Abstract: Recent advances in deep learning have contributed greatly to the field of parallel MR imaging, where a reduced amount of k-space data are acquired to accelerate imaging time. In our previous work, we have proposed a deep learning method to reconstruct MR images directly from k-space data acquired with Cartesian trajectories. However, MRI utilizes various non-Cartesian trajectories, such as radial trajectories, with various numbers of multi-channel RF coils according to the purpose of an MRI scan. Thus, it is important for a reconstruction network to efficiently unfold aliasing artifacts due to undersampling and to combine multi-channel k-space data into single-channel data. In this work, a neural network named ‘ETER-net’ is utilized to reconstruct an MR image directly from k-space data acquired with Cartesian and non-Cartesian trajectories and multi-channel RF coils. In the proposed image reconstruction network, the domain transform network converts k-space data into a rough image, which is then refined in the following network to reconstruct a final image. We also analyze loss functions including adversarial and perceptual losses to improve the network performance. For experiments, we acquired k-space data at a 3T MRI scanner with Cartesian and radial trajectories to show the learning mechanism of the direct mapping relationship between the k-space and the corresponding image by the proposed network and to demonstrate the practical applications. According to our experiments, the proposed method showed satisfactory performance in reconstructing images from undersampled single- or multi-channel k-space data with reduced image artifacts. In conclusion, the proposed method is a deep-learning-based MR reconstruction network, which can be used as a unified solution for parallel MRI, where k-space data are acquired with various scanning trajectories.

Keywords: image reconstruction; RNN; end to end; MRI

1. Introduction

k-space is the sensor domain signal in magnetic resonance imaging (MRI) systems, and its trajectory is usually Cartesian. For certain purposes, such as to speed up the scanning time or to improve motion robustness, k-space data can be acquired with non-Cartesian trajectories [1], e.g., radial and spiral trajectories. In such cases, Radon transformation [2] or non-uniform fast Fourier transform (NUFFT) [3] can be considered for image reconstruction. Parallel imaging techniques have been used in combination with both Cartesian and non-Cartesian trajectories to reduce the data-acquisition time by undersampling. To reconstruct an image from undersampled k-space data in parallel imaging, many reconstruction algorithms have been developed. For example, sensitivity encoding (SENSE) [4] or generalized autocalibrating partially parallel acquisitions (GRAPPA) [5] can be adopted to reconstruct an image from regularly undersampled k-space data. For irregularly undersampled k-space data, compressed sensing [6] can be used.

As deep learning techniques are widely adopted in various imaging processing fields [7–10], a number of deep learning techniques have been introduced in the medical imaging field. Generally, deep learning techniques are adopted to reconstruct images...
for conventional medical imaging systems, including MRI, computed tomography (CT), single-photon emission computed tomography (SPECT) and positron emission tomography (PET), and to improve efficiency in diagnostic and research processes [11–15]. For MR image reconstruction, deep learning approaches can be categorized into three types: estimation of full k-space data from undersampled k-space data, transformation of a corrupted image to an artifact-free image, and reconstruction of an image from k-space input data. For the estimation of un-acquired data from undersampled k-space data, a convolutional neural network (CNN) has been developed [16,17]. Then, to reconstruct an image from the estimated full k-space data, a conventional transformation, such as inverse Fourier transform (IFT), can be used. To transform a corrupted input image into an artifact-free output image, multi-layer perceptrons (MLP) or CNN-based networks have been proposed [18–21]. In these cases, an initial aliased input image is obtained by IFT of the undersampled k-space data, and the proposed network is used to generate an anti-aliased image. To further improve the quality of an anti-aliased output image, data consistency, adversarial loss, and perceptual loss terms can be properly added, where the original k-space data are utilized as known prior information.

Domain transform can be performed by deep learning approaches, where an output image can be directly reconstructed from input k-space data [22,23]. As an analytic solution to domain transformation already exists, deep learning-based image reconstruction methods that transform k-space data into an image are less frequently reported than image to image or k-space to k-space reconstruction methods. The k-space to image domain transform in MRI requires a full receptive field, because a single point of the k-space data is defined by integration of the entire spatial information [24]. For the deep learning-based direct mapping method, automated transform by manifold approximation (AUTOMAP) was proposed as an end-to-end image reconstruction architecture with multiple fully connected layers [22]. In AUTOMAP, a deep learning network was used to reconstruct an image from the measured k-space data to correct unknown artifacts and errors. However, fully connected layers of AUTOMAP require a large memory size, which restricts image resolutions by current hardware performances.

