

## Article

# Joint Image Reconstruction and Super-Resolution for Accelerated Magnetic Resonance Imaging

Wei Xu <sup>1,2</sup>, Sen Jia <sup>1</sup>, Zhuo-Xu Cui <sup>1</sup>, Qingyong Zhu <sup>1</sup>, Xin Liu <sup>1</sup>, Dong Liang <sup>1,\*</sup> and Jing Cheng <sup>1,\*</sup>

<sup>1</sup> Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; wei.xu@siat.ac.cn (W.X.); sen.jia@siat.ac.cn (S.J.); zx.cui@siat.ac.cn (Z.-X.C.); qy.zhu@siat.ac.cn (Q.Z.); xin.liu@siat.ac.cn (X.L.)

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 101408, China

\* Correspondence: dong.liang@siat.ac.cn (D.L.); jing.cheng@siat.ac.cn (J.C.)

**Abstract:** Magnetic resonance (MR) image reconstruction and super-resolution are two prominent techniques to restore high-quality images from undersampled or low-resolution k-space data to accelerate MR imaging. Combining undersampled and low-resolution acquisition can further improve the acceleration factor. Existing methods often treat the techniques of image reconstruction and super-resolution separately or combine them sequentially for image recovery, which can result in error propagation and suboptimal results. In this work, we propose a novel framework for joint image reconstruction and super-resolution, aiming to efficiently image recovery and enable fast imaging. Specifically, we designed a framework with a reconstruction module and a super-resolution module to formulate multi-task learning. The reconstruction module utilizes a model-based optimization approach, ensuring data fidelity with the acquired k-space data. Moreover, a deep spatial feature transform is employed to enhance the information transition between the two modules, facilitating better integration of image reconstruction and super-resolution. Experimental evaluations on two datasets demonstrate that our proposed method can provide superior performance both quantitatively and qualitatively.

**Keywords:** MR imaging; multi-task; image reconstruction; super-resolution; deep learning



**Citation:** Xu, W.; Jia, S.; Cui, Z.-X.; Zhu, Q.; Liu, X.; Liang, D.; Cheng, J. Joint Image Reconstruction and Super-Resolution for Accelerated Magnetic Resonance Imaging. *Bioengineering* **2023**, *10*, 1107. <https://doi.org/10.3390/bioengineering10091107>

Academic Editor: Jia Guo

Received: 1 August 2023

Revised: 7 September 2023

Accepted: 8 September 2023

Published: 21 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Magnetic resonance imaging (MRI) has been widely used in modern medical diagnosis as a non-invasive imaging modality. The raw data acquired in MRI are located in k-space; however, the imaging time for MR examination is relatively long due to the data acquisition scheme in k-space and the time for signal recovery. Prolonged scan time not only leads to patient discomfort but also affects the image quality, particularly in high-resolution (HR) imaging that demands even longer scan time. High-resolution images, due to their rich anatomical information and textural details, have the potential to significantly enhance the accuracy of medical diagnosis. Therefore, reducing imaging time and improving image quality are of paramount importance in MR imaging.

Acquiring less data in k-space is one of the most important ways to accelerate MR acquisition, and high-quality images can be obtained through image restoration from the partial k-space data. Generally, there are two mainstream approaches for image restoration: image reconstruction (Rec) and super-resolution (SR). The former aims to recover artifact-free images from undersampled k-space data, while the latter infers high-resolution (HR) images from low-resolution (LR) observations. Undersampling in k-space preserves imaging resolution in the stage of data acquisition; however, it introduces artifacts in the image due to the violation of the Nyquist sampling theorem. In the case of SR, although data are sampled with the Nyquist sampling ratio, the absence of data fidelity in the high frequency of k-space is prone to generating fake structures, creating puzzles in the images. Furthermore, image reconstruction usually takes place in k-space domain [1,2] or

hybrid domain [3,4], while super-resolution is performed in image domain. To improve image quality, prior information is employed in both tasks, such as sparse priors [3–5]. Although optimization-based techniques have achieved numerous successes in MR Rec and SR, there are still some limitations in practice, such as long restoration time, heavy computational burdens, and challenges in hyper-parameter selection, which mainly include sparse transformations, regularization parameters, and the number of iterations. Moreover, the use of simplistic priors restricts the performance of MR restoration in achieving better image quality.

Recently, deep learning (DL)-based methods have made significant advancements and demonstrated remarkable improvements in various fields [6–10]. Specifically, DL has gained popularity in image restoration tasks and has shown great potential in enhancing image quality. In MR reconstruction, neural networks are employed to eliminate undersampled artifacts in the image domain or utilized as regularization in an iterative algorithm with end-to-end training [11–27]. For SR, neural networks learn the mapping between LR and HR image pairs [28–37]. Research in SR focuses on designing advanced network architectures to enhance the structures of interest and preserve sensitive detail information in the generated HR images. However, most studies generate LR images by degrading in the image domain, overlooking the intrinsic properties of MR imaging, where raw data are acquired in k-space using multi-channel coils.

Although the implementation details are different, both Rec and SR tasks aim to obtain high-quality images with fine details from partially acquired k-space data. The Rec task ensures data consistency, while the SR task focuses on generating perceptual structures and improving the spatial resolution of the image. By leveraging the strengths of both tasks, it is possible to accelerate MR imaging and achieve higher image quality. Currently, there are two common approaches for integrating Rec and SR: performing the tasks separately or combining them sequentially using two separate sub-networks for Rec and SR. However, these approaches have limitations. Performing the tasks separately can lead to information loss and error propagation, while sequential approaches overlook the correlations between the two tasks. In this work, we propose a novel multi-task framework for joint MR Rec and SR. Instead of the sequential approaches of combining Rec and SR, we introduce a transmission of features from the Rec task to the SR task. This feature transmission facilitates the integration of information from both tasks, allowing the SR task to benefit from the enhanced features obtained during the Rec task. By jointly considering Rec and SR tasks and incorporating feature transmission, we aim to improve the overall performance of MR image restoration, resulting in higher image quality and faster MR imaging.

The main contributions of this work are as follows:

- We introduce a novel multi-task framework that integrates the model-based deep learning reconstruction (Rec) method with the super-resolution (SR) method, resulting in a reduction of the scan time for MR imaging and an enhancement in image quality.
- The proposed method performs image reconstruction and super-resolution simultaneously, demonstrating superior performance compared to various sequential combinations of state-of-the-art MR Rec and SR methods, as well as other multi-task approaches.
- We validate the efficiency of the proposed method on different datasets, including a 2D brain dataset and a 3D T1w brain and neck vessel wall dataset.

## 2. Related Works

### 2.1. DL-Based MR Image Reconstruction

DL-based MR reconstruction methods aim to remove artifacts caused by undersampling. In 2016, Wang et al. [11] applied CNN to MR reconstruction, setting a precedent for DL-based reconstruction. Subsequent studies focused on advanced networks to learn the mapping between the zero-filled image and the fully sampled one, and combining the imaging model with deep networks. For instance, Zbontar J et al. [12] proposed a Unet-based network, and Lee et al. [13] proposed a ResNet-based network for MR image reconstruction.

The main drawback of the data-driven approach is its lack of interpretability, as it does not consider the problem from the perspective of the MR imaging model. Model-driven methods then gained lots of attention in MR image reconstruction [14–27]. Model-based reconstruction algorithms have inspired network innovations in model-driven methods to make them more flexible and interpretable [38]. For example, ISTA-Net [15] and its variants [17] unroll the Iterative Shrinkage Threshold Algorithm (ISTA) to solve the problem of MR image reconstruction. Aggarwal et al. [16] proposed integrating a CNN-based denoiser into model-based reconstruction. ADMM-CSNet [18] unrolls the alternating direction multiplier method (ADMM) [39] for compressed sensing applications. Advanced networks, such as generative adversarial networks (GAN) [40] and attention mechanisms, are also employed in MR image reconstruction [14,24–27].

## 2.2. DL-Based MR Image Super-Resolution

With the success of deep learning in the field of image classification and recognition, deep learning has also found applications in image SR. The first DL attempt for image SR was the SRCNN [28], and since then, numerous DL-based SR methods have been proposed. For instance, Oktay et al. [41] introduced a residual convolutional neural network (CNN) to recover the resolution of cardiac images. This network handles multiple input data from different observation perspectives, thereby enhancing its performance. To address the challenge of partial loss of image details, Shi et al. [42] extended the SRCNN framework and designed a network that combined multi-scale global residual learning and local residual learning based on shallow network blocks. Local residual learning is utilized to recover high-frequency details of images. The lack of high-quality medical image training samples has been a concern, leading to many SR models suffering from over-fitting or under-fitting. To mitigate this issue, Zhao et al. [43] proposed the channel splitting network (CSN) to alleviate the representation burden of deep models. GAN has also been applied in the field of SR to recover texture details of SR images [29–34]. Building on the concept of GAN, MedSRGAN [44] was introduced for the SR of medical images, including MR images, allowing for the generation of more realistic texture details. Additionally, the architecture of transformer [7] has been applied to image SR [35–37]. Transformer can effectively capture remote and non-local features that facilitate image SR. In conclusion, deep learning has significantly advanced the field of image SR, and the incorporation of various advanced network architectures has contributed to more practical and efficient solutions in image super-resolution.

