IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 August 2024, accepted 16 August 2024, date of publication 23 August 2024, date of current version 3 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3448612

RESEARCH ARTICLE

DISTA-CSNet: Efficient Data Aware Deep Learning Model for CS MRI Recovery

HASSAAN HAIDER[1], JAWAD ALI SHAH[2], (Senior Member, IEEE), UMER JAVEED[3], SHARJEEL ABID BUTT[1], AND KUSHSAIRY A. KADIR[4], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, International Islamic University Islamabad, Islamabad 44000, Pakistan
[2]Cloud Computing Service Provider, Kuala Lumpur 59200, Malaysia
[3]Sydney, Australia
[4]Electrical Section, Universiti Kuala Lumpur British Malaysian Institute (UniKL BMI), Gombak 53100, Malaysia

Corresponding author: Hassaan Haider (hassaan.haider@iiu.edu.pk)

This work was supported in part by the Higher Education Commission, Pakistan, under Project 5797; and in part by the University of Kaula Lumpur British Malaysian Institute, Malaysia.

ABSTRACT This research paper explores the rising interest in utilising deep learning methods, specifically Convolution Neural Network (CNN), to improve the reconstruction of Compressively Sampled Magnetic Resonance Imaging (CS MRI) images from under-sampled data. By training deep learning architectures on extensive data sets of paired under sampled and fully sampled MR images, these models aim to capture intricate patterns and structures, ultimately enhancing the accuracy of MR image reconstructions. A novel deep learning model is proposed dubbed as ''Deep Iterative Shrinkage Thresholding Algorithm-Compressed Sensing Network'' (DISTA-CSNet), specifically designed for efficient recovery of CS MRI. Our model showcases impressive results with only 20 epochs and can be effortlessly trained on diverse datasets. To ensure robustness across different datasets, the dropouts are incorporated into the model, evident from testing results. The trained DISTA-CSNet exhibits remarkable performance in recovering CS MRI from various data sets, surpassing several advanced deep learning techniques with changing compression ratios consistently. The proposed model demonstrates significant enhancements in both Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics, affirming the efficacy of our proposed model. The capability of our DISTA-CSNet in accurately reconstructing CS MRI images from 5-fold undersampled data shows promise in improving medical imaging applications and advancing the field of compressed sensing MRI.

INDEX TERMS Compressed sensing, MRI, deep neural networks, iterative shrinkage, dropouts, soft thresholding, compressively sampled MRI.

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a non-invasive diagnostic technique based on nuclear magnetic resonance, providing high-quality soft tissue contrast. It is valuable for monitoring oxygen saturation levels in the brain, assessing blood flow, and measuring body temperature. Efforts to improve MRI speed have focused on hardware mechanisms, including faster data acquisition, improved pulse sequences, and parallel imaging with multiple coils. However, some techniques may result in reduced field of view, decreased signal-to-noise ratio, or unwanted nerve stimulation.

A. COMPRESSIVELY SAMPLED MRI

There is a challenge of further reducing MRI scan time due to hardware limitations. Compressed Sensing (CS) combined with non-linear reconstruction techniques offers a solution by reducing the number of samples collected in k-space, enabling high-quality MRI recovery with shorter scan times. Although MRI acquisition is devoid of potentially harmful waves, its long scanning duration, and image quality remains susceptible to object motion, thereby inducing discomfort in

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

patients. The temporal extent of MRI scanning is intricately tied to the number of samples acquired within the Fourier domain, commonly referred to as k-space. Compressed Sensing (CS) serves to mitigate these challenges by yielding clinically acceptable MR images utilising a reduced number of samples, consequently reducing scan duration. Due to its inherent sparse representation within established domains such as Finite Difference, Discrete Cosine Transform and Wavelet, MRI stands as a promising domain for CS application. CS MRI acquisition leverages non-uniform random undersampling techniques to maximise incoherence with the sparsifying domain. The efficacious recovery of MRI from compressively sampled k-space data poses computational complexities and temporal demands within the realm of CS, albeit executed offline to alleviate patient discomfort. The extant body of literature introduces several adept algorithms advanced for the purpose of CS MRI recovery [1], [2], [3], [4], [5], [6], [7], [8]. Equation (1) defines the optimisation problem posed by CS MRI recovery.

$$\hat{x} = \arg\min_{x} \|\Psi x\|_1 \text{ subject to } \|F_u x - y\|_2^2 < \epsilon \quad (1)$$

where $\hat{x}$ is the estimated image, *Psi* is the sparsifying transform, l_1 norm is the sparsity promoting constraint in estimated solution. F_u is the undersampling domain, y is the undersampled measurements in k-space and ϵ is the tolerance level for noise.

B. MACHINE LEARNING AND CS MRI

Recently, there has been notable interest in utilising deep learning techniques to enhance the reconstruction of CS MRI images from sparsely sampled data. Among these techniques, CNN have demonstrated their capacity to capture intricate image patterns and structures, facilitating the recovery of high-quality MRI images. Researchers have devised deep learning architectures to cater specifically to CS MRI reconstruction, trained to establish a mapping between undersampled input data and fully-sampled MRI images. These architectures leverage extensive data sets containing paired undersampled and fully sampled MR images, with the intent of assimilating knowledge about underlying image structures to enhance the accuracy of reconstruction [9], [10], [11], [12].

Deep learning strategies for CS MRI recovery often adopt an end-to-end learning approach, where the network takes undersampled data as input and directly produces fully sampled images as output, eliminating the requirement for intermediate iterative reconstruction stages. This framework not only enhances efficiency but also reduces computational intricacies [13]. Researchers have explored methods to address limited training data in various ways, including artificial dataset augmentation. Approaches encompass diverse techniques such as random transformations, patch extraction, and simulated undersampling patterns. Furthermore, regularisation methods, including sparsity and total variation constraints, have been integrated to improve the generalisation and reconstruction quality of deep learning models [14], [15], [16]. In response to limited annotated training data, researchers have explored transfer learning as a viable approach for addressing CS MRI reconstruction. This involves refining models pre-trained on extensive image datasets to cater to CS MRI reconstruction tasks. By initially training convolutional neural network (CNN) models on extensive natural image datasets such as ImageNet, foundational image features are acquired, which can subsequently be harnessed to enhance CS MRI reconstruction [17], [18], [19]. Nonetheless, these techniques exhibit limitations due to their utilisation of fixed measurement matrices and predetermined image dimensions. A solution to mitigate these limitations was introduced through a CNN-based approach [20]. This method facilitated the learning of mappings from initial CS reconstructions to significantly improved outcomes. In the pursuit of achieving superior quality CS reconstructions, the implementation of Generative Adversarial Networks (GANs) has emerged [17], [21]. These CNNs employed Mean Squared Error (MSE) as their underlying cost function. To enhance the quality of MR image reconstructions, a Deep Learning framework incorporating the Bayesian methodology was presented in [22], utilising prior probabilities as a training loss.

C. ITERATIVE SHRINKAGE METHODS IN DEEP LEARNING

The iterative sparse coding algorithm, introduced by [23], serves as the foundational concept for the discriminative learning techniques explored in this study. These methodologies incorporate insights from trained models into deep learning (DL) approaches. Notably, a single iteration within these methods bears resemblance to conventional Convolutional Neural Network (CNN) training. In contrast to conventional CNNs, these frameworks possess the ability to learn the required mapping during training by iteratively increasing the number of iterations, eliminating the need for extra parameters. Consequently, the depth of the network can be expanded without incurring the parameter overhead associated with traditional CNNs in [24], where proximal operators were replaced with a CNN, drawing inspiration from primal-dual hybrid gradient techniques.

A novel variant of projected gradient descent, designed to enforce measurement consistency between reconstructed images and their corresponding measurements, thus ensuring convergence under specific conditions in [25]. This technique demonstrated superior performance in reconstructing sparse-view computed tomography (CT) images. Furthermore, [26] proposed a gradient descent method for reconstructing knee MR images with a 4-fold undersampling factor, incorporating a variational model. This approach exhibited enhanced computational efficiency for undersampled reconstructions, utilizing a single graphics card. The unrolled approximate message passing (D-AMP) algorithm, introduced [27], incorporates a deep CNN to replace the denoising operator within each iteration of the learned-AMP method.

To overcome the challenge of proving the bounded properties of complex deep denoisers. Provable and trainable bounded denoiser using dual tight frames and spatial-variation thresholds are introduced in [28], Video snapshot compressive imaging (SCI) to recover multiple video frames from a single measurement, with recent plug-and-play (PnP) methods leveraging pre-trained deep Gaussian denoisers is proposed in [29]. A novel provable bounded denoiser, BMDual, which integrates a trainable denoiser using dual tight frames with the BM3D denoiser and incorporates multiple dual frames into a new regularization model for CSMRI is proposed in [30].