In our previous work, we have proposed a deep learning method to reconstruct MR images directly from the k-space data acquired while solving the issues related to memory size and hardware performance. In this work, we take one step further and demonstrate that the proposed method can be used to reconstruct images acquired with various non-Cartesian trajectories, such as radial trajectories, with various numbers of multi-channel RF coils according to the purpose of the MRI scan. The novel contribution of this work is that the proposed method can be utilized as a unified solution for image reconstruction because it can be used regardless of the trajectories while overcoming the hardware limitations for computation and thereby enabling the reconstruction of higher-resolution images. The proposed network achieved this goal by replacing fully connected network with a bi-directional recurrent neural network (RNN) for end-to-end image reconstruction [25]. The core layers of the proposed network perform domain transformation through RNN-based networks, and the additional CNN-based network performs refinement for the final image. In the following sections, the proposed network will be explained and analyzed in detail.

2. Methods

A schematic illustration of the proposed end-to-end MR image reconstruction network (ETER-net) is presented in Figure 1. The ETER-net is an RNN-based image reconstruction technique for undersampled k-space data acquired with either Cartesian or non-Cartesian trajectories. The deep learning architecture in the proposed method comprises two parts: a domain transform network and a refinement network. In the domain transform network, the input k-space data are transformed into multiple latent features in the image domain. Then, the refinement network generates a final image from the multiple spatial features. A detailed description of the domain transformation and the refinement networks is provided in the following sections.
2.1. Domain Transform Network

The domain transform network should transform the \( k \)-space input data into an image. In a global transform, the entire \( k \)-space information should be reflected for the generation of a single pixel in an image. Even if some fully connected layers can learn global transformation, fully connected layers require huge parameters, which increases the amount of computation and GPU memory size. The proposed method utilizes four recurrent neural networks that sweep both horizontally and vertically in forward and reverse directions across the input \( k \)-space data, which reduces the demand on the number of parameters while utilizing the entire data. The domain transform network is designed to produce the output with the same width and height as the final image. Because the proposed network has different input and output sizes, it can be used to reconstruct images from \( k \)-space data acquired with non-Cartesian, undersampled, and arbitrary trajectories. The recurrent layers in the proposed method can be expressed as:

\[
\begin{align*}
    v_i^F &= f_{FWD}(z_{i-1}^F, x_i), \quad i = 1, \ldots, I, \\
    v_i^R &= f_{REV}(z_{i+1}^R, x_i), \quad i = I, \ldots, 1.
\end{align*}
\]

where \( f_{FWD} \) is the forward sweeping recurrent layer and \( f_{REV} \) is the reverse sweeping recurrent layer. The forward sweeping is performed from left to right in the horizontal recurrent layer and from top to bottom in the vertical recurrent layer. Similarly, the reverse sweeping is performed from right to left in the horizontal recurrent layer and from bottom to top in the vertical recurrent layer. The input vector is indicated as \( x_i \). The output feature vectors and the hidden state vectors generated by the forward and reverse recurrent layers are represented as \( v_i^F, v_i^R, z_i^F \) and \( z_i^R \), respectively.
In this work, the input \( k\)-space data are assumed as complex-valued data with a dimension of \( K_x \times K_y \times N_c \), where \( K_x \) is the number of readout samples, \( K_y \) is the number of phase encoding in the Cartesian coordinates (or the number of radial views in the radial coordinates), and \( N_c \) is the number of Rx coils. By modifying the complex-valued data into two sets of real-valued data, input \( X \in \mathbb{R}^{K_x \times K_y \times 2N_c} \) has 3D data of \( K_x \times K_y \times 2N_c \).

Each 2D matrix of the input 3D matrix \( X \) along the \( k_x \) direction is inserted (and flattened) into a horizontal recurrent layer, as illustrated in Figure 1. The horizontal recurrent layer sweeps the horizontal (\( k_x \)) axis in the forward and reverse bi-directions, and the length of generated outputs is \( K_x \), which is equivalent to the input sequence length. For the horizontal recurrent layer, the number of neurons in the horizontal recurrent layer is \( N_y \times N_f \), where \( N_f \) is a hyper-parameter defined by user. Then, the input and output of the bi-directional horizontal recurrent layers can be expressed as:

\[
\begin{align*}
v^{F,\text{hor}}_i &= f_{\text{FWD}}(z^{F,\text{hor}}_{i-1}, x_i^{\text{col}}), \quad i = 1, \ldots, K_x, \\
v^{R,\text{hor}}_i &= f_{\text{REV}}(z^{R,\text{hor}}_{i+1}, x_i^{\text{col}}), \quad i = K_x, \ldots, 1.
\end{align*}
\]

