Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(9): 15982–15998.
DOI: 10.3934/mbe.2023712
Received: 15 June 2023
Revised: 25 July 2023
Accepted: 27 July 2023
Published: 04 August 2023

Research article

Time-Resolved 3D cardiopulmonary MRI reconstruction using spatial transformer network

Qing Zou[1,2,3,*], Zachary Miller[4], Sanja Dzelebdzic[1], Maher Abadeer[1], Kevin M. Johnson[5,6] and Tarique Hussain[1,2,3]

[1] Division of Pediatric Cardiology, Department of Pediatrics, The University of Texas Southwestern Medical Center, Dallas, TX, USA
[2] Department of Radiology, The University of Texas Southwestern Medical Center, Dallas, TX, USA
[3] Advanced Imaging Research Center, The University of Texas Southwestern Medical Center, Dallas, TX, USA
[4] Department of Biomedical Engineering, University of Wisconsin, Madison, WI, USA
[5] Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
[6] Department of Radiology, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA

* **Correspondence:** Email: Qing.Zou@UTSouthwestern.edu; Tel: +12146455024; Fax: +12144562154.

Abstract: The accurate visualization and assessment of the complex cardiac and pulmonary structures in 3D is critical for the diagnosis and treatment of cardiovascular and respiratory disorders. Conventional 3D cardiac magnetic resonance imaging (MRI) techniques suffer from long acquisition times, motion artifacts, and limited spatiotemporal resolution. This study proposes a novel time-resolved 3D cardiopulmonary MRI reconstruction method based on spatial transformer networks (STNs) to reconstruct the 3D cardiopulmonary MRI acquired using 3D center-out radial ultra-short echo time (UTE) sequences. The proposed reconstruction method employed an STN-based deep learning framework, which used a combination of data-processing, grid generator, and sampler. The reconstructed 3D images were compared against the start-of-the-art time-resolved reconstruction method. The results showed that the proposed time-resolved 3D cardiopulmonary MRI reconstruction using STNs offers a robust and efficient approach to obtain high-quality images. This method effectively overcomes the limitations of conventional 3D cardiac MRI techniques and has the potential to improve the diagnosis and treatment planning of cardiopulmonary disorders.

Keywords: cardiopulmonary MRI; spatial transformer network; 3D UTE sequence

1. Introduction

Cardiovascular and respiratory diseases are leading causes of morbidity and mortality worldwide [1–3]. Accurate and timely diagnosis is crucial for the effective management and treatment of these disorders, necessitating the development of advanced diagnostic tools. MRI has emerged as a powerful non-invasive tool for the assessment of cardiac and pulmonary structures [4–6], allowing for a comprehensive evaluation of anatomy [7], function [8], and perfusion [9] in three dimensions (3D). However, the acquisition of high-quality 3D images remains challenging due to the complex and dynamic nature of the cardiovascular system, motion artifacts from respiratory and cardiac cycles, and the inherently long acquisition times associated with MRI.

Conventional 3D cardiac MRI techniques, such as 3D balanced steady-state free precession (bSSFP) [10] and 3D fast gradient echo (FGRE) [11], are widely used in clinical practice. These techniques offer improved signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) compared to their 2D counterparts [12,13]. However, they suffer from several limitations that may compromise their diagnostic accuracy. First, they are sensitive to motion artifacts due to the long acquisition times, which can lead to image degradation and impede the detection of subtle anatomical details. Second, these techniques often exhibit limited spatiotemporal resolution, which may hinder the precise assessment of cardiac and pulmonary function. Lastly, the breath-hold requirements and scan duration of conventional 3D cardiac MRI techniques can be challenging for patients, particularly those with compromised respiratory or cardiac function. Consequently, there is an unmet need for novel 3D cardiopulmonary MRI acquisition methods and reconstruction methods that overcomes these limitations and provides improved image quality and diagnostic accuracy.