In a similar vein, previous work [31] presented a compressed sensing (CS) reconstruction strategy for MR and natural images using an unrolled alternate direction method of multipliers (ADMM) algorithm. The discriminatory learned CS recovery and ADMM parameters of this model yielded impressive results for both real-valued natural images and complex-valued MR images.

An innovative approach was proposed in [10], where an iterative thresholding technique (ISTA) was employed to recover real images from undersampled observations. This ISTA-Nets framework seamlessly integrates optimization-based and network-based methodologies, showcasing a well-structured topology optimized for CS image reconstruction. Notably, ISTA-Nets offer interpretability, shedding light on their operational mechanics. All parameters within ISTA-Nets are end-to-end discriminatively learned, ensuring efficient harnessing of the network's potential for image reconstruction. By fusing the strengths of optimization-based and network-based approaches and emphasizing interpretability, ISTA-Nets emerge as a promising avenue for advancing CS image recovery techniques.

A novel high-throughput deep unfolding network (HiTDUN) [32] is proposed, capable of transmitting multi-channel information between adjacent network stages, resulting in superior performance compared to state-of-the-art DUNs in CS-MRI applications. HiTDUN investigates the optimization algorithms unfolded into deep unfolding networks (DUNs) for CS-MRI, addressing the lack of discussion on the optimal algorithm and existing DUN bottlenecks.

In the context of Compressed Sensing MRI (CS-MRI) restoration, the Multi-Layer Convolutional Sparse Coding (ML-CSC) system [11] leverages iterative thresholding techniques to extract nonlinear mapping parameters from CS MRI k-space data. This framework effectively learns the requisite mapping from CS measurements, specifically benefiting knee and brain MRI reconstruction. Notably, the incorporation of additional learn-able parameters in the deep neural network yields improved CS reconstruction results with minimal parameter overhead.

Some of the limitations of modern existing techniques are listed below:

- Presently, contemporary deep learning architectures heavily lean on heuristic methodologies for training. However, the absence of meticulous theoretical analysis obstructs the optimal enhancement of feature acquisition. This concern takes on heightened significance within the realm of biomedical images, where precise solutions to inverse problems are pivotal for clinical comprehension and diagnosis.
- The proposed frameworks designed to tackle inverse problems via deep learning necessitate extensive training on specialised datasets. This requisite poses considerable challenges when it comes to adapting these models to real-world applications, particularly in scenarios where tailored datasets are not immediately accessible or practical to construct.
- A notable research gap emerges from the necessity for a versatile architecture capable of being tailored to diverse datasets and imaging protocols. The creation of such adaptable models could significantly elevate the integration of AI-assisted tools into the expansive biomedical imaging landscape, facilitating consistent and efficacious solutions across a spectrum of scenarios.
- Despite substantial efforts directed towards mastering model parameters, the translation of these acquired parameters into pragmatic implementations within clinical settings remains a formidable obstacle. The demand is for streamlined methodologies that facilitate the seamless deployment of testing and restoration frameworks, ensuring their usability within authentic medical environments devoid of unnecessary intricacies.
- A significant number of current methodologies experience extended training duration, leading to potential vulnerability when confronted with novel environments or circumstances. Swift adaptability to evolving conditions remains pivotal in clinical practice, prompting research endeavours aimed at formulating algorithms that strike a harmonious equilibrium between training time and resilience, thus guaranteeing efficient performance within dynamic medical contexts.
- While deep learning models exhibit exceptional performance, their intricacy frequently impedes interpretability and explanation. Bridging the gap between achieving exemplary accuracy and providing clinicians with intelligible insights remains an enduring trial. The evolution of techniques that generate outcomes possessing both precision and comprehensibility for medical practitioners is an ongoing frontier of investigation.

The proposed DISTA-CSNet model introduces a novel approach by incorporating an iterative soft thresholding algorithm, enabling the augmentation of network depth without incurring additional feature complexity and computational overhead. This innovative strategy not only enhances the model's capacity to capture intricate image features but also minimises computational demands. The efficacy of the proposed model is demonstrated through its impressive performance in accurately reconstructing under-sampled Brain and Knee MRI data. This signifies its potential as a valuable tool for advancing the field of medical imaging

by providing robust solutions for challenging reconstruction tasks. Following are the major contributions of our proposed model.

- *Versatile Training and Swift Convergence:* Our proposed DISTA-CSNet model offers the advantage of easy adaptation to various datasets, showcasing its versatility.The Data Consistency Constraint is added at every stage of proposed model that ensures the estimated solution is consistent with the actual samples taken during scanning process of MRI. It significantly improved the convergence in just 20 epochs, indicating swift convergence during training.
- *Shrinkage and l_1-norm approximation by tanh():* We have introduced hyperbolic tangent based thresholding as well as l_1-norm approximation that gives us flexibility to make our shrinkage function flexibility. Hyperbolic tangent based shrinkage function gradually become soft thresholding to hard thresholding function as the model progressed towards finding the optimal solution. Similarly, l_1-norm approximation by hyperbolic tangent allows us to gradually move from smooth approximation of l_1-norm to more accurate approximation that avoids local minima and converges to optimal solution more swiftly.
- *Robustness through Dropout Integration:* The DISTA-CSNet has proven to be Robust by introduction of Dropout. The model's robustness across diverse datasets i.e. Two Brain MRI and One Knee MRI by incorporating dropouts into its architecture. This integration ensures a more resilient performance during testing, contributing to its reliability in handling different imaging scenarios.
- *Consistent Excellence in CS MRI Recovery:* The trained DISTA-CSNet stands out for its remarkable ability to recover Compressive Sensing (CS) MRI from a range of datasets, a feat accomplished through the strategic application of dropouts. This consistent excellence underscores the model's potential to provide accurate and reliable image reconstructions.
- *Quantifiable Improvements in Quality Metrics:* Our approach yields substantial improvements in both PSNR and SSIM metrics, underscoring its efficacy in enhancing image quality. This quantifiable boost in image fidelity reaffirms the valuable contribution of our DISTA-CSNet model in advancing the field of CS MRI recovery.

Section II defines the proposed DISTA-CSNet Model alongwith the DISTA-CSNet Training and Testing mechanism of the proposed model. Training and Testing results are discussed in Section III. This article culminates with conclusions and future work directions in Section IV and Section V respectively.

II. DIST-CSNet MODEL

In the context of deep learning architectures, training often relies on heuristic techniques, necessitating theoretical analysis for improved feature learning and accurate solutions to inverse problems, especially pertinent in biomedical imaging where reconstruction quality is pivotal for accurate diagnoses. Custom dataset-intensive training precedes the real-world deployment of frameworks, highlighting the demand for adaptable, broadly applicable architectures to integrate machine learning into MRI procedures. Ensuring ease of use in clinical settings is crucial once model parameters are acquired to promote wider usability.

The novel DISTA-CSNet model capitalises on the effectiveness of iterative shrinkage-based techniques in addressing CS-MRI's inverse problem, showcasing its suitability within the deep learning domain. It harnesses the concept of CNN unfolding, a strategy that augments network depth without introducing additional parameters or computational complexities, setting it apart from conventional CNN-based approaches. This innovation leads to substantial improvements in reconstruction performance, as substantiated by extensive experiments conducted across diverse datasets, consistently yielding high-quality reconstructions.

The model's prowess extends to its generalisation capabilities, demonstrated through training on three benchmark datasets and subsequent restoration of images using this adaptable framework. Rigorous testing on both GPU and CPU platforms confirms commendable restoration times and showcases elevated PSNR/SSIM metrics, affirming the model's potential for real-world medical imaging scenarios. Overall, the DISTA-CSNet model emerges as a compelling solution with the ability to efficiently and accurately reconstruct MR images, holding promise for transformative impacts in the field.

The proposed CS-MRI restoration approach introduced in this study integrates novel insights from theoretical deep-learning research [10], [33]. This innovative method combines an iterative shrinkage-based technique with a multi-layered convolutional neural network to swiftly learn the mapping between fully sampled MR images and corresponding CS MRI k-space. A key advantage of the DISTA-CSNet model lies in its rapid training capability (20 epochs) and comprehensive global pursuit strategy, enhancing the efficiency of learning the mapping function.

During the training phase, the DISTA-CSNet model benefits from GPU acceleration, ensuring fast training times and effective parameter learning. Once trained, the model showcases its versatility by restoring CS-MR images across various CS ratios and random masks, exhibiting robustness and accuracy under diverse undersampling conditions. Notably, the framework's ability to perform test set reconstructions on a CPU without relying on GPUs is a notable advantage. This feature ensures feasible restoration times, highlighting the efficiency of the trained parameters. As a result, the approach holds practical value for clinical settings, where GPU availability might be limited, providing an efficient solution for CS-MRI restoration and significantly reducing processing times.

