X. Fu, Y. Huang, X. Ding, and J. Paisley. A deep information sharing network for multi-contrast compressed sensing MRI reconstruction. *IEEE Transactions on Image Processing*, 28(12):6141–6153, 2019.

[48] C. Trabelsi et al. Deep complex networks. In *International Conference on Learning Representations (ICLR)*, 2018.

[49] P. Virtue, S. X. Yu, and M. Lustig. Better than real: Complex-valued neural nets for MRI fingerprinting. In *IEEE International Conference on Image Processing (ICIP)*, pages 3953–3957, 2017.

[50] Shanshan Wang et al. DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution. *Magnetic Resonance Imaging*, 68:136 – 147, 2020.

[51] X. Wang et al. ESRGAN: Enhanced super-resolution generative adversarial networks. In *European Conference on Computer Vision (ECCV) Workshops*, pages 63–79, 2018.

[52] S. Wisdom, T. Powers, J. Hershey, J. Le Roux, and L. Atlas. Full-capacity unitary recurrent neural networks. In *Advances in Neural Information Processing Systems*, pages 4880–4888, 2016.

[53] J. Xie, L. Xu, and E. Chen. Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems*, pages 341–349. 2012.

[54] G. Yang et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1310–1321, 2018.

[55] Y. Yang, J. Sun, H. Li, and Z. Xu. Deep ADMM-Net for compressive sensing MRI. In *Advances in Neural Information Processing Systems*, pages 10–18. 2016.

[56] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[57] Jure Zbontar et al. fastMRI: An open dataset and benchmarks for accelerated MRI. *ArXiv*, abs/1811.08839, 2018.

[58] Z. Zhang, H. Wang, F. Xu, and Y. Jin. Complex-valued convolutional neural network and its application in polarimetric SAR image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):7177–7188, 2017.