where \( i \) is the index in the horizontal direction and \( x_i^{\text{col}} \) is the flattened form of the \( i \)th 2D matrix of the input 3D matrix \( X \) along the \( k_x \) direction. For the horizontal forward recurrent layer, \( z^{F,\text{hor}}_i \) is the \( i \)th hidden state and \( v^{F,\text{hor}}_i \) is the \( i \)th output vector. For the horizontal reverse recurrent layer, \( z^{R,\text{hor}}_i \) is the \( i \)th hidden state and \( v^{R,\text{hor}}_i \) is the \( i \)th output. The entire output of the horizontal recurrent layer can be represented as:

\[
v^{\text{hor}} = \left\{ v^{F,\text{hor}}_1, \ldots, v^{F,\text{hor}}_{K_x}, v^{R,\text{hor}}_1, \ldots, v^{R,\text{hor}}_{K_x} \right\}.
\]

To transfer the output of the horizontal recurrent layer to the following vertical recurrent layer, \( V^{\text{hor}} \) is reshaped as \( V' \) (\( \in \mathbb{R}^{K_x \times N_y \times N_f} \)). By vertically sweeping \( V' \) in the forward and reverse directions, the vertical recurrent layer generates an output with a sequence length \( N_y \) and the number of neurons \( N_x \times N_f \). The relationship between the input and output of the bi-directional vertical recurrent layers can be written as:

\[
\begin{align*}
v^{F,\text{ver}}_j &= f_{\text{FWD}}(z^{F,\text{ver}}_{j-1}, x_j^{\text{row}}), \quad j = 1, \ldots, N_y, \\
v^{R,\text{ver}}_j &= f_{\text{REV}}(z^{R,\text{ver}}_{j+1}, x_j^{\text{row}}), \quad j = N_y, \ldots, 1.
\end{align*}
\]

where \( j \) is the index in the vertical direction and \( x_j^{\text{row}} \) is the flattened form of the \( j \)th 2D matrix of the input 3D matrix \( V' \) along the \( n_y \) direction. For the vertical forward recurrent layer, \( z^{F,\text{ver}}_j \) is the \( j \)th hidden state and \( v^{F,\text{ver}}_j \) is the \( j \)th output. For the vertical reverse recurrent layer, \( z^{R,\text{ver}}_j \) is the \( j \)th hidden state and \( v^{R,\text{ver}}_j \) is the \( j \)th output. Then, the entire output of the vertical RNN (\( V^{\text{ver}} \)) is reshaped as \( N_x \times N_y \times N_f \) to be used as an input for the following refinement network.

To summarize, the domain transform network generates output latent features (\( \in \mathbb{R}^{N_x \times N_y \times N_f} \)) from the input \( k\)-space data (\( \in \mathbb{R}^{K_x \times K_y \times 2N_c} \)).

### 2.2. Refinement Network

The latent features generated as an output of the domain transform network can be regarded as a roughly reconstructed image. To merge and refine the latent features into a final image, the refinement network should perform a simple image-domain to image-domain task. Thus, various types of neural network architectures can be used to generate a single channel magnitude image from multiple channel spatial features, where the widths and heights of the spatial features and the final image are equivalent. We
evaluated CNN-based networks for the refinement network because CNN can provide the best performance in terms of image quality. To find an optimal network for the ETER-net, we compared the single convolutional layer, the multiple convolutional layers inspired by winner-take-all autoencoder (WTA) [26], and a dual frame U-net (DFU) [27], which is one of the most commonly used architectures.

2.3. Loss Function

The most frequently used loss function, particularly for regression tasks, is the Euclidean distance between the given value of the label and the output of the network. However, using only the Euclidean distance as a loss function tends to blur the final image [28–31]. In order to overcome this problem, additional losses have been examined, including the adversarial loss, which uses a discriminator network to classify real and generated data [32], and the perceptual loss, which utilizes the distance of feature levels between the real and generated data examined by a pre-trained network [33]. To optimize the performance of the proposed reconstruction network, we utilize Euclidean distance, the adversarial loss based on the patch generative adversarial network (GAN) discriminator [29], and perceptual loss based on the pretrained "VGG" network [34]. The total loss function can be written as:

\[ L_{\text{total}} = \lambda L_1 L_1 + \lambda_{\text{adv}} L_{\text{adv}} + \lambda_{\text{VGG}} L_{\text{VGG}}. \]  

(8)

3. Experiments

3.1. Data Preparation

As explained, the proposed network is applicable for the reconstruction of images from k-space data acquired with Cartesian or non-Cartesian trajectories. To demonstrate image reconstruction for various k-space trajectories, both Cartesian and non-Cartesian k-space data were acquired using FSE at a 3T MRI (Siemens Verio) with the following parameters: For Cartesian k-space data, FOV = 220 × 220 mm², matrix size = 216 (FE) × 216 (PE), slice thickness = 5 mm, TR/TE = 5000/90 ms, and echo train length = 18. For radial k-space data, FOV = 256 × 256 mm², matrix size = 256 (frequency encoding) × 400 (view angles), slice thickness = 5 mm, TR/TE = 5000/52 ms, and echo train length = 8.