In general, the iterative shrinkage thresholding-based algorithms in the context of CS MRI recovery recast the cost function defined by (1) as follows:

$$\hat{x} = \arg\min_{x} \frac{1}{2} \left\| y - \Phi\Psi^H x \right\|_2^2 + \lambda \left\| x \right\|_1 \tag{2}$$

The layered iterative shrinkage methods proposed by [30] and [143] attempt to unfold neural networks by simultaneously minimising the above cost function by iterative thresholding methods. This approach is aimed at achieving a global multi-layer sparse model, where it is not attainable. Instead, the focus is on obtaining representation estimates that solely describe the local layer defined by (3)

$$\hat{x} = \arg\min_{x_i} \frac{1}{2} \left\| y_i - \Phi\Psi^H x_{i-1} \right\|_2^2 + \lambda \left\| x_{i-1} \right\|_1 \tag{3}$$

where x_i represents the current layer of the model.

A. DISTA-CSNet TRAINING MODEL

The suggested approach processes images in k-space with related CS measurements using multi-layer Iterative Shrinkage Thresholding Algorithms (ML-ISTA). The proposed DISTA-CSNet aims to effectively train to quickly map fully sampled MR images and matching CS MRI k-space and store these mappings as model parameters. The model trains with only 20 epochs while achieving better PSNR and SSIM than other state-of-the-art algorithms as evident from experimental results. The testing module then reconstructs test images using the trained ML-ISTA learning parameters and computes the recovered images' PSNR and SSIM. For all layers, it uses Xavier initialisation to initialise the dictionaries. *Algorithm-1* presents the multi-layered basis pursuit algorithm, which makes use of the iterative thresholding process. To train the model parameters, the framework collects ground truth images, k-space data, and masks for under-sampling. Without the use of additional parameters, the learning process, Algorithm-1, uses ML-ISTA unfolding to deepen the learning framework. The model was trained separately on Brain MR image and Knee MR image. Both models were tested for CS MRI recovery based on PSNR and SSIM.

The first dataset comes from earlier work [10], and it comprises 850 brain MRIs that were captured in the sagittal and axial planes as a pair of (CS MRI, Label (Ground Truth Image) (Dataset 1). 621 Knee MRIs from patients comprise the second dataset used to train the CS-MRI framework. These Knee MRIs are fully sampled of the coronal view with PD fat saturation and were taken using a 1.5T imaging device (Siemens MAGNETOM Symphony) at the Hospital Kuala Lumpur (Dataset 2). A certified radiologist classified the MRIs into three categories: full tear, partial tear, and normal knee. The images were collected to diagnose anterior cruciate ligament (ACL) damage. The performance was evaluated using a test set of 50 and 21 MRIs randomly from both datasets, respectively. The third dataset comprises 123 slices thick, 256 × 256 size MRIs of ageing-normal brains in the coronal plane [34] (Dataset 3), this dataset was used for testing purposes only. It was tested by using the model trained on the first dataset. The Harvard Medical School's AANLIB database makes this dataset accessible to the public.

The incorporation of Monte Carlo (MC) Dropout has notably enhanced the robustness of the model, evident from its performance across diverse datasets by mitigating overfitting. Dropout, initially introduced by Gal and Ghahramani [35], offers a practical and computationally efficient approach to estimate model uncertainty in deep neural networks. This technique involves applying dropout during both training and inference, facilitating the computation of

Algorithm 1

Task: Training of DISTA-CS Net with CS MRI data along with its fully sampled labels

$$\hat{x} = \arg\min_{x_i} \frac{1}{2} \left\| y - F_u\Psi^H x \right\|_2^2 + \lambda \sum_{i=1}^{n} x_i tanh(\gamma x_i)$$

Input: Training Data (CS MRI along with labels),Dictionary $\Phi \epsilon R^{m\times n}$,
Masks for Undersampling, Thresholing Parameter λ, β.
Adams Optimizer learning rate μ, undersampled data $\mathbf{y}$.
Output: Trained DISTA - CSNet Model for CS-MRI Recovery

1) Initialisation: Initialise Wieghts and Conv filters using Xavier Initialisation
   Set thresholding λ and β to a learn-able parameter.
   initial residual: $r_0 = x_0 - \alpha F_u^{-1}(y_0 - F_u x_0)$
2) DISTA-CSNet Main Iteration: Increment i by 1,and apply the following steps:
   a) Residual: Compute $r_i = F_u^{-1}(y_{i-1} - F_u x_{i-1})$
   b) Data Consistency in Fourier Domain:

$$X_i[j] = F_u(X_i)$$

$$X_{DC}[j] = \begin{cases} x_i[j] & \text{if } x[j] = 0 \\ x[j] & \text{otherwise} \end{cases}$$

   c) Convolution: Perform convolution as shown in Figure 1.
   d) Dropouts: Perform dropouts as shown in Figure 1.
   e) Convolutions: Perform convolution transpose as shown in Figure 1.
   f) ReLU: Apply ReLU as shown in Figure 1.
   g) Shrinkage: Compute

$$S_{\lambda,\beta}(x) \cong \begin{cases} x_i(tanh(\beta(|x_i| - \lambda))) & \text{if } |x| > \lambda \\ 0 & \text{otherwise} \end{cases}$$

   h) Adam Optimisation: Compute updated x_i using learning rate μ
   i) Residual Update: Compute $r_{i+1} = y - F_u x_i$
   j) Stopping Rule: Go to step (2) until the maximum number of epochs criteria is met.
3) Output: Trained DISTA-CSNet Model

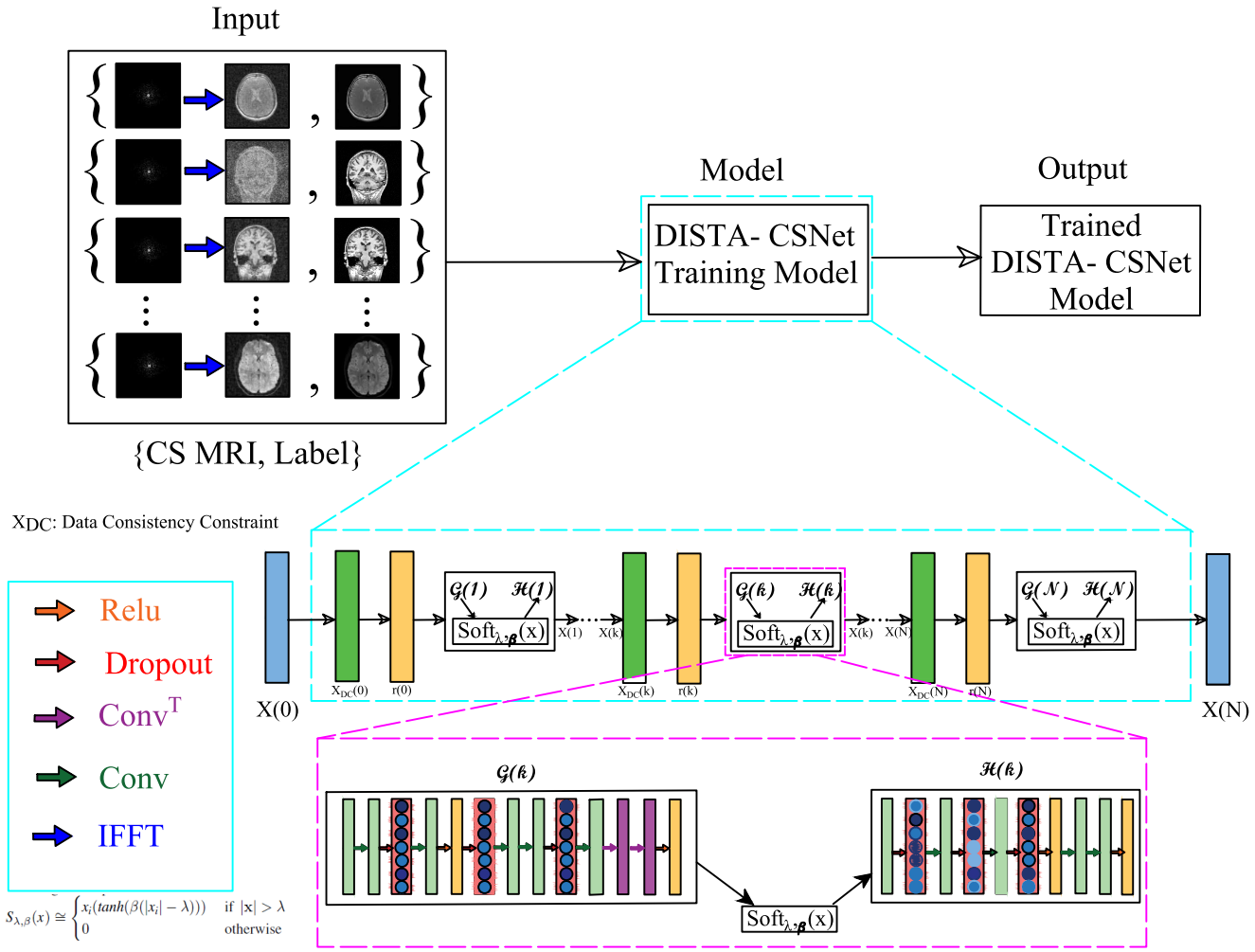

FIGURE 1. DISTA-CSNet training model.

predictive uncertainty and effectively tackling overconfidence issues in deep learning models. Through the repetition of dropout during inference, MC Dropout introduces prediction variability, thereby enhancing the reliability and effectiveness of deep neural networks.