## 3. Methods

### 3.1. Problem Statement

The multi-coil super-resolution MR forward imaging model can be formed as:

$$b = PFCx + \delta \quad (1)$$

where  $b$  is the acquired undersampled, low-resolution k-space data,  $x$  is the MR image to be recovered,  $P$  is the sampling mask that  $P = MH$ ,  $M$  is the binary undersampling mask,  $H$  is the downsampling operator to crop out the high-frequency components of the k-space domain,  $F$  is the Fourier transform, and  $C$  is the coil sensitivity map;  $\delta$  denotes the measurement error which can be well modeled as noise. Our objective is to reconstruct an artifact-free SR image  $x$  from the given partial k-space data  $b$ .

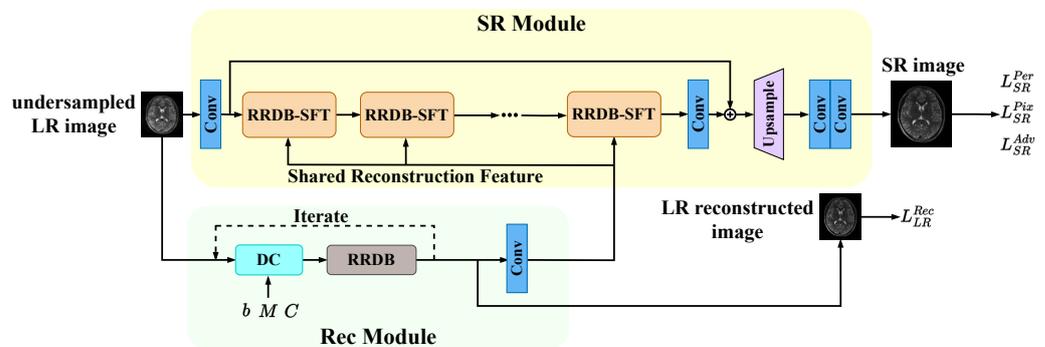
However, Equation (1) is an ill-posed problem making it challenging to directly obtain the image through operator inversion. With the aid of deep learning, we propose a network  $G$  with parameter  $\theta$  to learn the best estimate of the inversion process. The network  $G$  takes undersampled, low-resolution MR image  $x^{LR}$  as input and generates artifact-free SR image  $x$  and the corresponding fully sampled high-resolution image  $x^{GT}$  as ground truth, and the

loss function  $L$  is used to measure the error between  $x$  and  $x^{GT}$ . Therefore, the training process of the network can be formulated in the following supervised manner:

$$\theta_G^* = \arg \min_{\theta_G} \mathbb{E}_x [L(G(x^{LR}; b; M; C; \theta_G), x^{GT})] \tag{2}$$

### 3.2. Overall Architecture

Our proposed network, shown in Figure 1, can simultaneously perform image super-resolution and MR image reconstruction. The network takes an undersampled, low-resolution MR image as input, along with the input coil sensitivity map, undersampling mask, and undersampled k-space, and outputs super-resolution MR images corresponding to fully sampled, high-resolution images as ground truth. The network comprises two main components: an SR Module and a Rec Module. The SR Module extracts features from the undersampled and low-resolution MR image and restores the resolution of the image. The Rec Module performs the MR image reconstruction process in LR space. Since the SR Module depends on the output of the Rec Module, we first introduce the Rec Module, followed by the SR Module, and finally the loss function used to optimize the network.



**Figure 1.** The overall architecture of our proposed network. RRDB: Residual in Residual Dense Block, SFT: Spatial Feature Transform, DC: Data Consistency.

#### 3.2.1. Rec Module

For the Rec Module, the reconstruction problem can be modeled as solving the following minimization problem:

$$x_{rec}^* = \arg \min_{x_{rec}} \frac{1}{2} \|Ax_{rec} - b\|_2^2 + R(x_{rec}) \tag{3}$$

$x_{rec}$  is the LR image to be reconstructed,  $R(x_{rec})$  is the regularization term,  $b$  is the k-space data, and  $A = MFC$  is the coding matrix in MRI.

Following the strategy of unrolling, we select the proximal gradient descent (PGD) algorithm [45] to solve the problem (3), resulting the following iterations:

$$\begin{aligned} x_{rec}^{k+1} &= prox_{R,\eta} (x_{rec}^k - \eta \nabla \frac{1}{2} \|Ax_{rec} - b\|_2^2) \\ &= prox_{R,\eta} (x_{rec}^k - \eta A^H (Ax_{rec} - b)) \end{aligned} \tag{4}$$

where  $\eta$  is a step size and the proximal operator  $prox$  with respect to the function  $R$  is defined as follows:

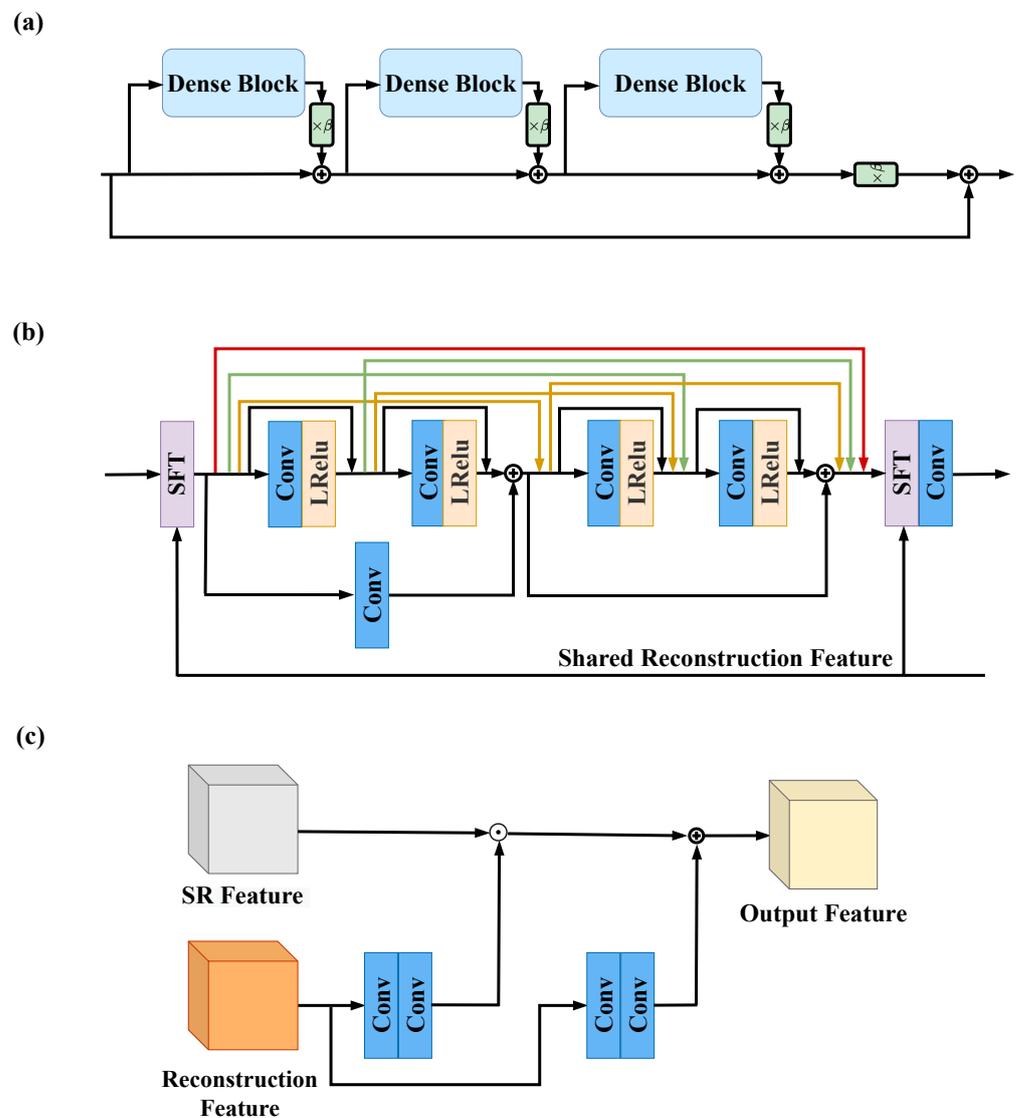
$$prox_{R,\tau}(p) = \arg \min_p \left\{ R(t) + \frac{\|t - p\|_2^2}{2\tau} \right\} \tag{5}$$

The proximal operator  $prox_{R,\tau}(\cdot)$  is a nonlinear operator that can be replaced by a trainable network  $\Gamma$ , so the final iterative process is as follows:

$$\begin{cases} s^k &= x_{rec}^k - \eta A^H(Ax_{rec}^k - b) \\ x_{rec}^{k+1} &= \Gamma(s^k) \end{cases} \quad (6)$$

We make step size  $\eta$  a trainable parameter and then the iterative process in (6) is refined into a trainable network, called PGD Net. In PGD Net, the first step involves data consistency, which enforces data fidelity. For the proximal operation, we employ the Residual in Residual Dense Block (RRDB) [31], as shown in Figure 2a. To enhance the stability of the Reconstruction Module, we utilize the deep equilibrium model [46] with Jacobian-Free back-propagation [47] to ensure the convergence of iterations and reduce the training memory.