Then, for MRI in Cartesian coordinates, the label images were prepared by Fourier transform of the full k-space data of each receive coil channel, taking the root mean square of the multi-channel images. For radial MRI, images from each receive coil channel were reconstructed by the filtered back projection, and the root mean square of the multi-channel images was calculated to generate the label images. For stabilization of the learning process, the k-space values and the intensity values of the label images were normalized to [−1, 1] and [0, 1], respectively.

From the acquired k-space data, accelerated parallel acquisition was simulated by retrospective sub-sampling of k-space data along the phase encoding direction for Cartesian k-space data or along the view angles for radial k-space data with various reduction factors. For Cartesian MRI, both regular and irregular undersampling patterns were used, as illustrated in Figure 2a. The number of channels of the receive coil was eight. A total of 1520 images from 14 subjects were used for training, and 64 images from four subjects were used for testing. For radial MRI, 2538 images from 26 subjects were used for training and 48 images from four subjects were used for testing. The number of channels of the receive coil for radial MRI was four. To verify the performance with respect to various acceleration factors, we examined the proposed method with reduction factors R of 2, 4, 6 and 8, respectively.
Accelerated MRI usually utilizes multi-channel receive coils to incorporate the local sensitivity of each channel of each receive coil for image reconstruction. To verify the reconstruction performance of the proposed method in case of accelerated imaging with a single-channel receive coil, we attempted to reconstruct images from a single-channel k-space data. For this experiment, the label image was reconstructed from fully acquired single-channel k-space data, and the input k-space data of the deep network were undersampled with a reduction factor of four.

3.2. Experimental Settings

From the acquired k-space datasets, images were reconstructed by the proposed ETER-net, two image-domain to image-domain reconstruction methods (dual frame U-net [27] and DAGAN [19]), and a k-domain to k-domain reconstruction method (k-space deep learning [16]). In the domain transformation network of the ETER-net, the user-defined parameter $N_f$ was 32, which was the maximum possible value, and it was limited by the memory size of the GPU we used. If $N_f$ increases as the model size becomes larger, the loss values of the network converge at lower values. As illustrated in Figure 3, we find that the proposed model with a small $N_f$ is not enough for end-to-end image reconstruction. In this work, we set the $N_f$ as the maximum possible value to maximize the size of the proposed model.

For the refinement network, experiments were performed with the following settings. In case of the single convolutional layer, it was designed to merge its input features with $N_f = 32$ channels into a single channel output. For the multiple convolutional layers (WTA), which were also used in AUTOMAP [22,26], the output channels of convolutional layers were set as 256 and the final deconvolutional layer was designed to merge the features of 256 channels into a single channel output. While the vanilla U-net only utilized a skip connection, the dual frame U-net (DFU) utilized an additional by-pass connection for residual learning [27] in combination with the skip connection, resulting in improved learning performance.
Various loss functions were analyzed, where the adversarial loss and the perceptual loss were added to the ETER-net, while the pixel-wise loss function with a weighting factor ($\lambda$) of 1 was utilized as a baseline. For the perceptual loss, the feature-level difference was used as a loss based on the pre-trained VGG network. The loss functions utilizing either adversarial loss, perceptual loss, or both were compared to evaluate the effectiveness of the loss functions. Other parameters are described as follows. For Cartesian MRI, $\lambda$ for adversarial loss was also set as 1 when only the adversarial loss was additionally used in combination with the baseline loss function. If only the perceptual loss was added to the baseline loss function, $\lambda = 50$ was used for perceptual loss. When both adversarial loss and perceptual loss were additionally used, $\lambda = 0.02$ and $\lambda = 5$ were used for adversarial loss and perceptual loss, respectively. Then, the total number of parameters of the ETER-net utilizing the DFU for the refinement network is 312,367,393. For radial MRI, $\lambda = 0.001$ was used for adversarial loss where the adversarial loss was additionally used in combination with the baseline loss function. When only the perceptual loss was added to the baseline loss function, $\lambda = 0.01$ was used for perceptual loss. When both adversarial loss and perceptual loss were additionally used, $\lambda = 0.001$ and $\lambda = 1$ were used for adversarial loss and perceptual loss, respectively. Then, the total number of parameters of the ETER-net was 453,460,769. The training was commonly performed as in Adam [35] with $lr = 0.0001$.