Using a new thresholding method based on the *tanh* was used for training [7]. The *tanh* offers the advantage of an adjustable slope near the origin and bounded function, making it a preferable choice as an estimation for soft thresholding (ST). Consequently, the mathematical description of *tanh*-based ST is represented as:

$$S_{\lambda,\beta}(x) \cong \begin{cases} x_i(tanh(\beta(|x_i| - \lambda))) & \text{if } |\mathbf{x}| > \lambda \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The proposed shrinkage function includes a thresholding parameter λ, and a parameter β that controls the shape of the *tanh*. 2 and 3 illustrate the behavior of regularization parameters β and λ respectively. In 2, when β increases, the thresholding function behavior changes from soft thresholding to hard thresholding, When β is close to zero, (4) approximates the behaviour of a soft thresholding function. As β approaches infinity, the (4) transforms into a hard thresholding function. We utilize this property in iterative manner, where β grows with each iteration results in faster convergence. In 3, the Lagrangian multiplier λ defines the sparsity level in the estimated solution. Higher value of λ yields more sparsity in the solution and lower value of λ will result in denser solution. So, we have to choose λ appropriately by evaluating it over the range of values and choosing the optimal value for the specific problem. Both λ and β are learnable parameters.

The data consistency constraint (DCC) is introduced in the training model which significantly improves the learning rate of the training model. DCC in the frequency domain is a fundamental principle in CS MR recovery [6]. It ensures that the samples acquired in the k-space domain remain constant throughout the recovery process. In other words, the measured data points in the frequency domain, which correspond to the acquired MRI measurements, are

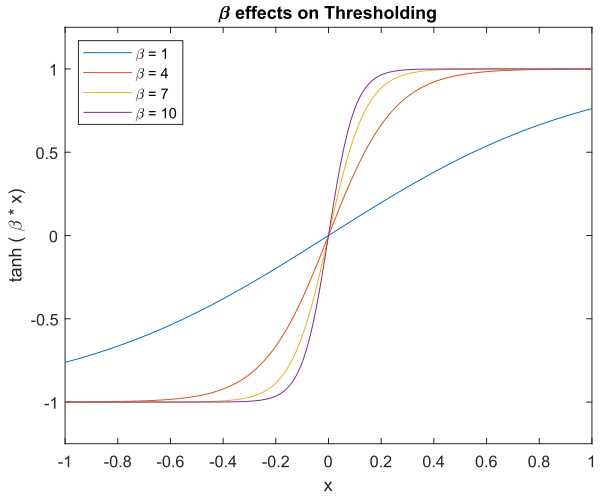

FIGURE 2. Shapes of *tanh* for varying β.

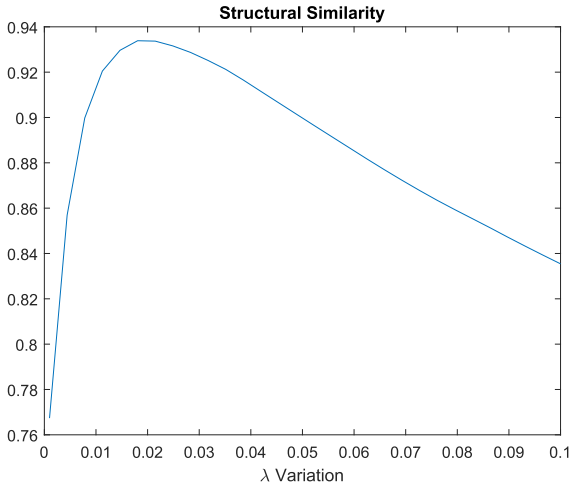

FIGURE 3. SSIM achieved against different values of λ.

preserved, and not altered during the iterations of the recovery algorithm. The data consistency constraint is crucial because it guarantees that the recovered MRI stay consistent with sampled data. It prevents the algorithm from introducing spurious information or modifying the original measurements, which could lead to erroneous results. By preserving the actual samples from the k-space domain, the algorithm ensures that the recovered image accurately represents the underlying structure of the imaged object, even with limited or under-sampled data. The data consistency constraint in the frequency domain is a vital aspect of CS MRI recovery, as it maintains the integrity of the acquired data and helps achieve reliable and accurate image reconstruction. Let $X_i[j] = F_u(X_i)$ be the FFT of estimated image.

$$X_{DC}[j] = \begin{cases} x_i[j] & \text{if } x[j] = 0 \\ x[j] & \text{otherwise} \end{cases} \tag{5}$$

In deep neural networks, convolutions are a fundamental operation used in convolutional layers to process and extract features from input data. A filter or kernel is slid across the input data during the convolutional procedure, and element-wise multiplication is done between the filter and kernel and the corresponding local input region. The result is then summed to produce a single value in the output feature map. Whereas, convolution transpose, also known as deconvolution, is an operation used in deep neural networks for tasks like image segmentation and image generation. It is the opposite of standard convolution and is used to increase the spatial dimensions of the feature maps.

ReLU is used as an activation function commonly used in deep neural networks to introduce non-linearity. It returns the input if it is positive, and zero otherwise. Xavier Initialisation was first proposed in [36]. The technique is widely used to initialise weights in neural network layers, especially in networks using the ReLU activation function, as it helps in maintaining signal variance and promoting stable and efficient training.

The optimisation approach used in training the proposed DISTA-CSNet model is Adam (Adaptive Moment Estimation) [37] which is frequently used in deep learning to update the weights of a neural network while it is being trained. To adapt the learning rate for each parameter based on their historical gradients and second moments, it utilises the advantages of both the RMSprop and momentum methods. Adam dynamically adjusts the learning rate, which makes it ideal for training complicated, high-dimensional neural networks and promotes faster convergence and improved performance across a range of deep learning applications.

The discrepancy loss evaluates the disparity between predicted neural network outputs and true target values, quantifying the alignment of model predictions with task-specific ground truth. It encompasses mean squared error (MSE) for regression and cross-entropy loss for classification tasks. The constraint loss integrates extra constraints or regularisation components, such as $l_1 or l_2$ norm terms, to shape the model's behaviour in training, encouraging weight sparsity and a more interpret-able model. These two losses combined form the total loss function, striking a balance between accurate data fitting (minimising discrepancy loss) and meeting constraints or regularisation (minimising constraint loss), culminating in an optimised model aligned with task objectives and specified preferences.

$$L_T = L_d + \gamma L_C \tag{6}$$

$$L_d = \frac{1}{N_d N} \sum_{i=1}^{N_d} \left\| x_i^{N_p} - x_i \right\|_2^2 \tag{7}$$

$$L_C = \frac{1}{N_d N} \sum_{i=1}^{N_d} \sum_{k=1}^{N_p} \left\| G^k(H^k(x_i)) - x_i \right\|_2^2 \tag{8}$$

where L_T, L_d and L_C are a total loss, discrepancy loss and constraint loss respectively. N, N_d, N_p and γ are the size of each block x_i, total training blocks, total phases, and regularisation constraint respectively. All these three losses during the training process by our model are depicted in Fig. 4 and 5. Fig. 4 shows the convergence of total loss,

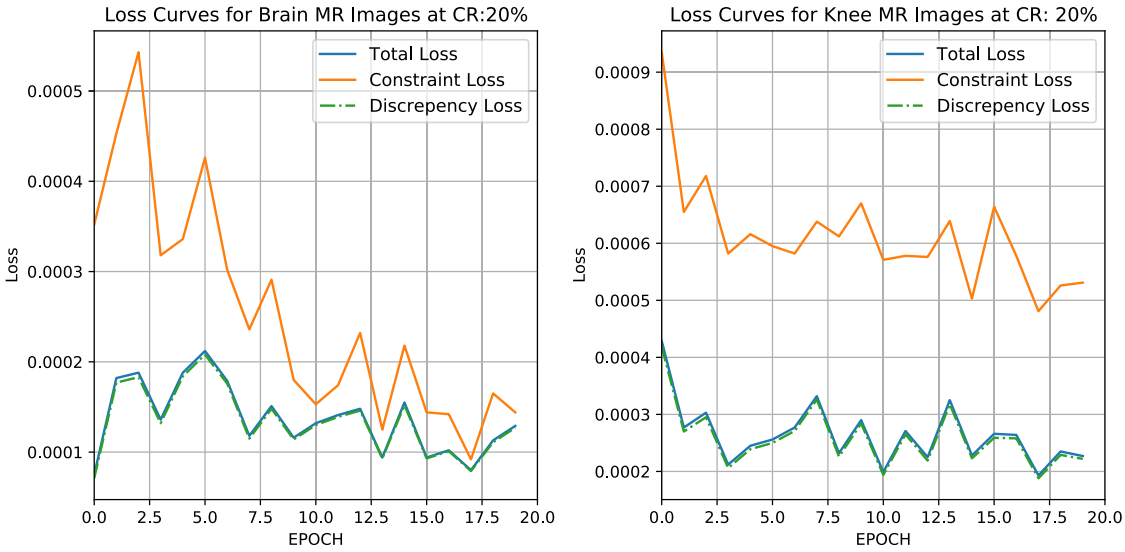

FIGURE 4. Loss functions vs Epochs for 5-fold compression of Brain and Knee MRIs.