The PGD Net will generate LR reconstructed images after completing the iteration process. The reconstructed LR image is represented as a reconstruction feature by a convolutional layer, and finally transmitted to the SR module.



**Figure 2.** (a) The architecture of Residual in Residual Dense Block (RRDB).  $\beta = 0.2$  is a scaling parameter. (b) The architecture of Dense Block in RRDB. (c) The detailed structure of Spatial Feature Transform (SFT) layer.

### 3.2.2. SR Module

The SR Module comprises two parts: feature extraction and upsampling recombination. In the feature extraction part, an initial convolutional layer is utilized to represent the input image as a high-dimensional feature map, facilitating subsequent deep feature extraction. Following the convolutional layer, 23 SR blocks are connected. These blocks can be any feature extraction blocks, and in this work, we use the architecture of RRDB. The RRDB benefits from the combination of residual network [48] and dense connections [49], enhancing its ability to extract complex features and patterns in images, thereby improving the model’s super-resolution capability. In addition, we add two skip connections in each Dense Block of RRDB to enhance the mobility of the features. To enable the SR Module to receive the reconstruction features from the Rec Module, we configure the Spatial Feature Transform (SFT) layer [30] within the RRDB, thus constituting the RRDB-SFT (shown in Figure 2). The SFT layer performs an affine transformation on the intermediate features, with the transformation parameters obtained by transforming two convolutional layers under an additional prior condition. Here, the prior condition is the reconstruction feature from the Rec Module. To improve efficiency, the reconstruction feature is broadcasted to all SFT layers, meaning that all SFT layers share the same reconstruction feature. This approach better guides the SR Module with the assistance of the Rec Module, resulting in a higher-quality image recovery process.

In the upsampling recombination part, there is an upsampling operation on the low-resolution feature map obtained earlier, and the upsampling method used here is nearest neighbor interpolation. The following two convolutional layers are used to suppress the artifacts caused by the upsampling operation and recombine the feature map to generate the final SR image.

### 3.2.3. Loss Function

To reduce the average pixel difference between recovered images  $x$  and Ground Truth  $x^{GT}$ , pixel loss is introduced:

$$L_{SR}^{Pix} = \mathbb{E}_x ||G(x^{LR}) - x^{GT}||_1 \tag{7}$$

Perceptual loss [50] can enhance the semantic correlation between recovered images  $x$  and Ground Truth  $x^{GT}$ , and it contributes to the generation of more realistic texture details [29], which is defined as follows:

$$L_{SR}^{Per} = \mathbb{E}_x ||\phi(G(x^{LR})) - \phi(x^{GT})||_1 \tag{8}$$

where  $\phi$  is a pre-trained VGG-19 network [51].

To enhance the contribution of MR image reconstruction, reconstruction loss  $L_{LR}^{Rec}$  is introduced:

$$L_{LR}^{Rec} = \mathbb{E}_{x_{rec}} ||x_{rec} - x_{LR}^{GT}||_1 \tag{9}$$

where  $x_{rec}$  is the output of PGD Net;  $x_{LR}^{GT}$  is the image obtained by inverse Fourier transform of the LR k-space without undersampling.

GAN can facilitate the network to generate realistic texture details close to the Ground Truth image [29]. We use the network  $G$  as the generator and VGG-19 network [51] as the discriminator  $D$ , and let them perform adversarial training to improve the realism of generated SR images. Therefore, the adversarial loss of generator  $L_{SR}^{Adv}$  and the adversarial loss of discriminator  $L_{SR}^{Dis}$  are introduced:

$$L_{SR}^{Adv} = -\mathbb{E}_x [\log D(G(x^{LR}))] \tag{10}$$

$$L_{SR}^{Dis} = -\mathbb{E}_x [\log(1 - D(x))] - \mathbb{E}_{x^{GT}} [\log D(x^{GT})] \tag{11}$$

Therefore, the overall loss function of the generator  $G$  is as follows:

$$L^G = L_{SR}^{Per} + \gamma_{Pix} L_{SR}^{Pix} + \gamma_{Rec} L_{LR}^{Rec} + \gamma_{Adv} L_{SR}^{Adv} \quad (12)$$

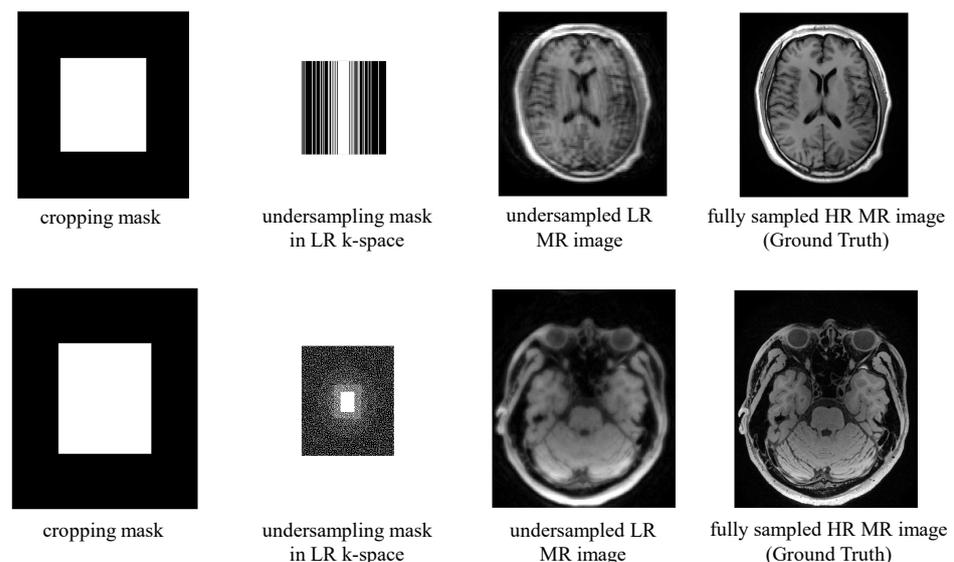
where  $\gamma_{Pix}$ ,  $\gamma_{Rec}$ ,  $\gamma_{Adv}$  are trade-off parameters of different losses.

### 3.3. Datasets

Two sets of data were used to evaluate the feasibility of the proposed method. One is a set of 2D multi-contrast brain data; these contrasts include T1w, T2w, T1 FLAIR and T2 FLAIR. The other is a set of 3D T1w brain and neck vessel wall imaging (VWI) data. The two datasets were acquired with Cartesian, rectilinear k-space sampling. Informed consent was obtained from the imaging subjects in compliance with the Institutional Review Board (IRB) policy.

The fully sampled HR k-space data were acquired, and we simulated to generate LR undersampled data by cropping and undersampling. The LR-HR paired images were generated from the corresponding LR-HR k-space data through coil combination with ESPIRiT map. This process allowed us to obtain LR images and their corresponding HR counterparts for training and evaluation purposes.

The two datasets are acquired on a 3.0T MR imaging system (uMR790, United Imaging Healthcare, Shanghai, China). The brain dataset contains 1900 fully sampled k-space slices acquired from 12 subjects. To reduce the computational complexity, we used the coil compression technique GCC [52] to compress the coils of the original data into 12. Each k-space slice has dimensions of  $12 \times 256 \times 232$  (coils  $\times$  FE  $\times$  PE). We used a 1D 2.5x Gaussian undersampling mask and 2x SR scaling factor to generate the undersampled LR brain data. Since cropping in the FE direction cannot accelerate MRI, the final imaging acceleration is  $2 \times 2.5 = 5x$ , and 1600 slices were used for training, 150 for validation, and 150 for testing. The 3D VWI data contains 1472 fully sampled k-space slices, which were acquired from four healthy volunteers using T1-weighted MATRIX sequences with non-selective excitation. The dimension of the k-space slices of the VWI data is  $18 \times 336 \times 280$  (coils  $\times$  PE  $\times$  SPE). A 2D 4x Poisson undersampling mask and 2x SR scaling factor were used to generate LR data, resulting in a  $2 \times 2 \times 4 = 16x$  acceleration. In the VWI dataset, there are 1104, 184, and 184 slices for the training set, validation set, and test set, respectively. The cropping masks, undersampling masks used to generate the datasets, and the examples of LR-HR image pairs for both datasets are presented in Figure 3.