The proposed architecture was implemented on Pytorch [36] and computed by an Intel(R) Core (TM) i5-8400 2.80GHz CPU and NVIDIA TITAN XP GPU. We used the gated recurrent unit (GRU) [37] for recurrent units to reduce the number of parameters. There was no significant difference in the performance compared with the long short-term memory (LSTM) [38].

3.3. Quantitative Evaluation

The normalized mean square error (nMSE) and structural similarity index (SSIM) [39] are deployed for the quantitative evaluation. They are defined as follows:
\[ nMSE = \frac{\sum\sum \left[ x_{ij}^{GT} - x_{ij}^{PRED} \right]^2}{\sum\sum \left[ x_{ij}^{GT} \right]^2}, \]  

\[ SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)}. \]  

where \( x_{ij}^{GT} \) is the ground truth image, \( x_{ij}^{PRED} \) is the predicted image, \( \mu_x \) and \( \mu_y \) are the means of \( x \) and \( y \), respectively, \( \sigma_x^2 \) is the variances of \( x \), \( \sigma_y^2 \) is the variances of \( y \), and \( \sigma_{xy} \) is the covariance of \( x \) and \( y \). To prevent divergence by dividing by zero, small constants are included such as \( c_1 \) and \( c_2 \).

4. Results

4.1. Variants of the Refinement Network

Figure 4 presents the reconstructed images from three different refinement networks: the single convolution layer, multiple convolutional layers, and dual frame U-net. Figure 5 presents the quantitative analysis of the three refinement networks. As shown in Figure 4, the images refined by the single convolution layer contain a substantial amount of structural artifacts. In addition, the image details are also deteriorated. In contrast, the images refined by the WTA and DFU have considerably improved image quality, and the image details are clear. However, the images refined by the DFU have better image details. The quantitative measures also demonstrate that the DFU can be efficiently used for the refinement of the latent features in the proposed method. The single convolutional layer shows considerably higher \( nMSE \) and lower \( SSIM \) values than the other two methods, and the DFU shows improved performance in terms of \( nMSE \) and \( SSIM \).

![Figure 4. Comparison of three different refinement networks (Section 4.1). Two example slices of experiment results are shown. (a): Ground truth magnitude images (label images). (b–d): Reconstructed images using (b) single convolutional layer, (c) WTA, and (d) DFU.](image-url)
4.2. Acceleration with a Single-Channel RF Coil

While most parallel MRI methods utilize information from multi-channel RF coils to reconstruct images from undersampled k-space data, deep learning-based reconstruction approaches can generate unaliased images from an undersampled dataset without using the information of the multi-channel receive coil [40]. To demonstrate the performance of the proposed ETER-net, we performed image reconstruction from the undersampled k-space data obtained with a single receive channel. To compare the results with a commonly used deep learning reconstruction method, we also simulated the image-domain to image-domain reconstruction using the dual frame U-net. In this experiment, the label images were reconstructed from fully sampled k-space data of a single receive channel. Figure 6 shows three different slices reconstructed from a single channel k-space data, which were undersampled with a reduction factor of four. All images show non-uniform intensities because the k-space data are acquired from a single receive channel, which has non-uniform sensitivity. The images reconstructed by the DFU (column (b)) and the ETER-net (column (d)) are very similar to the label images in column (a). However, when the images are brightened to emphasize artifacts (columns (c) and (e)), it is clear that the proposed ETER-net generates images with less artifacts than the DFU. In particular, the images reconstructed by the DFU have more blurred structural edges and apparent aliasing artifacts, while the ETER-net generates alias-free images with improved image details.

4.3. Loss Functions

To further optimize the proposed network for image reconstruction, two different losses were additionally calculated along with the baseline loss function of the Euclidean distance: adversarial loss and perceptual loss. The effects of the loss functions were analyzed for the undersampled k-space data with R = 4. The ETER-net with different loss functions was also compared with two different image-domain to image-domain reconstruction methods, DFU [27] and DAGAN [19], and a k-domain to k-domain reconstruction method, k-space deep learning for accelerated MRI [16]. Figure 7 shows the result images and errors, and Figure 8 shows the quantitative analysis of the reconstructed images. Figure 7b,c show a substantial amount of image artifacts and aliasing patterns, owing to the regular subsampling of the input k-space data. In contrast, the proposed methods show quality results, and the aliasing artifacts are eliminated in these images (columns (e–h)). By comparing the images reconstructed with various loss functions, it can be concluded that the loss functions, including the Euclidean distance, adversarial loss, and perceptual loss, provided the best performance in terms of image quality.
Figure 6. Images reconstructed from a single channel k-space data, which are undersampled with a reduction factor of four: (a): label images, (b,c): images reconstructed by the DFU and the corresponding brightened images, (d,e): images reconstructed by the ETER-net and the corresponding brightened images. Brightening was performed by taking the root of each value in the image.
Figure 7. Comparison of previous methods with the proposed method. **first and third row**: Magnitude images. **second and fourth row**: Corresponding error maps in the 1/10 range. (a): Ground truth magnitude images (label images). The images are reconstructed by (b): DFU, (c): DAGAN, (d): k-space deep learning, (e): baseline loss, (f): baseline loss + perceptual loss, (g): baseline loss + adversarial loss, and (h): baseline loss + adversarial loss + perceptual loss.