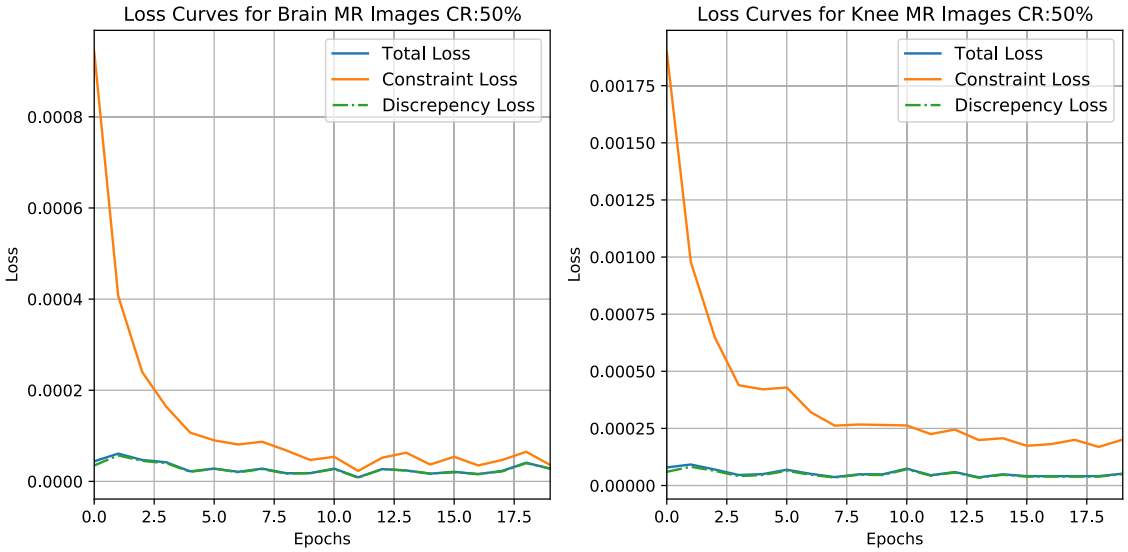

FIGURE 5. Loss functions vs Epochs for 2-fold compression of Brain and Knee MRIs.

discrepancy loss and constraint loss with respect to Epochs when the compression ratio (CR) is set to only 20% that gives five fold compression. Whereas Fig. 5 shows the convergence of discrepancy loss, constraint loss and total loss with compression ratio is set to 50% that result in half the samples acquired during MRI scanning process.

B. DISTA-CSNet TESTING MODEL

After training two separate DISTA-CSNet models for Compressed Sensing MRI reconstruction of Brain and Knee MRIs. The testing phase evaluated the models' accuracy by reconstructing MRIs from three test datasets: Brain MRI (50 slices), Knee MRI (21 images), and an additional Brain MRI dataset (123 slices). The trained models performed with great accuracy regarding PSNR and SSIM, demonstrating their ability to produce high-quality MRI estimates from compressed measurements. Overall, the DISTA-CSNet models proved effective in accurately reconstructing MRIs, showcasing their potential for practical medical imaging applications. The process of testing the model is illustrated in Fig. 6. *Algorithm 2* describes the testing process.

III. DISTA-CSNet MODEL TESTING RESULTS AND DISCUSSION

The DISTA-CSNet Model was subjected to testing using three distinct datasets: the first dataset consisted of Brain

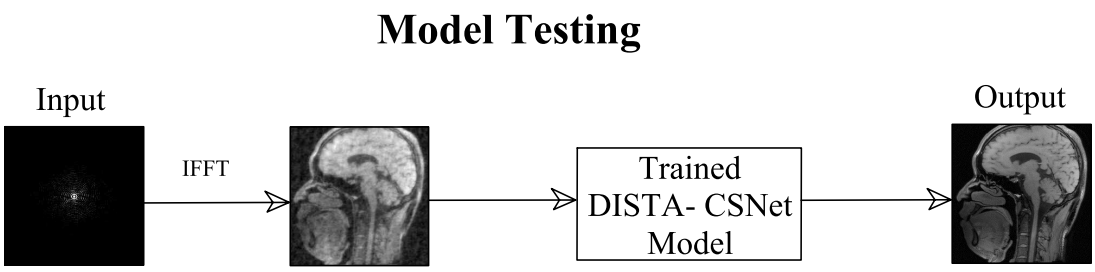

FIGURE 6. DISTA-CSNet testing model.

Algorithm 2 Testing DISTA-CSNet Process

Task: Test the Trained DISTA-CSNet for CS MRI Reconstruction.
Inputs: Trained Model, Test MR images, Undersampled Masks.
Output: Reconstructed MRI from CS-MRI, PSNR and SSIM of reconstructed image, Recovery time by GPU / CPU

1) *Initialisation*: Initialise Weights and Convolution filters using Xavier Initialisation
   Set λ and β to a learn-able parameter for thresholding.
   initial residual: $r_0 = x_0 - \alpha F_u^{-1}(y_0 - F_u x_0)$
2) Loading Model: Load pre-trained DISTA-CSNet Model
3) DISTA-CSNet Main Iteration: For every CS MRI in the test dataset:
   a) *Loading Model:* Load pre-trained learnt dictionary
   b) *Masking*: Apply masking to input MRI at appropriate CR in k-space
   c) *Reconstruction*: Apply a pre-trained model for MRI reconstruction.
   d) *Calculate Performance Measures*: Calculate PSNR and SSIM of the recovered image.
   e) *Stopping Rule*: If all images are tested in dataset go to Step 4 otherwise go to Step 3a.
4) *Output*: Reconstructed Image, Recovery Time, PSNR and SSIM in comparison to Ground Truth.

MRIs with 50 slices [10] (Dataset 1), the second dataset included 21 Knee MR images (Dataset 3), and the third dataset contained 123 Brain MRI slices [34] (Dataset 3). To assess the performance of the proposed algorithm, the reconstruction losses are calculated using performance metrics like Structural Similarity Index (SSIM) and Peak Signal to Noise Ratio (PSNR). If SSIM is equal to 1 that means exact compressively sampled MRI is recovered while if SSIM is equal to zero means very poor recovery and there is no similarity in reconstructed image and original image. The reconstructed MRI has value very close to 1 e.g., 0.9899 in case of Brain MRI with compression ratio of 50% that means recovered image is almost equal to original image and reconstruction losses are negligibleBy measuring the PSNR and SSIM values, the experiments aimed to quantitatively evaluate how accurately the DISTA-CSNet Model reconstructed the MRIs compared to the original ground truth images, providing insights into the algorithm's efficacy and potential for medical imaging applications.

The training and testing of the proposed model were performed on the Lambda (λ) Quad AI Workstation with a single Nvidia GeForce RTX 3090 GPU using JupyterLab platform. The batch size was set to 4 for training for Brain and Knee MRI training datasets.

Fig. 7 presents a selection of randomly recovered images, which were obtained during the testing phase using three distinct datasets. These datasets were subjected to a 5-fold compression in the sampling process, resulting in a significant 5-time reduction in scanning time compared to conventional methods. The PSNR / SSIM of the recovered images are also shown. The visual evidence from the recovered images indicates that they meet the required standards for clinical treatment. The reconstructed images demonstrate high quality and accuracy, making them suitable for practical medical applications with confidence.

Table 1 presents the effectiveness of the proposed DISTA-CSNet model that was trained on 800 Brain MR images, showcasing its superior performance in terms of PSNR and SSIM across various compression ratios. Even at a 5-fold compression, the model achieves an impressive average SSIM value of 0.9634, indicating its proficiency in producing high-quality reconstructions. During testing, the computational efficiency of the model demonstrated significant improvements when executed on a GPU, with computation times in the sub-second range. However, even on a CPU, the computational time has been reduced, making it feasible to test the model without relying on a GPU. Moreover, the training time for the DISTA-CSNet model was substantially reduced when trained on Dataset 1, requiring only 20 epochs and approximately 16 minutes to reach a near-optimal performance level. This observation highlights the fast-learning capability of the proposed method compared to other state-of-the-art approaches, further solidifying its potential for practical and time-efficient applications in medical image reconstruction.