**Figure 3.** The cropping masks, undersampling masks in LR k-space and examples of undersampled LR MR image and fully sampled HR MR image sample from the brain dataset (upper) and the VWI dataset (lower).

### 3.4. Evaluation Metrics

For a fair quantitative comparison, we use Peak Signal-to-Noise Ratio (PSNR), Structure Similarity Index Measure (SSIM) [53], and Learned Perceptual Image Patch Similarity (LPIPS) [54] to evaluate the quality of the generated images.

The PSNR is calculated based on the mean square error (MSE) between the generated image  $x$  and the Ground Truth image  $y$ . The larger PSNR indicates that the distortion of the generated image is smaller, and the specific calculation formula is as follows:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [x(i, j) - y(i, j)]^2 \quad (13)$$

$$PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad (14)$$

where  $M$  and  $N$  denote the size of the image, and  $MAX$  is the maximum possible pixel value in the image.

The SSIM measures the similarity of the generated image  $x$  and the Ground Truth image  $y$  based on brightness, contrast and structure, and a larger SSIM indicates better quality of the generated image, as calculated by the following formula:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (15)$$

where  $\mu$  and  $\sigma$  represent the mean and variance of the image.  $\sigma_{xy}$  represents the mutual covariance of the two images  $x$  and  $y$ , and  $C_1 = (0.01MAX)^2$  and  $C_2 = (0.03MAX)^2$  are two constants.

The LPIPS is calculated based on the  $L_2$  distance of the generated image  $x$  and the Ground Truth image  $y$  in the feature space of a pre-trained network. This metric is more in accordance with human perception than the above two metrics. The specific calculation formula is as follows:

$$LPIPS = \sum_l \frac{1}{H_l W_l} \sum_{i=1}^{H_l} \sum_{j=1}^{W_l} \left\| w_l \odot (\hat{x}^l(i, j) - \hat{y}^l(i, j)) \right\|_2^2 \quad (16)$$

where  $w_l$  is a channel-wise vector,  $\hat{x}^l$  and  $\hat{y}^l$  are the output feature map of the  $l$ -th layer of a pre-trained network, by convention we use AlexNet [6] and  $H_l$  and  $W_l$  are the size of the feature maps.

### 3.5. Training Details

In our work, we used the ADAM optimizer [55] with default parameters  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  and  $\epsilon = 1 \times 10^{-8}$  to optimize the parameters of the generator and discriminator. Both the generator and discriminator were trained with a learning rate of  $1 \times 10^{-4}$ , and the learning rates were reduced by half after 50k, 100k, 200k, and 300k iterations. The batch size for training is 4, and the sizes of the input undersampled LR MR images for the brain dataset and VWI dataset are  $128 \times 116$  and  $168 \times 140$ , respectively. As for the trade-off parameters of different losses, to ensure the perceptual quality of the generated super-resolution images and to highlight the role of MR image reconstruction, we set  $\gamma_{Rec}$  to 1,  $\gamma_{Pix}$  to  $1 \times 10^{-2}$ , and  $\gamma_{Adv}$  to  $5 \times 10^{-3}$ . All the experiments are implemented by the PyTorch framework on two NVIDIA Tesla V100 GPUs.

## 4. Experimental Results and Analysis

### 4.1. Ablation Study

In this subsection, we present the results of ablation experiments conducted to validate our motivation that integrating Rec into SR can generate images with higher quality. The quantitative results on the brain and VWI datasets are presented in Tables 1 and 2,

respectively. In these two tables, the following notations are used: “SR” and “Recon” represent using only the SR or Rec Module to obtain fully sampled HR images; “Recon-SR” refers to training reconstruction and SR separately; “Recon+SR” denotes the sequential combination of reconstruction and SR with end-to-end training. The main difference between the proposed method and “Recon+SR” is that in our approach, the Rec Module and SR Module are not simply connected in series, but they are interactively connected using SFT layers. This interactive connection allows for more effective information exchange between the two modules, resulting in improved image quality.

**Table 1.** Quantitative results of ablation study on the brain dataset. The average value and standard deviation are given for each metric. The figures in bold are the best results.

Metric	PSNR	SSIM	LPIPS
SR	28.88 ± 2.40	0.8462 ± 0.0446	0.0559 ± 0.0155
Recon	30.23 ± 2.43	0.8483 ± 0.0416	0.1429 ± 0.0178
Recon-SR	29.27 ± 2.45	0.8320 ± 0.0479	0.0537 ± 0.0146
Recon+SR	29.40 ± 2.62	0.8625 ± 0.0403	0.0482 ± 0.0117
Proposed	<b>30.32 ± 2.42</b>	<b>0.8682 ± 0.0418</b>	<b>0.0409 ± 0.0111</b>

**Table 2.** Quantitative results of ablation study on the VWI dataset. The average value and standard deviation are given for each metric. The figures in bold are the best results.

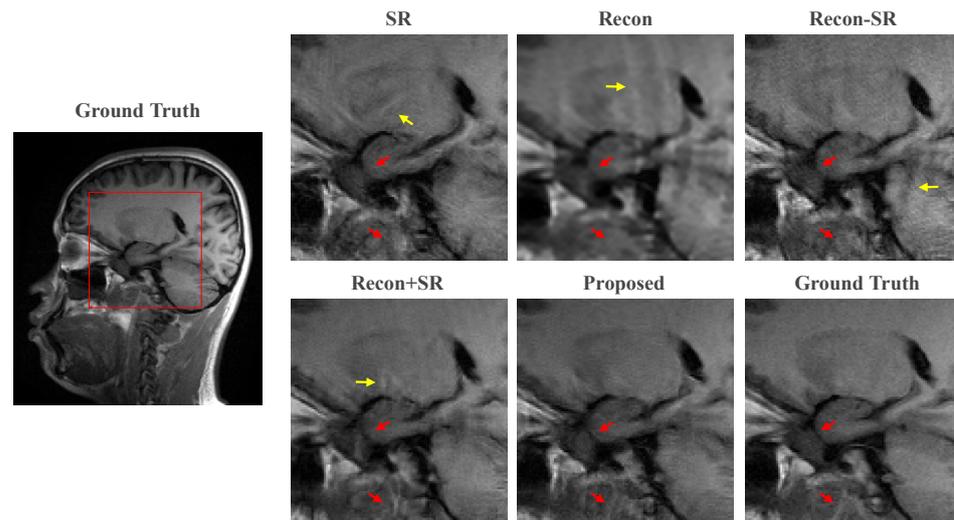
Metric	PSNR	SSIM	LPIPS
SR	31.93 ± 3.11	0.8657 ± 0.0324	0.0568 ± 0.0114
Recon	32.34 ± 4.60	0.8912 ± 0.0355	0.1089 ± 0.0306
Recon-SR	32.32 ± 3.22	0.8750 ± 0.0264	0.0513 ± 0.0102
Recon+SR	32.72 ± 3.00	0.8856 ± 0.0293	0.0425 ± 0.0074
Proposed	<b>32.91 ± 3.03</b>	<b>0.8948 ± 0.0288</b>	<b>0.0383 ± 0.0062</b>

Comparing the results of the SR scheme and the Recon scheme, it can be observed that the PSNR and SSIM of the SR scheme are relatively lower, while the LPIPS is significantly higher than that of the Recon scheme, indicating that the MR images generated by the SR scheme are closer to the original images in terms of visual perception but hold a higher level of distortion.

In our attempt to combine super-resolution with MR image reconstruction, we developed the Recon-SR scheme. However, this scheme performs poorly in terms of PSNR and SSIM compared to the Recon scheme. In particular, it can be seen that there is a substantial decrease in SSIM in the brain dataset, indicating that the Recon-SR scheme did not result in accurate MR images. The reason for this result is that super-resolution and MR image reconstruction are still two independent processes, and the results of reconstruction in the LR space produce errors that are amplified by the SR network. These errors cannot be reduced by optimizing the SR network, and at the same time, the SR network also produces errors, resulting in the accumulation of errors. As a result, the Recon-SR scheme was not effective in improving the accuracy of MR image reconstruction and super-resolution.

The Recon+SR scheme performs better than the Recon-SR scheme in all metrics. This is because the errors generated by the Rec network can be optimized with end-to-end training, alleviating the problem of error accumulation that occurs when the two processes are performed independently. Comparing our proposed method with other schemes, our proposed method performs the best in all metrics, demonstrating that the MR images generated by our network have precise structural information and texture details, making the visual perception close to the original image. These results suggest that the interactive connectivity scheme employed by our network is superior to the scheme that simply connects the Rec module and the SR module in series and that the scheme that combines MR image reconstruction with super-resolution is superior to the scheme that uses either technique alone.