4.4. **Radial Sampling Data**

The proposed method can be applied to the data acquired with a non-Cartesian trajectory because it is a unified solution for image reconstruction. As an example of non-Cartesian trajectories, we examined the proposed method for k-space data obtained with a radial trajectory. For the reconstruction of images from the radially under-sampled k-space data with a reduction factor of four, we used the ETER-net with the DFU as the refinement network. The effectiveness of various loss functions was also analyzed for the reconstruction of radially sampled k-space data by comparing the baseline loss function of the Euclidean distance with the additional adversarial and perceptual losses.
Figure 8. Quantitative analysis of DFU, DAGAN, and the proposed methods with/without additional losses. nMSE and SSIM were measured from 64 slices of four subjects. U-net, DAGAN, kspDL: comparison methods, ETER: with baseline loss function of pixel-wise Euclidean distance, GAN: with additional adversarial loss function, VGG: with additional perceptual loss via pre-trained VGG network, GAN + VGG: with additional adversarial and perceptual loss functions.

The reconstructed images from radially sampled k-space data are shown in Figure 9. When the filtered back projection was used to reconstruct the images from the under-sampled radial k-space data, the streaking artifacts could be observed in the images, as shown in Figure 9 (rows 2). In contrast, the ETER-net provided improved image quality (rows 3 and 4). Moreover, the comparison of the images in rows 3 and 4 demonstrates that the loss function including Euclidean distance, adversarial loss, and perceptual loss improved the image quality. The improvement owing to the additional loss functions can also be validated by the quantitative measures in Table 1. The average nMSE is 3.51% for the filtered back projection, 1.98% for the proposed method with L1 loss only, and 1.49% for the proposed method with additional losses. In terms of SSIM, the average score of the filtered back projection is 0.786, the average score of the proposed method with L1 loss only is 0.922, and the average score of the proposed method with additional losses is 0.938. The proposed method with additional losses shows the least error and highest SSIM.

Table 1. Statistics for the proposed and the inverse Radon transform from 100 view angles of radial k-space data. Test set is 48 images from four subjects.

| Case                        | nMSE (%) | SSIM     |
|-----------------------------|----------|----------|
| Radon                       | 3.51 ± 0.63 | 0.786 ± 0.019 |
| Baseline                    | 1.98 ± 0.40 | 0.922 ± 0.013 |
| Baseline + Adversarial loss | 1.49 ± 0.45 | 0.938 ± 0.015 |
| Baseline + Perceptual loss  |          |          |
5. Discussion

In this study, a neural network architecture that reconstructs an MR image from \textit{k-space} data was proposed by utilizing the RNNs. The proposed method can be used to reconstruct an image from \textit{k-space} data acquired with various \textit{k-space} trajectories, including Cartesian, non-Cartesian, fully sampled and under-sampled datasets, because the proposed network is a unified solution for image reconstruction. Moreover, it is applicable both for \textit{k-space} data acquired with single or multi-channel RF coils.

Likewise, AUTOMAP can learn a direct sensor-to-image mapping. However, AUTOMAP has a major drawback, i.e., the limitation of the size of the reconstructed image, owing to the hardware performance. Because of the fully connected layers, AUTOMAP requires the number of parameters as \(O(n^2)\) to produce an \(n \times n\) image. However, the
The proposed method requires considerably fewer parameters by utilizing recurrent neurons, thus enabling the reconstruction of high-resolution images. To produce an \( n \times n \) image, the proposed method requires the number of parameters as \( O(n) \). Thus, the main advantage of the proposed method in comparison with the AUTOMAP is a considerable reduction in the number of parameters for the network.

The proposed method is a deep learning architecture that generates a final image from the input \( k\)-space data; this implies that it requires a domain transform to the image domain. We assumed that bi-directional recurrent layers can be used to perform the domain transform. To verify this, we examined the output features of the recurrent layers. The domain transform network of the proposed method successfully conducts domain transformation from sensor-to-image.