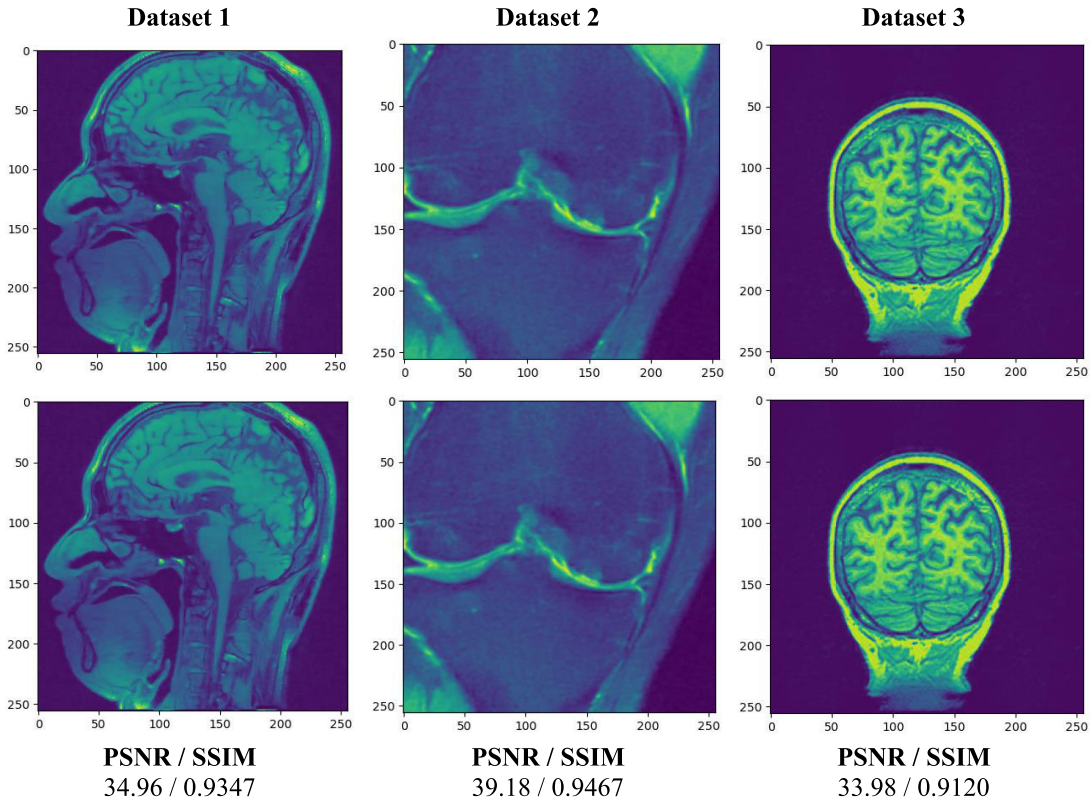

FIGURE 7. Sample recovered images from testing with different datasets having 5-fold compression.

TABLE 1. Dataset 1: Brain MRI with 50 slices testing results.

Algorithms	Compression Ratio								GPU / CPU Testing Time	GPU Training Time	Epochs
	20%		30%		40%		50%				
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM			
ADMM-Net	37.17	0.9374	39.84	0.9526	41.56	0.9664	43.00	0.9734	0.9535s /5.2s	03:31:23	200
ISTA-Net+	38.70	0.9484	40.97	0.9639	42.65	0.9729	44.12	0.9792	0.1437s /4.8s	02:38:15	200
ML-CSC	39.25	0.9551	41.50	0.9689	43.66	0.9774	45.96	0.9855	0.0688s / 2.8s	01:09:22	50
HiT-DUNS	39.27	0.9529	41.37	0.9660	43.66	0.9774	46.06	0.9801	0.0118s /1.57s	01:34:20	200
DISTA-CSNet	**40.36**	**0.9634**	**42.67**	**0.9755**	**45.24**	**0.9839**	**47.51**	**0.9899**	**0.0114s /1.36s**	**00:16:18**	**20**

TABLE 2. Dataset 2: Knee MRI testing results with 21 MR images.

Algorithms	Compression Ratio								GPU / CPU Testing Time	GPU Training Time	Epochs
	20%		30%		40%		50%				
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM			
ADMM-Net	34.30	0.9385	38.29	0.9385	40.78	0.9665	41.33	0.9775	0.8935s/ 2.32s	02:38:27	200
ISTA-Net+	33.51	0.8533	36.23	0.9089	39.42	0.9502	41.86	0.9723	0.1261s /2.05s	01:58:07	200
ML-CSC	36.93	0.9262	39.68	0.9569	42.04	0.9738	43.44	0.9824	0.0538s/ 2.18s	00:49:35	50
HiT-DUNS	37.14	0.9299	39.85	0.9658	42.19	0.9761	44.09	0.9839	0.0119 /1.67s	01:10:29	200
DISTA-CSNet	**38.06**	**0.9395**	**41.04**	**0.9661**	**43.42**	**0.9799**	**46.21**	**0.9898**	**0.0112s/ 1.34s**	**00:12:15**	**20**

Table 2 illustrates the effectiveness of the proposed DISTA-CSNet model, which was trained on a dataset containing 600 Knee MR images, showcasing its superior performance in terms of PSNR and SSIM across various compression ratios. Remarkably, even at a 5-fold compression, the model achieves an impressive average SSIM value of 0.9395, demonstrating its proficiency in generating high-quality reconstructions. The model's computational

TABLE 3. Dataset 3: Testing results of brain MR images with 123 slices.

Algorithms	Compression Ratio 20% PSNR	20% SSIM	30% PSNR	30% SSIM	40% PSNR	40% SSIM	50% PSNR	50% SSIM	GPU/CPU Testing Time
ADMM-Net	29.29	0.84386	33.00	0.89599	36.02	0.92406	38.95	0.94158	0.9730s /5.4s
ISTA-Net+	30.50	0.8786	33.94	0.9214	36.96	0.9483	39.96	0.9661	0.1528s /4.94s
ML-CSC	31.68	0.8782	35.21	0.9234	38.35	0.9554	41.60	0.9724	0.0732s/3.79s
DISTA-CSNet	**33.32**	**0.8975**	**37.9**	**0.9522**	**41.37**	**0.9749**	**44.50**	**0.9863**	**0.0429s/1.40s**

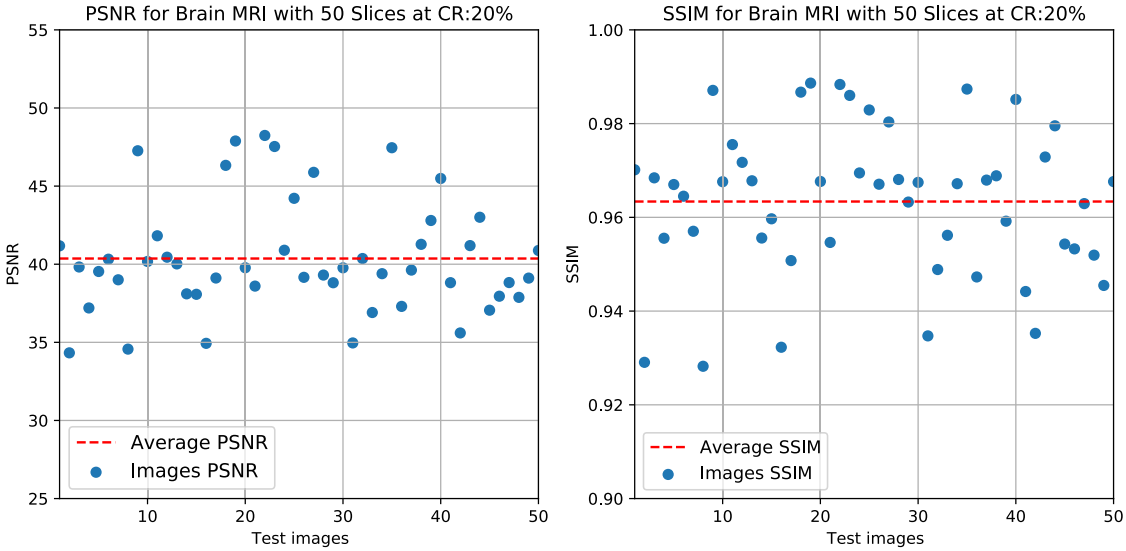

FIGURE 8. PSNR and SSIM for dataset 1 of brain MRI with 20 slices at a compression ratio of 20%.