The qualitative comparisons are shown in Figure 4. It can be seen from the zoomed-in images that the proposed method can effectively preserve fine details and exhibits the best consistency with the ground truth. In the case of the only SR task, the restoration tends to produce fake structures, as indicated by the yellow arrow. Furthermore, visible artifacts are present in the image generated by the Recon method. Comparing Recon+SR with Recon-SR, it can be observed that the end-to-end training in Recon+SR yields better results, indicating that the joint training approach benefits image restoration more than separate training. The proposed method demonstrates its capability to enhance image quality and maintain the fidelity of fine details, outperforming the other approaches in the qualitative evaluation.



**Figure 4.** Zoomed-in view of the ablation experiment. The yellow arrows point to the fake structures in the images and the red arrows represent the fine details that can be recovered by our proposed method compared to other methods.

#### 4.2. Comparison Experiment

To comprehensively evaluate the performance of our proposed model, we conducted a series of comparison experiments on the two datasets. Specifically, we compared our method with ten other models, which include nine Recon+SR methods and one state-of-the-art multi-task method called T2Net [56]. For the nine Recon+SR methods, we selected three reconstruction algorithms (CG-SENSE [57], MoDL [16], and DL-ESPIRiT [20]) and three super-resolution algorithms (bicubic, MedSRGAN [44], and BebyGAN [34]). These algorithms were combined sequentially to form the nine comparison methods. To ensure a fair comparison, each of these methods was carefully optimized to guarantee the best performance and ensure equitable evaluations.

The quantitative results of different methods on the two datasets are presented in Tables 3 and 4. These results were obtained on the entire testing dataset. Across both datasets, our proposed method outperforms all other competing methods. The highest PSNR values indicate that our method can generate images with the least distortion, while the highest SSIM values suggest that it retains more global and local structural information. Moreover, the lowest LPIPS scores indicate that our method achieves the closest visual perception to the ground truth.

**Table 3.** Quantitative results of different models on the brain dataset. The average value and standard deviation are given for each metric. The figures in bold are the best results. The average computation time in seconds for all models on the brain dataset is also provided.

Metric	PSNR	SSIM	LPIPS	Computation Time
CG-SENSE+bicubic	22.41 ± 1.90	0.5805 ± 0.0606	0.4425 ± 0.0533	0.0241
CG-SENSE+MedSRGAN	23.04 ± 2.01	0.6038 ± 0.0567	0.1812 ± 0.0396	0.0517
CG-SENSE+BebyGAN	23.33 ± 3.76	0.7151 ± 0.0593	0.1570 ± 0.0354	0.0647
MoDL+bicubic	29.33 ± 2.48	0.8463 ± 0.0397	0.0929 ± 0.0190	0.0659
MoDL+MedSRGAN	29.62 ± 2.47	0.8466 ± 0.0438	0.0467 ± 0.0121	0.0935
MoDL+BebyGAN	29.84 ± 2.64	0.8590 ± 0.0464	0.0476 ± 0.0129	0.1048
DL-ESPIRiT+bicubic	29.44 ± 2.88	0.8538 ± 0.0388	0.0996 ± 0.0240	0.0220
DL-ESPIRiT+MedSRGAN	29.43 ± 2.93	0.8319 ± 0.0488	0.0477 ± 0.0146	0.0486
DL-ESPIRiT+BebyGAN	29.69 ± 3.07	0.8596 ± 0.0455	0.0455 ± 0.0141	0.0591
T2Net	28.05 ± 2.39	0.8236 ± 0.0436	0.0685 ± 0.0174	0.1820
Proposed	<b>30.32 ± 2.42</b>	<b>0.8682 ± 0.0418</b>	<b>0.0409 ± 0.0111</b>	0.0775

**Table 4.** Quantitative results of different models on the VWI dataset. The average value and standard deviation are given for each metric. The figures in bold are the best results. The average computation time in seconds for all models on the VWI dataset is also provided.

Metric	PSNR	SSIM	LPIPS	Computation Time
CG-SENSE+bicubic	26.15 ± 2.95	0.7323 ± 0.1015	0.2952 ± 0.0769	0.0540
CG-SENSE+MedSRGAN	24.71 ± 4.40	0.7616 ± 0.0835	0.1799 ± 0.0416	0.1068
CG-SENSE+BebyGAN	25.31 ± 3.41	0.7886 ± 0.0681	0.1466 ± 0.0366	0.1236
MoDL+bicubic	32.24 ± 3.31	0.8753 ± 0.0284	0.0953 ± 0.0209	0.1375
MoDL+MedSRGAN	32.30 ± 3.03	0.8651 ± 0.0278	0.0448 ± 0.0088	0.1873
MoDL+BebyGAN	32.71 ± 3.46	0.8933 ± 0.0312	0.0481 ± 0.0103	0.2056
DL-ESPIRiT+bicubic	32.49 ± 3.49	0.8727 ± 0.0278	0.0859 ± 0.0192	0.0402
DL-ESPIRiT+MedSRGAN	31.82 ± 4.32	0.8707 ± 0.0307	0.0454 ± 0.0076	0.0902
DL-ESPIRiT+BebyGAN	32.43 ± 3.22	0.8707 ± 0.0260	0.0491 ± 0.0088	0.1073
T2Net	30.75 ± 3.59	0.8322 ± 0.0353	0.0808 ± 0.0191	0.3426
Proposed	<b>32.91 ± 3.03</b>	<b>0.8948 ± 0.0288</b>	<b>0.0383 ± 0.0062</b>	0.1385

For a more intuitive comparison, we provide visual comparisons for each method. In Figures 5 and 6, we illustrate the axial and sagittal views of the 2D brain dataset, respectively. In Figures 7 and 8, we display the brain image and neck image of the 3D VWI dataset. For each method, we show the restorations along with the corresponding error maps. The images are placed on the left, the error maps on the right, and the PSNR and LPIPS metrics of each image are shown at the bottom right corner. From the error maps, it is evident that our proposed method consistently produces high-quality results with improved visual fidelity and better retention of structural information compared to the other methods.

Figures 9 and 10 present the zoomed-in images of the enclosed parts in the ground truth. Our proposed method can faithfully recover the fine structures, as indicated by the red arrows, and successfully preserve the image contrast, as pointed out by the yellow arrow. The high-fidelity restoration achieved by our method in these zoomed-in regions further demonstrates its ability to faithfully recover fine details and enhance image quality.

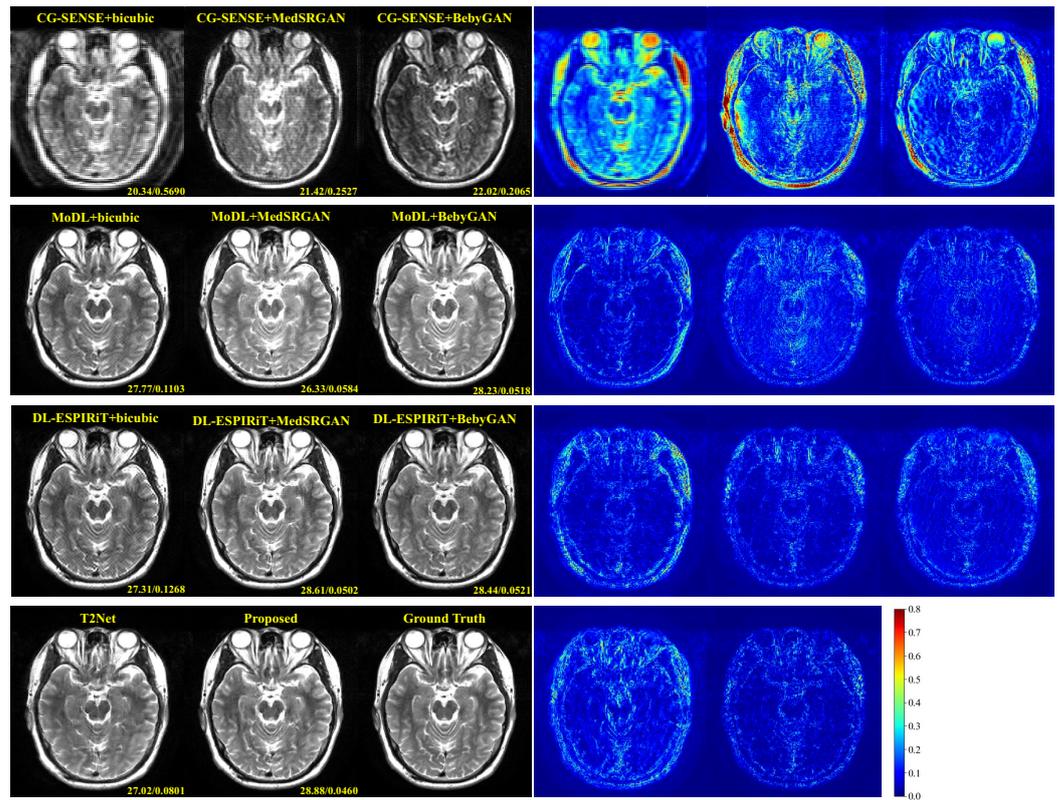


Figure 5. Visual comparison of each method for a slice of axial view in the brain dataset.