For refinement network, experiments were performed by applying three different CNNs. The results showed that there was a significant difference in the performance depending on which network was used for the refinement network. The performance was better when the DFU was applied, which is an advanced CNN-based network, compared with when the single convolutional layer was used. A comparable performance was shown when multiple convolutional layers were used.

The proposed method was applied to the radial trajectory as an example of non-Cartesian trajectory. For the experiments on radial trajectory, the label was defined as the reconstructed image from the Radon transform using 400 views. As demonstrated by the experiments, the view and readout axes in the radial trajectory correspond to the phase encoding and frequency encoding axes in the Cartesian trajectories, respectively, and the image reconstruction using the proposed ETER-net was successfully performed.

In general, the proposed method also provided blurry results owing to the regression problem in which the Euclidean distance was used as a loss function. To address this, learning experiments were conducted by including adversarial loss and perceptual loss. The adversarial loss encouraged the network to favor realistic solutions by trying to fool the discriminator network. Furthermore, perceptual loss allowed the network to reconstruct fine details and edges. Through these experiments, blurry results were obtained again when only the Euclidean distance was used as the loss function, and it was verified that this could be overcome by using additional losses.

6. Conclusions

In this work, we have proposed an efficient alternative for a deep learning-based direct mapping method that reconstructs images from undersampled \( k\)-space data. While previous direct mapping methods could be selectively used in cases of relatively low-resolution images due to limitations, we have demonstrated that the proposed method can be used for higher-resolution images by utilizing the inductive bias that comes from the structural characteristics of bi-RNN, where shared weights are used for sequentially inserted inputs, and information from the past can be integrated through the hidden state. The proposed method can be used as a unified solution applicable to various scanning trajectories, such as radial, spiral, and Cartesian trajectories, and it can be applied to single and multi-channel data. To verify the proposed method, experiments were conducted with an actual Cartesian and radial dataset, and both quantitative and qualitative results showed that the proposed method is effective for MR image reconstruction.

Author Contributions: Conceptualization, C.O.; methodology, C.O. and Y.H.; software, C.O.; validation, J.-Y.C. and Y.H.; data curation, C.O.; writing—original draft preparation, C.O.; writing—review and editing, Y.H.; visualization, C.O.; supervision, J.-Y.C. and Y.H.; funding acquisition, J.-Y.C. All authors have read and agreed to the published version of the manuscript.
Funding: This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2022R1A2C2010363), and partly supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR14C0002).

Institutional Review Board Statement: All in vivo experiments were approved by the KAIST Institutional Review Board (IRB), IRB Number: KH2012-005.

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

Data Availability Statement: Not applicable.

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

References
1. Dikaios, N.; Protonotarios, N.E.; Fokas, A.S.; Kastis, G.A. Quantification of T1, T2 relaxation times from Magnetic Resonance Fingerprinting radially undersampled data using analytical transformations. Magn. Reson. Imaging 2021, 80, 81–89. [CrossRef] [PubMed]

2. Deans, S.R. The Radon Transform and Some of Its Applications; Courier Corporation: Chelmsford, MA, USA, 2007.

3. Fessler, J.A. On NUFFT-based gridding for non-Cartesian MRI. J. Magn. Reson. 2007, 188, 191–195. [CrossRef]

4. Pruessmann, K.P.; Weiger, M.; Scheidegger, M.B.; Boesiger, P. SENSE: Sensitivity encoding for fast MRI. Magn. Reson. Med. 1999, 42, 952–962. [CrossRef]

5. 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]

6. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med. 2007, 58, 1182–1195. [CrossRef] [PubMed]

7. Wu, D.; Wu, C. Research on the Time-Dependent Split Delivery Green Vehicle Routing Problem for Fresh Agricultural Products with Multiple Time Windows. Agriculture 2022, 12, 793. [CrossRef]

8. Chen, H.; Miao, F.; Chen, Y.; Xiong, Y.; Chen, T. A hyperspectral image classification method using multifeature vectors and optimized KELM. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2021, 14, 2781–2795. [CrossRef]

9. Zhao, H.; Liu, J.; Chen, H.; Chen, J.; Li, Y.; Xu, J.; Deng, W. Intelligent diagnosis using continuous wavelet transform and gauss convolutional deep belief network. IEEE Sens. J. 2022, 22. [CrossRef]

10. Deng, W.; Ni, H.; Liu, Y.; Chen, H.; Zhao, H. An adaptive differential evolution algorithm based on belief space and generalized opposition-based learning for resource allocation. Appl. Soft Comput. 2022, 127, 109419. [CrossRef]

11. Kawauchi, K.; Furuya, S.; Hirata, K.; Katoh, C.; Manabe, O.; Kobayashi, K.; Watanabe, S.; Shiiga, T. A convolutional neural network-based system to classify patients using FDG PET/CT examinations. BMC Cancer 2020, 20, 227. [CrossRef]