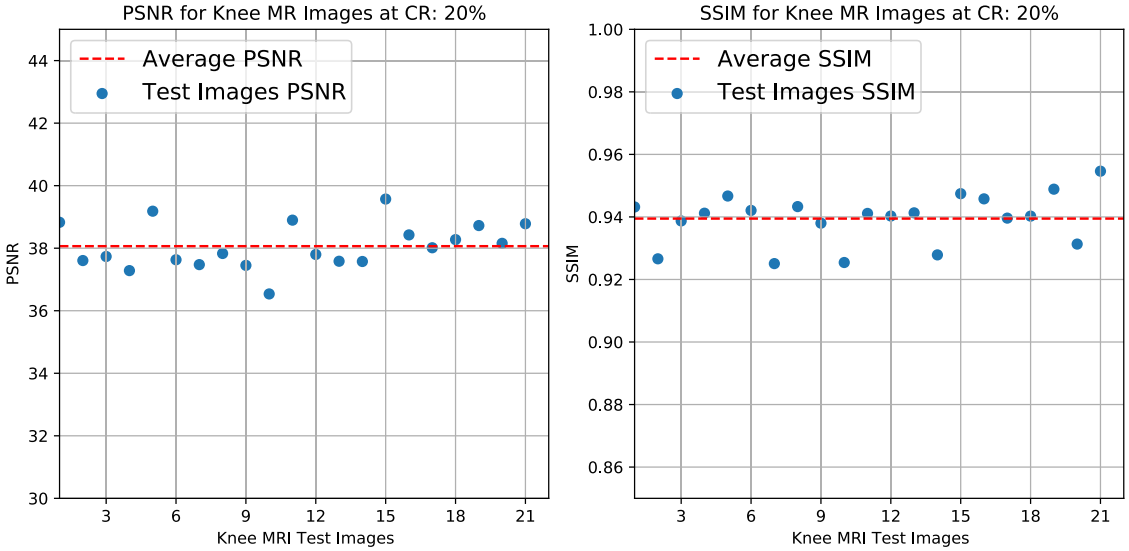

FIGURE 9. PSNR and SSIM for dataset 2 of knee MRI at a compression ratio of 20%.

efficiency during testing was notably improved when utilising a GPU, resulting in computation times within the sub-second range. Nevertheless, the computational time on a CPU has also been significantly reduced, making it viable to test the model without relying on a GPU. Additionally, the training time for the DISTA-CSNet model on Dataset 2 showed substantial reductions, requiring only 20 epochs and approximately 12 minutes to achieve near-optimal

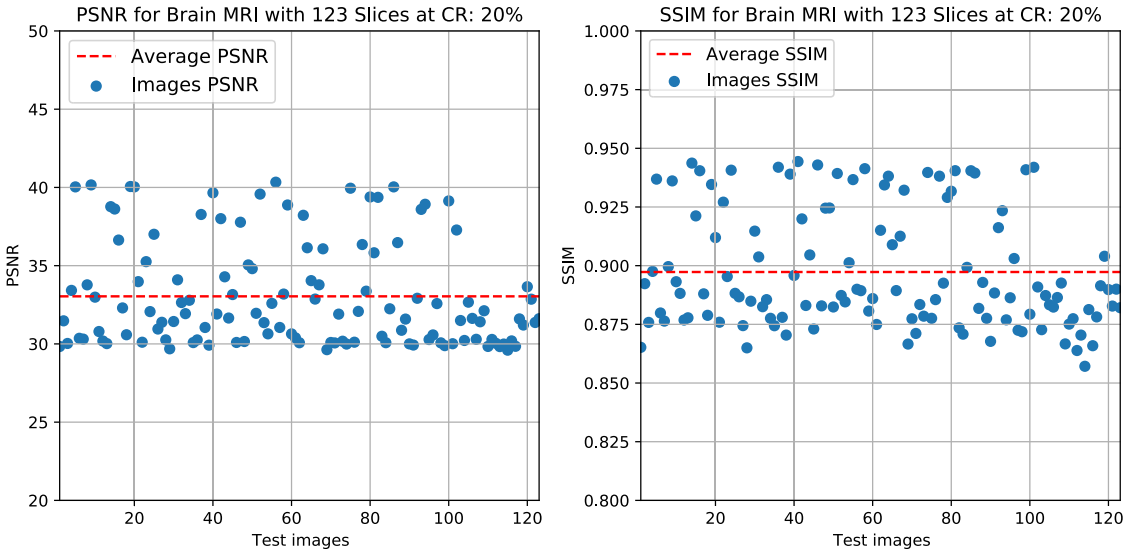

FIGURE 10. PSNR and SSIM for dataset 3 of 123 brain MRI slices at a compression ratio of 20%.

performance. The reduced training time compared to Brain MRI training time was mainly due to fewer MR images being available for training i.e., 600 instead of 800 MRIs. This observation emphasises the proposed method's fast-learning capability compared to other state-of-the-art approaches, further validating its potential for practical and time-efficient medical image reconstruction applications at clinical standards.

Table 3 demonstrates the effectiveness of the proposed DISTA-CSNet model, which was tested using Dataset 3 on the same model that was trained on a dataset comprising 800 Brain MR images, showcasing its exceptional performance concerning PSNR and SSIM across different compression ratios and robustness against varying test data. Even at a 5-fold compression, the model achieves remarkable average SSIM and PSNR scores, underscoring its ability to generate high-quality reconstructions. Notably, during testing, the model exhibited significant improvements in computational efficiency when employed on a GPU, achieving computation times in the sub-second range. However, the computational time on a CPU has also been reduced, making it feasible to test the model without relying solely on a GPU. These finding highlights the proposed method's adaptability and robustness mainly by introducing dropouts, compared to other state-of-the-art approaches, further reinforcing its potential for practical and time-efficient applications in medical image reconstruction.

In Fig. 8, the performance of the trained DISTA-CSNet model is depicted when it was tested on Dataset 1, which consisted of 50 slices of Brain MRI used for testing. The graph shows the PSNR and SSIM values achieved by the model at different compression ratios. Notably, the red dotted line highlights the average PSNR and SSIM attained by the proposed model when the compression ratio was set at only 20%. This emphasises the efficacy of the proposed model in recovering Compressed Sensing (CS) MRI data. Even at a relatively low compression ratio of 20%, the model achieves high PSNR and SSIM values, indicating its ability to produce accurate and high-quality MRI reconstructions. This result demonstrates the robustness and effectiveness of the DISTA-CSNet model in handling CS MRI recovery tasks, making it a promising solution for medical imaging applications that involve compressed data.

Fig. 9 illustrates the performance of the trained DISTA-CSNet model during testing on Dataset 2, comprising 21 MRIs of Knee MRI used for evaluation. The graph showcases the PSNR and SSIM values achieved by the model at various compression ratios. The red dotted line highlights the average PSNR and SSIM obtained by the proposed model when the compression ratio was set to just 20%. This emphasises the model's effectiveness in recovering Compressed Sensing (CS) MRI data. Even at a relatively 5-fold compression, the model can produce accurate and high-quality MRI reconstructions with better PSNR and SSIM. These results underscore the robustness and efficacy of the DISTA-CSNet model in handling CS MRI recovery tasks, rendering it a promising solution for medical imaging applications that deal with compressed data.

Fig. 10 presents the performance of the trained DISTA-CSNet model during testing on Dataset 3, which consists of 123 slices of Brain MRIs used for evaluation. The graph showcases the PSNR and SSIM values achieved by the model at different compression ratios. Of particular significance is the red dotted line, indicating the average PSNR and SSIM obtained by the proposed model at a compression ratio of 20%. This highlights the model's effectiveness in recovering Compressed Sensing (CS) MRI data. Even at a relatively high 5-fold compression, the model

can produce accurate and high-quality MRI reconstructions, with improved PSNR and SSIM values. These results underscore the robustness and efficacy of the DISTA-CSNet model in handling CS MRI recovery tasks, making it a promising solution for medical imaging applications that involve compressed k-space data.

IV. CONCLUSION

The proposed DISTA-CSNet model has introduced data consistency constraints and a flexible tanh-based shrinkage technique, leading to remarkable advancements in terms of reduced training time, requiring only 20 epochs adapt to different datasets for specific applications efficiently. Despite the reduced training time, the model's testing performance significantly improves in both SSIM and PSNR values that are acceptable for clinical applications. Moreover, incorporating dropouts in the proposed model has demonstrated its robustness when confronted with varying datasets, that can significant where dataset is scarce. This resilience is evident from the testing results on Dataset 3, Brain MRI with 123 slices, where the proposed model outperforms other state-of-the-art methods in recovering Compressed Sensing MRI data. These findings highlight the effectiveness and versatility of the DISTA-CSNet model, making it a promising and competitive solution for medical image reconstruction tasks.