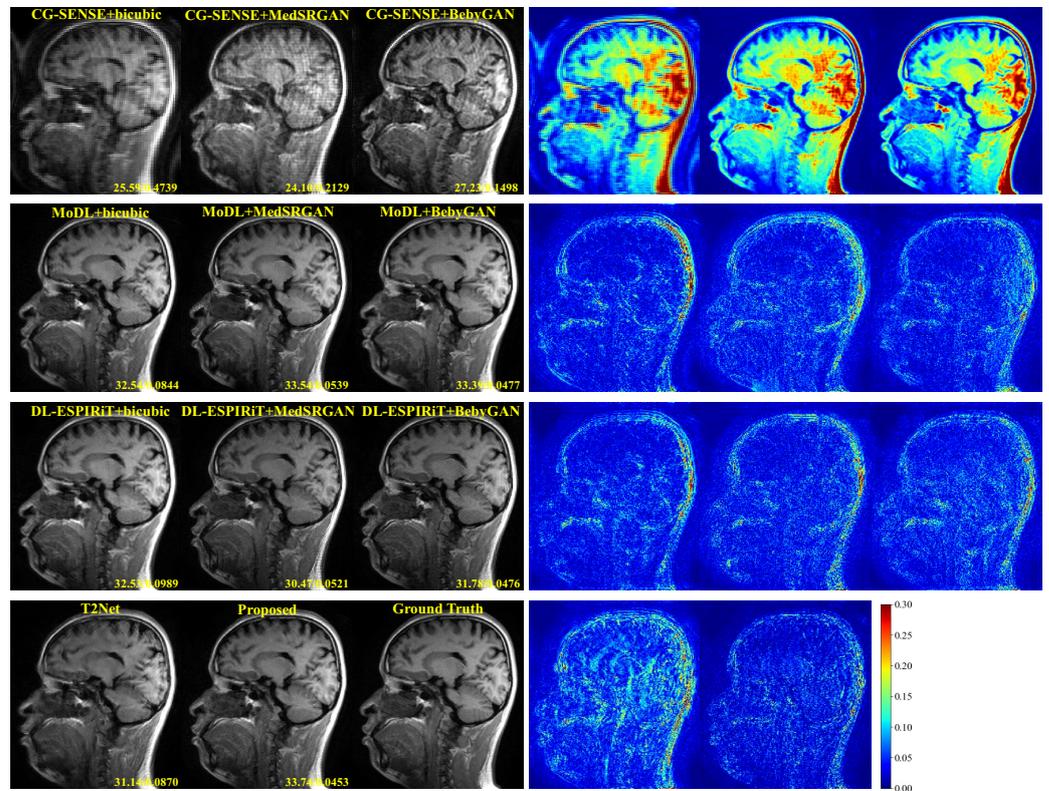


Figure 6. Visual comparison of each method for a slice of sagittal view in the brain dataset.

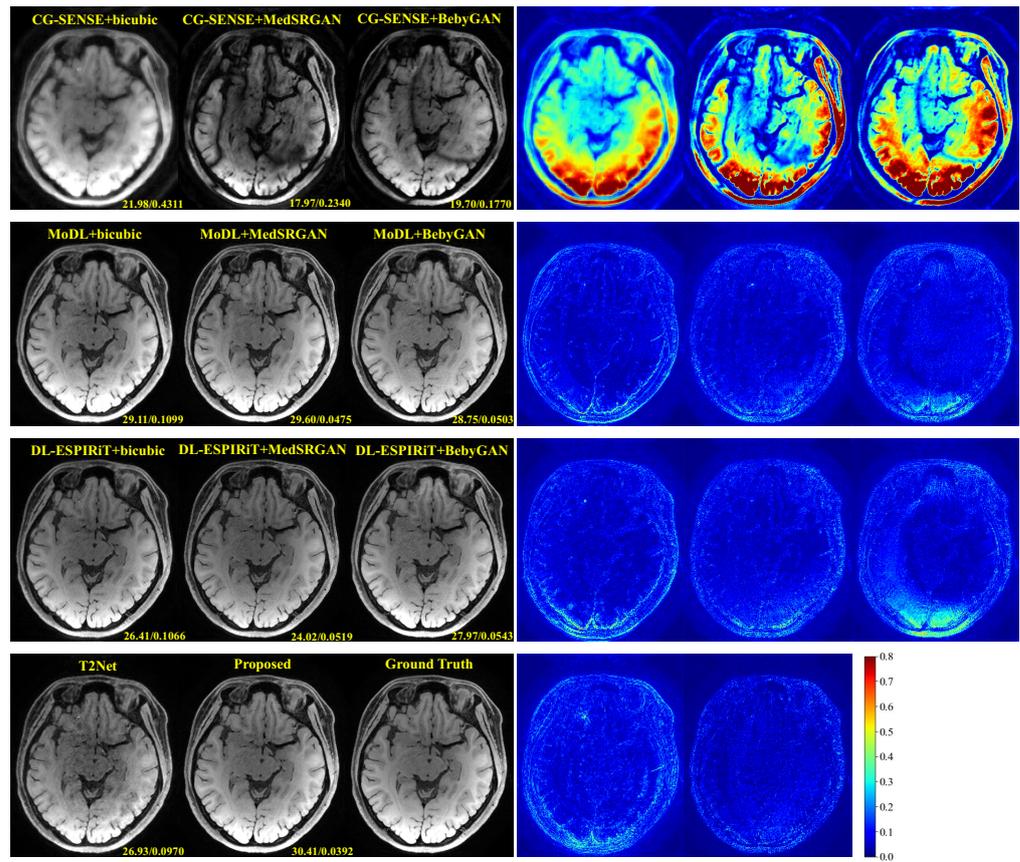


Figure 7. Visual comparison of each method for a brain image in the VWI dataset.

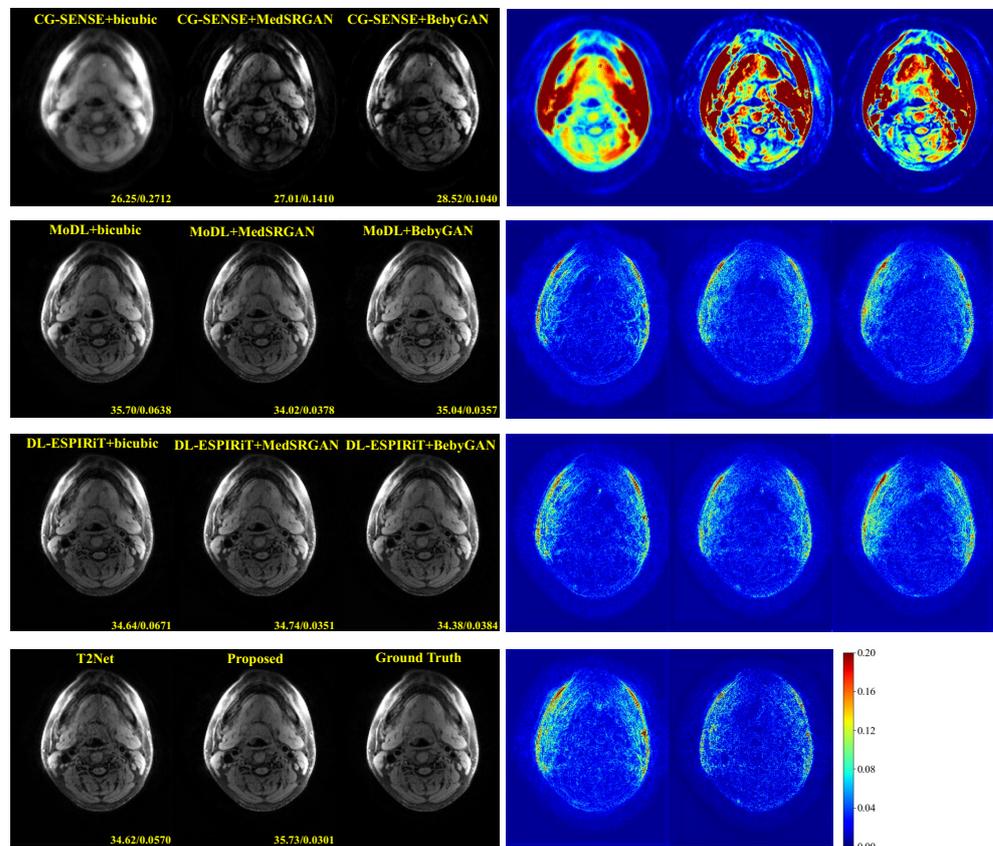


Figure 8. Visual comparison of each method for a neck image in the VWI dataset.