12. Arabi, H.; AkhavanAllaf, A.; Sanaat, A.; Shiga, T. The promise of artificial intelligence and deep learning in PET and SPECT imaging. Phys. Medica 2021, 83, 122–137. [CrossRef] [PubMed]

13. Kumar, A.; Fulham, M.; Feng, D.; Kim, J. Co-learning feature fusion maps from PET-CT images of lung cancer. IEEE Trans. Med. Imaging 2019, 39, 204–217. [CrossRef] [PubMed]

14. Protonotarios, N.E.; Katsamenis, I.; Sykiotis, S.; Dikaio, N.; Kastis, G.A.; Chatzioannou, S.N.; Metaxas, M.; Doulamis, N.; Doulamis, A. A few-shot U-Net deep learning model for lung cancer lesion segmentation via PET/CT imaging. Biomed. Phys. Eng. Express 2022, 8, 025019. [CrossRef]

15. Varoquaux, G.; Cheplygina, V. Machine learning for medical imaging: Methodological failures and recommendations for the future. NPJ Digit. Med. 2022, 5, 48. [CrossRef] [PubMed]

16. Han, Y.; Ye, J.C. k-Space Deep Learning for Accelerated MRI. arXiv 2018, arXiv:1805.03779.

17. Eo, J.; Jun, Y.; Kim, T.; Jang, J.; Lee, H.J.; Hwang, D. KIKI-net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magn. Reson. Med. 2018, 5, 2188–2201. [CrossRef]

18. Kwon, K.; Kim, D.; Park, H. A parallel MR imaging method using multilayer perceptron. Med. Phys. 2017, 44, 6209–6224. [CrossRef]

19. 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]

20. Hammernik, K.; Klatzer, T.; Kobler, E.; Recht, M.P.; Sodickson, D.K.; Pock, T.; Knoll, F. Learning a variational network for reconstruction of accelerated MRI data. Magn. Reson. Med. 2018, 79, 3055–3071. [CrossRef]

21. Schlemper, J.; Caballero, J.; Hajnal, J.V.; Price, A.N.; Rueckert, D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging 2018, 37, 491–503. [CrossRef]

22. Zhu, B.; Liu, J.Z.; Catley, S.F.; Rosen, B.R.; Rosen, M.S. Image reconstruction by domain-transform manifold learning. Nature 2018, 555, 487. [CrossRef] [PubMed]
23. Oh, C.; Kim, D.; Chung, J.Y.; Han, Y.; Park, H. A k-space-to-image reconstruction network for MRI using recurrent neural network. *Med. Phys.* 2021, 48, 193–203. [CrossRef] [PubMed]

24. Wang, G.; Ye, J.C.; Mueller, K.; Fessler, J.A. Image reconstruction is a new frontier of machine learning. *IEEE Trans. Med. Imaging* 2018, 37, 1289–1296. [CrossRef] [PubMed]

25. Visin, F.; Kastner, K.; Cho, K.; Matteucci, M.; Courville, A.; Bengio, Y. Renet: A recurrent neural network based alternative to convolutional networks. *arXiv* 2015, arXiv:1505.00393.

26. Makhzani, A.; Frey, B.J. Winner-take-all autoencoders. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 2791–2799.

27. Han, Y.; Ye, J.C. Framing U-Net via deep convolutional framelets: Application to sparse-view CT. *IEEE Trans. Med. Imaging* 2018, 37, 1418–1429. [CrossRef]

28. 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 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.

29. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.

30. Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context encoders: Feature learning by inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2536–2544.

31. Zhang, R.; Isola, P.; Efros, A.A. Colorful image colorization. In *Proceedings of the European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 649–666.

32. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.

33. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In *Proceedings of the European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 694–711.

34. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* 2014, arXiv:1409.1556.

35. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* 2014, arXiv:1412.6980.

36. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in PyTorch. *NIPS-W 2017*. Available online: https://openreview.net/forum?id=BJJsrmICZ (accessed on 22 September 2022).

37. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* 2014, arXiv:1406.1078.

38. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* 1997, 9, 1735–1780. [CrossRef]

39. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 2004, 13, 600–612. [CrossRef]

40. Muckley, M.J.; Riemenschneider, B.; Radmanesh, A.; Kim, S.; Jeong, G.; Ko, J.; Jun, Y.; Shin, H.; Hwang, D.; Mostapha, M.; et al. Results of the 2020 fastmri challenge for machine learning mr image reconstruction. *IEEE Trans. Med. Imaging* 2021, 40, 2306–2317. [CrossRef] [PubMed]