V. FUTURE WORK

This research work can be further enhanced by replacing analytical sparsifying transforms (dictionary) by integrating an adaptive dictionary learning mechanism could further elevate recovery outcomes. Moreover, the research suggests extending CS recovery to patch-based reconstruction, facilitating localized and adaptive modelling that is anticipated to yield superior reconstruction quality compared to traditional global modelling approaches. Within the domain of deep learning, the transfer learning methodology could be used in development of more universally applicable CS recovery models capable of adaptation to diverse datasets without necessitating extensive initial training. Additionally, the proposed DISTA-CSNet model, with its demonstrated efficacy, could be optimized for parallel processing using multiple GPUs, thereby accelerating its training process. Furthermore, an intriguing avenue for future exploration lies in the modification of the DISTA-CSNet model to incorporate classification capabilities for biomedical images within the CS framework, potentially offering a holistic solution for both reconstruction and analysis tasks in medical imaging. Through these advancements, the research endeavors to propel the field of CS-MRI towards more efficient, adaptable, and comprehensive solutions with significant implications for medical imaging.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[3] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 14–20, Mar. 2008.

[4] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.

[5] G. Kutyniok and W.-Q. Lim, "Optimal compressive imaging of Fourier data," *SIAM J. Imag. Sci.*, vol. 11, no. 1, pp. 507–546, Jan. 2018.

[6] H. Haider, J. A. Shah, I. M. Qureshi, H. Omer, and K. Kadir, "Compressively sampled MRI recovery using modified iterative-reweighted least square method," *Appl. Magn. Reson.*, vol. 47, no. 9, pp. 1033–1046, Sep. 2016.

[7] H. Haider, J. A. Shah, K. Kadir, and N. Khan, "Sparse reconstruction using hyperbolic tangent as smooth l1-norm approximation," *Computation*, vol. 11, no. 1, p. 7, Jan. 2023.

[8] J. A. Jahanshahi, H. Danyali, and M. S. Helfroush, "Compressive sensing based the multi-channel ECG reconstruction in wireless body sensor networks," *Biomed. Signal Process. Control*, vol. 61, Aug. 2020, Art. no. 102047.

[9] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 521–538, Mar. 2020.

[10] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1828–1837.

[11] A. Wahid, J. A. Shah, A. U. Khan, M. Ahmed, and H. Razali, "Multi-layer basis pursuit for compressed sensing MR image reconstruction," *IEEE Access*, vol. 8, pp. 186222–186232, 2020.

[12] S. Saadain, J. A. Shah, A. Wahid, and K. Kadir, "Comparative analysis of ML-CSC based CS-MRI framework with state of the art," in *Proc. IEEE 8th Int. Conf. Smart Instrum., Meas. Appl. (ICSIMA)*, Sep. 2022, pp. 175–180.

[13] Q. Huang, Y. Xian, D. Yang, H. Qu, J. Yi, P. Wu, and D. N. Metaxas, "Dynamic MRI reconstruction with end-to-end motion-guided network," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101901.

[14] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, "A review of medical image data augmentation techniques for deep learning applications," *J. Med. Imag. Radiat. Oncol.*, vol. 65, no. 5, pp. 545–563, Aug. 2021.

[15] Z. Fabian, R. Heckel, and M. Soltanolkotabi, "Data augmentation for deep learning based accelerated MRI reconstruction with limited data," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 3057–3067.

[16] W. Ren, A. H. Bashkandi, J. A. Jahanshahi, A. Qasim M. AlHamad, D. Javaheri, and M. Mohammadi, "Brain tumor diagnosis using a step-by-step methodology based on courtship learning-based water strider algorithm," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104614.

[17] M. Yaqub, F. Jinchao, S. Ahmed, K. Arshid, M. A. Bilal, M. P. Akhter, and M. S. Zia, "GAN-TL: Generative adversarial networks with transfer learning for MRI reconstruction," *Appl. Sci.*, vol. 12, no. 17, p. 8841, Sep. 2022.

[18] F. Knoll, K. Hammernik, E. Kobler, T. Pock, M. P. Recht, and D. K. Sodickson, "Assessment of the generalization of learned image reconstruction and the potential for transfer learning," *Magn. Reson. Med.*, vol. 81, no. 1, pp. 116–128, Jan. 2019.

[19] J. Lv, G. Li, X. Tong, W. Chen, J. Huang, C. Wang, and G. Yang, "Transfer learning enhanced generative adversarial networks for multi-channel MRI reconstruction," *Comput. Biol. Med.*, vol. 134, Jul. 2021, Art. no. 104504.

[20] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[21] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanawala, G. Zaharchuk, L. Xing, and J. M. Pauly, "Deep generative adversarial neural networks for compressive sensing MRI," *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 167–179, Jan. 2019.

[22] G. Luo, N. Zhao, W. Jiang, E. S. Hui, and P. Cao, "MRI reconstruction using deep Bayesian estimation," *Magn. Reson. Med.*, vol. 84, no. 4, pp. 2246–2261, Oct. 2020.

[23] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Jun. 2010, pp. 399–406.
[24] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018.
[25] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.
[26] 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," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, Jun. 2018.
[27] C. Metzler, A. Mousavi, and R. Baraniuk, "Learned D-AMP: Principled neural network based compressive image recovery," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–12.
[28] B. Shi, Y. Wang, and D. Li, "Provable general bounded denoisers for snapshot compressive imaging with convergence guarantee," *IEEE Trans. Comput. Imag.*, vol. 9, pp. 55–69, 2023.
[29] B. Shi, D. Li, Y. Wang, Y. Su, and Q. Lian, "Provable deep video denoiser using spatial–temporal information for video snapshot compressive imaging: Algorithm and convergence analysis," *Signal Process.*, vol. 214, Jan. 2024, Art. no. 109236.
[30] B. Shi and K. Liu, "Regularization by multiple dual frames for compressed sensing magnetic resonance imaging with convergence analysis," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 11, pp. 2136–2153, Nov. 2023.
[31] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.
[32] J. Zhang, Z. Zhang, J. Xie, and Y. Zhang, "High-throughput deep unfolding network for compressive sensing MRI," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 4, pp. 750–761, Jun. 2022.
[33] J. Sulam, A. Aberdam, A. Beck, and M. Elad, "On multi-layer basis pursuit, efficient algorithms and convolutional neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 1968–1980, Aug. 2020.
[34] A. Keith. (1999). *The Whole Brain Atlas*. [Online]. Available: http://www.med.harvard.edu/AANLIB/home.html
[35] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1050–1059.
[36] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

HASSAAN HAIDER received the B.S. degree in electronic engineering from International Islamic University Islamabad (IIUI), Pakistan, in 2008, the M.S. degree in electrical engineering from the National University of Science and Technology (NUST), Pakistan, in 2014, and the Ph.D. degree from IIUI, in 2024.

From 2008 to 2009, he worked in networks/systems with TechnoEd Pvt. Ltd. He was a Research Assistant with the National Institute of Electronics, Islamabad, from 2009 to 2010. Since 2010, he has been the Laboratory Engineer with the Department of Electrical Engineering, IIUI. He has published three journal research articles and presented three research papers at international conferences. His research interests include compressed sensing, sparse signal processing, machine learning, and deep learning. He is registered with Pakistan Engineering Council. He has served as a Reviewer for IEEE Access and MDPI journals.

JAWAD ALI SHAH (Senior Member, IEEE) received the B.S. degree in electrical engineering and the M.S. degree in telecom engineering from UET Peshawar, in 2001 and 2009, respectively, and the Ph.D. degree in electronic engineering from International Islamic University, Pakistan.

He has been a Visiting Researcher with the University of Colorado Anschutz Medical Campus, Denver, USA, through the HEC IRSIP Program. He is currently working with world leading Cloud Computing Service Provider. He has over eight years of experience in the telecom industry and has received professional training in Pakistan, China, Singapore, the USA, and Malaysia. His research interests include sparse signal processing, compressed sensing, machine learning, cloud computing, and biomedical signal processing.

UMER JAVEED is currently pursuing the Ph.D. degree in electronic engineering (machine learning and computer vision). He is a Machine Learning Engineer in a reputable medical technology development firm. He has actively supervised many postgraduate students and has provided consultancy to financial firms and other research and development organizations.

SHARJEEL ABID BUTT received the bachelor's degree in electrical engineering from Air University, Islamabad, Pakistan, in 2009, and the M.S. degree in electronics engineering from International Islamic University (IIU) Islamabad, in 2013, where he is currently pursuing the Ph.D. degree in electronics engineering. He is an Instructor with IIU Islamabad. His research interests include speech recognition and processing, signal and image processing, machine learning, and evolutionary heuristic techniques.

KUSHSAIRY A. KADIR (Senior Member, IEEE) received the Bachelor of Science degree in engineering from the University of the West of England, Bristol, in 1998, the Master of Science degree in mechatronic from International Islamic University, Malaysia, in 2007, and the Ph.D. degree in electronic and electrical engineering from Strathclyde University, in 2012.

He is currently an Associate Professor with the Electrical Technology Section, Universiti Kuala Lumpur British Malaysian Institute. He is also the Dean of Universiti Kuala Lumpur British Malaysian Institute. His research interest includes signal and image processing. He served as the Past Chair for the IEEE IMS Malaysia Chapter.