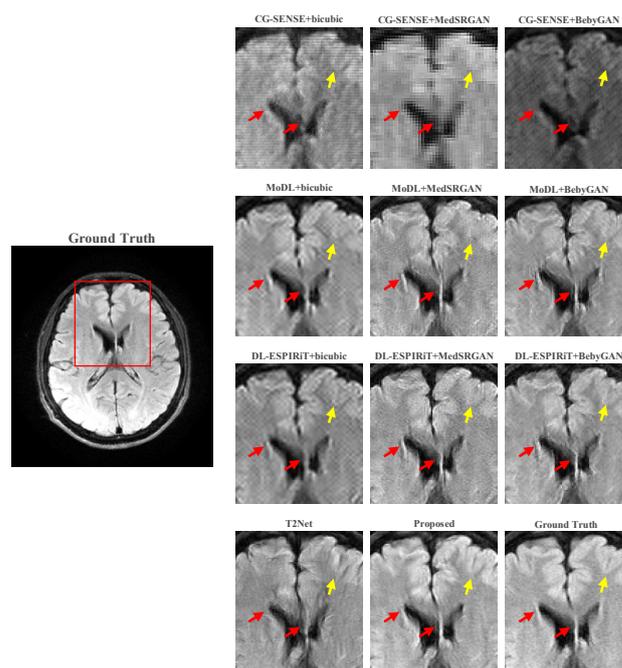


Figure 9. Zoomed-in view of the brain dataset comparison experiment.

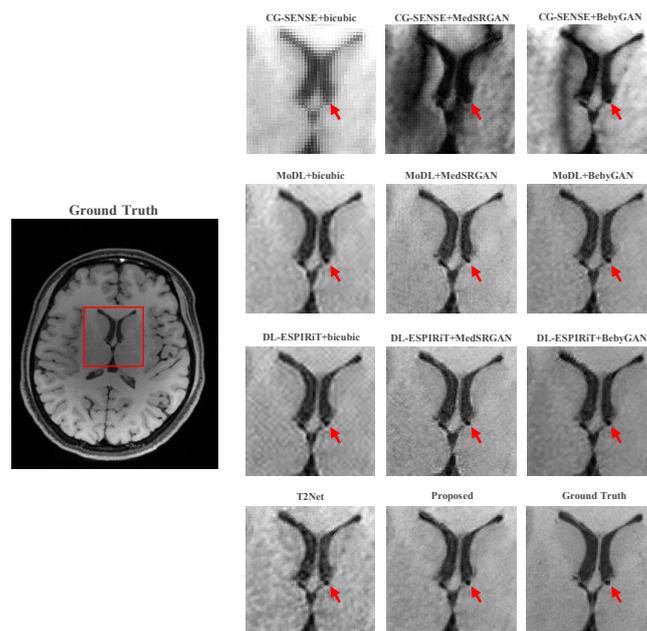


Figure 10. Zoomed-in view of the VWI dataset comparison experiment.

### 5. Discussion

Although our experimental results have demonstrated the validity of our proposed model, there are some limitations that should be noted. Firstly, the VWI dataset used in our study is a 3D dataset, while our model is designed as a 2D model, neglecting the correlation information between slices. Secondly, all the images used in our experiments are brain or neck images, and the performance of our model on images of other parts of the human body remains to be investigated. Finally, while our model has shown promising results in simulated scenarios, its performance in real medical settings needs to be further explored.

To address these limitations, we plan to develop better models that can leverage the correlation information between slices in 3D datasets and consider introducing data from other parts of the human body. Additionally, we aim to validate the effectiveness of

our proposed methods in clinical settings by testing them on real undersampled LR data from a diverse range of participants. This will further validate the potential and utility of our method in medical imaging and contribute to more accurate and efficient MR image restoration for clinical diagnosis and research.

## 6. Conclusions

In this paper, we have introduced a novel multi-task framework for joint MR reconstruction and image super-resolution. Our proposed method incorporates two main modules: the Rec Module responsible for image reconstruction in the LR space and the SR Module which extracts features in the LR space and recovers the SR image. To enhance information exchange between the modules, we introduced the SFT layer to transmit features of the reconstructed image to the SR Module, facilitating super-resolution. The experimental results conducted on 2D brain data and 3D VWI data have demonstrated the superior performance of our proposed method in both quantitative and qualitative evaluations. The comparisons with other state-of-the-art methods have shown that our approach consistently generates high-quality images with enhanced visual fidelity and structural information.

**Author Contributions:** Conceptualization, W.X. and J.C.; methodology, W.X., J.C. and Z.-X.C.; software, W.X. and J.C.; validation, S.J. and Q.Z.; formal analysis, J.C. and Z.-X.C.; investigation, W.X. and J.C.; resources, X.L.; data curation, S.J., Q.Z. and X.L.; writing—original draft preparation, W.X.; writing—review and editing, J.C. and D.L.; visualization, W.X. and Q.Z.; supervision, J.C. and D.L.; project administration, D.L.; funding acquisition, X.L., J.C. and D.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by the National Key R&D Program of China under Grant No. 2021YFF0501503; the National Natural Science Foundation of China under Grant Nos. 62106252, 62206273 and 81830056; the Key Laboratory for Magnetic Resonance and Multimodality Imaging of Guangdong Province under Grant No. 2023B1212060052.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (IRB) of Shenzhen Institutes of Advanced Technology (protocol code No. SIAT-IRB-180315-H0228, 13 March 2018).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** All of our data is unavailable due to privacy or ethical restrictions.

**Acknowledgments:** The authors sincerely appreciate the constructive suggestions provided by the editor and reviewers, which have greatly improved the quality of our paper. The quality of English language in this manuscript has been improved with the help of ChatGPT.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Griswold, M.A.; Jakob, P.M.; Heidemann, R.M.; Nittka, M.; Jellus, V.; Wang, J.; Kiefer, B.; Haase, A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med.* **2002**, *47*, 1202–1210. [[CrossRef](#)] [[PubMed](#)]
2. Lustig, M.; Pauly, J.M. SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magn. Reson. Med.* **2010**, *64*, 457–471. [[CrossRef](#)] [[PubMed](#)]
3. Haldar, J.P.; Hernando, D.; Liang, Z.P. Compressed-sensing MRI with random encoding. *IEEE Trans. Med Imaging* **2010**, *30*, 893–903. [[CrossRef](#)] [[PubMed](#)]
4. Chen, C.; Huang, J. Compressive sensing MRI with wavelet tree sparsity. In Proceedings of the Twenty-Sixth Conference on Neural Information Processing Systems (NeurIPS 2012), Lake Tahoe, NV, USA, 3–6 December 2012; Volume 25.
5. Wang, Y.H.; Qiao, J.; Li, J.B.; Fu, P.; Chu, S.C.; Roddick, J.F. Sparse representation-based MRI super-resolution reconstruction. *Measurement* **2014**, *47*, 946–953. [[CrossRef](#)]
6. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Twenty-Sixth Conference on Neural Information Processing Systems (NeurIPS 2012), Lake Tahoe, NV, USA, 3–6 December 2012; Volume 25.

7. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Volume 30.
8. Hamdia, K.M.; Ghasemi, H.; Bazi, Y.; AlHichri, H.; Alajlan, N.; Rabczuk, T. A novel deep learning based method for the computational material design of flexoelectric nanostructures with topology optimization. *Finite Elem. Anal. Des.* **2019**, *165*, 21–30. [[CrossRef](#)]
9. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
10. Hu, Y.; Yang, J.; Chen, L.; Li, K.; Sima, C.; Zhu, X.; Chai, S.; Du, S.; Lin, T.; Wang, W.; et al. Planning-oriented Autonomous Driving. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 18–22 June 2023; pp. 17853–17862. [[CrossRef](#)]
11. Wang, S.; Su, Z.; Ying, L.; Peng, X.; Zhu, S.; Liang, F.; Feng, D.; Liang, D. Accelerating magnetic resonance imaging via deep learning. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; pp. 514–517.
12. Zbontar, J.; Knoll, F.; Sriram, A.; Murrell, T.; Huang, Z.; Muckley, M.J.; Defazio, A.; Stern, R.; Johnson, P.; Bruno, M.; et al. fastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv* **2018**, arXiv:1811.08839.
13. Lee, D.; Yoo, J.; Tak, S.; Ye, J.C. Deep residual learning for accelerated MRI using magnitude and phase networks. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 1985–1995. [[CrossRef](#)]
14. Yang, G.; Yu, S.; Dong, H.; Slabaugh, G.; Dragotti, P.L.; Ye, X.; Liu, F.; Arridge, S.; Keegan, J.; Guo, Y.; et al. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. *IEEE Trans. Med Imaging* **2018**, *37*, 1310–1321. [[CrossRef](#)]
15. Zhang, J.; Ghanem, B. ISTA-Net: Interpretable Optimization-Inspired Deep Network for Image Compressive Sensing. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1828–1837. [[CrossRef](#)]
16. Aggarwal, H.K.; Mani, M.P.; Jacob, M. MoDL: Model-Based Deep Learning Architecture for Inverse Problems. *IEEE Trans. Med. Imaging* **2019**, *38*, 394–405. [[CrossRef](#)]
17. You, D.; Xie, J.; Zhang, J. ISTA-NET++: Flexible Deep Unfolding Network for Compressive Sensing. In Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME), Shenzhen, China, 5–9 July 2021; pp. 1–6. [[CrossRef](#)]
18. Yang, Y.; Sun, J.; Li, H.; Xu, Z. ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 521–538. [[CrossRef](#)]
19. Huang, Q.; Yang, D.; Wu, P.; Qu, H.; Yi, J.; Metaxas, D. MRI Reconstruction Via Cascaded Channel-Wise Attention Network. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 1622–1626. [[CrossRef](#)]
20. Sandino, C.M.; Lai, P.; Vasanawala, S.S.; Cheng, J.Y. Accelerating cardiac cine MRI using a deep learning-based ESPIRiT reconstruction. *Magn. Reson. Med.* **2021**, *85*, 152–167. [[CrossRef](#)] [[PubMed](#)]
21. Cheng, J.; Cui, Z.X.; Huang, W.; Ke, Z.; Ying, L.; Wang, H.; Zhu, Y.; Liang, D. Learning data consistency and its application to dynamic MR imaging. *IEEE Trans. Med. Imaging* **2021**, *40*, 3140–3153. [[CrossRef](#)] [[PubMed](#)]
22. Huang, W.; Ke, Z.; Cui, Z.X.; Cheng, J.; Qiu, Z.; Jia, S.; Ying, L.; Zhu, Y.; Liang, D. Deep low-rank plus sparse network for dynamic MR imaging. *Med. Image Anal.* **2021**, *73*, 102190. [[CrossRef](#)]
23. Ke, Z.; Huang, W.; Cui, Z.X.; Cheng, J.; Jia, S.; Wang, H.; Liu, X.; Zheng, H.; Ying, L.; Zhu, Y.; et al. Learned low-rank priors in dynamic MR imaging. *IEEE Trans. Med. Imaging* **2021**, *40*, 3698–3710. [[CrossRef](#)] [[PubMed](#)]
24. Qu, W.; Cheng, J.; Zhu, Y.; Liang, D. Deep MR parametric imaging with the learned L+ S model and attention mechanism. *IET Image Process.* **2023**, *17*, 969–978. [[CrossRef](#)]
25. Huang, J.; Xing, X.; Gao, Z.; Yang, G. Swin Deformable Attention U-Net Transformer (SDAUT) for Explainable Fast MRI. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2022, Singapore, 18–22 September 2022; pp. 538–548.
26. Yang, H.; Li, J.; Lui, L.M.; Ying, S.; Zeng, T. Fast MRI Reconstruction via Edge Attention. *Commun. Comput. Phys.* **2023**, *33*, 1409–1431. [[CrossRef](#)]
27. Lyu, J.; Sui, B.; Wang, C.; Tian, Y.; Dou, Q.; Qin, J. DuDoCAF: Dual-Domain Cross-Attention Fusion with Recurrent Transformer for Fast Multi-contrast MR Imaging. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2022, Singapore, 18–22 September 2022; pp. 474–484.
28. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; pp. 184–199.
29. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 105–114. [[CrossRef](#)]
30. Wang, X.; Yu, K.; Dong, C.; Change Loy, C. Recovering Realistic Texture in Image Super-Resolution by Deep Spatial Feature Transform. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 606–615. [[CrossRef](#)]

31. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Loy, C.C. ESRGAN: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision Workshops (ECCVW), Munich, Germany, 8–14 September 2018.
32. Ma, C.; Rao, Y.; Cheng, Y.; Chen, C.; Lu, J.; Zhou, J. Structure-Preserving Super Resolution With Gradient Guidance. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2020; pp. 7766–7775. [[CrossRef](#)]
33. Rakotonirina, N.C.; Rasoanaivo, A. ESRGAN+: Further Improving Enhanced Super-Resolution Generative Adversarial Network. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 3637–3641. [[CrossRef](#)]
34. Li, W.; Zhou, K.; Qi, L.; Lu, L.; Lu, J. Best-buddy gans for highly detailed image super-resolution. In Proceedings of the AAAI Conference on Artificial Intelligence, Online, 22 February–1 March 2022; Volume 36, pp. 1412–1420.
35. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. SwinIR: Image Restoration Using Swin Transformer. In Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME), Shenzhen, China, 5–9 July 2021; pp. 1833–1844. [[CrossRef](#)]
36. Lu, Z.; Li, J.; Liu, H.; Huang, C.; Zhang, L.; Zeng, T. Transformer for Single Image Super-Resolution. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 19–24 June 2022; pp. 456–465. [[CrossRef](#)]
37. Chen, X.; Wang, X.; Zhou, J.; Qiao, Y.; Dong, C. Activating More Pixels in Image Super-Resolution Transformer. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 18–22 June 2023; pp. 22367–22377.
38. Liang, D.; Cheng, J.; Ke, Z.; Ying, L. Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks. *IEEE Signal Process. Mag.* **2020**, *37*, 141–151. [[CrossRef](#)]
39. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends® Mach. Learn.* **2011**, *3*, 1–122.
40. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14), Cambridge, MA, USA, 8–13 December 2014; Volume 2, pp. 2672–2680.
41. Oktay, O.; Bai, W.; Lee, M.; Guerrero, R.; Kamnitsas, K.; Caballero, J.; de Marvaio, A.; Cook, S.; O'Regan, D.; Rueckert, D. Multi-input Cardiac Image Super-Resolution Using Convolutional Neural Networks. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016, Athens, Greece, 17–21 October 2016; pp. 246–254.
42. Shi, J.; Liu, Q.; Wang, C.; Zhang, Q.; Ying, S.; Xu, H. Super-resolution reconstruction of MR image with a novel residual learning network algorithm. *Phys. Med. Biol.* **2018**, *63*, 085011. [[CrossRef](#)]
43. Zhao, X.; Zhang, Y.; Zhang, T.; Zou, X. Channel Splitting Network for Single MR Image Super-Resolution. *IEEE Trans. Image Process.* **2019**, *28*, 5649–5662. [[CrossRef](#)] [[PubMed](#)]
44. Gu, Y.; Zeng, Z.; Chen, H.; Wei, J.; Zhang, Y.; Chen, B.; Li, Y.; Qin, Y.; Xie, Q.; Jiang, Z.; et al. MedSRGAN: Medical images super-resolution using generative adversarial networks. *Multimed. Tools Appl.* **2020**, *79*, 21815–21840. [[CrossRef](#)]
45. Parikh, N.; Boyd, S. Proximal algorithms. *Found. Trends® Optim.* **2014**, *1*, 127–239. [[CrossRef](#)]
46. Gilton, D.; Ongie, G.; Willett, R. Deep Equilibrium Architectures for Inverse Problems in Imaging. *IEEE Trans. Comput. Imaging* **2021**, *7*, 1123–1133. [[CrossRef](#)]
47. Fung, S.W.; Heaton, H.; Li, Q.; McKenzie, D.; Osher, S.; Yin, W. Jfb: Jacobian-free backpropagation for implicit networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Online, 22 February–1 March 2022; Volume 36, pp. 6648–6656.
48. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778. [[CrossRef](#)]
49. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269. [[CrossRef](#)]
50. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; pp. 694–711.
51. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
52. Zhang, T.; Pauly, J.M.; Vasanawala, S.S.; Lustig, M. Coil compression for accelerated imaging with Cartesian sampling. *Magn. Reson. Med.* **2013**, *69*, 571–582. [[CrossRef](#)]
53. Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
54. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 586–595. [[CrossRef](#)]
55. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.

56. Feng, C.M.; Yan, Y.; Fu, H.; Chen, L.; Xu, Y. Task Transformer Network for Joint MRI Reconstruction and Super-Resolution. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2021, Strasbourg, France, 27 September–1 October 2021; pp. 307–317.
57. Uecker, M.; Lai, P.; Murphy, M.J.; Virtue, P.; Elad, M.; Pauly, J.M.; Vasanawala, S.S.; Lustig, M. ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. *Magn. Reson. Med.* **2014**, *71*, 990–1001. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.