From the data acquisition perspective, non-Cartesian (such as radial) ultrashort echo time (UTE) sequences have recently emerged as a tool for imaging the cardiopulmonary system. 3D radial acquisition has emerged as a powerful technique for cardiopulmonary MRI [14–16], offering numerous advantages in terms of image quality, motion robustness and scan efficiency. 3D radial acquisition has emerged as a powerful technique for cardiopulmonary MRI, offering numerous advantages in terms of image quality, motion robustness and scan efficiency. 3D radial acquisition also enables more efficient sampling of k-space, reducing scan times and increasing patient throughput. Employing the radial pattern, 3D radial acquisition optimizes k-space sampling efficiency and minimizes the required number of acquisitions. The non-Cartesian nature of 3D radial acquisition necessitates the development and optimization of reconstruction algorithms that can efficiently handle radial k-space data. By interleaving the radial trajectories, we are able to get the time-resolved images, and by binning the radial trajectories corresponding to different motion state, we are able to get motion-resolved reconstruction. For the UTE sequence, it enables us to capture the short T_2 tissues [17]. The key feature of UTE sequence is its ability to acquire images with extremely short echo times, typically on the order of microseconds, which enables the visualization of tissues that are otherwise "invisible" with standard MRI sequences (bSSFP or FGRE). The 3D UTE sequence also enables the possibility of providing high-resolution isotropic data, which allows for multi-planar reformatting and improved visualization of complex anatomical structures [18]. The rapid acquisition of images with 3D non-Cartesian UTE sequence reduces the sensitivity to motion artifacts [19, 20], which can be particularly beneficial when imaging areas prone to movement, such as the lungs or joints. This feature results in improved image quality and more reliable assessment of the underlying pathology.

With respect to reconstruction methods, deep learning techniques, particularly convolutional neural networks (CNNs) have achieved great success in many aspects [21, 22], and have shown promise in various medical image processing tasks, including image reconstruction [23, 24], segmentation [25, 26],

and classification [27, 28]. CNN models have demonstrated the ability to learn complex hierarchical features and patterns in image data, which has led to significant improvements in the accuracy and efficiency of medical imaging applications [29]. Specifically, deep learning-based image reconstruction methods have shown potential in addressing the limitations of conventional MRI techniques by reconstructing high-quality images from undersampled k-space data, reducing acquisition times, and improving the robustness to motion artifacts, see for recent example of the work [30, 31] and [32]. Different than the previous work, we propose a time-resolved reconstruction method without replying on priori available reference image [30], or data binning [32]. In this work, we propose to use spatial transformer networks (STNs) [33] for the reconstruction of time-resolved MRI. STNs are a class of deep learning models that are specifically designed to handle spatial transformations. The original STNs consist of a localization network, a grid generator and a sampler, which together enable the model to learn and apply spatial transformations to input data and output the data which is manipulated by the transformation. STN is known to provide improved performance by allowing the model to learn the optimal spatial transformations for a task and STN is differentiable and hence it can be trained using standard back-propagation techniques. Furthermore, STNs add robustness to the model by allowing it to focus on the most relevant parts of the image and ignore the rest, thereby reducing the impact of background noise or irrelevant details. By incorporating STNs into the MRI reconstruction process, we can address the challenges posed by motion artifacts and limited spatiotemporal resolution, ultimately enhancing the diagnostic utility of 3D cardiopulmonary MRI.

In this work, we employ a 3D center-out radial UTE sequence, in combination with a ferumoxytol-based contrast agent, for the acquisition of cardiopulmonary MRI data. The specific parameters used in the sequence are elucidated in the data acquisition portion of this paper. Subsequent to the data acquisition, we develop a method for time-resolved 3D reconstruction of cardiopulmonary MRI data based on STNs. The algorithm aims to achieve a temporal resolution of approximately 200 milliseconds and a reconstruction matrix size of $256 \times 256 \times 256$, providing highly detailed volumetric data over time. In order to accomplish this, we utilized a time-dependent STN to transform the acquired volumetric data into a temporal series of reconstructed 3D images. The implementation of the STN was carried out via a deep CNN, which facilitates complex image transformations. The input volume for this transformation process can be rendered as either modifiable, to facilitate training, or be fixed, using a static image obtained from other rapid reconstruction methods. In this investigation, we chose to fix the input image volume, using the Conjugate Gradient SENSE (CG-SENSE) reconstruction technique to ensure data consistency. Next, time-dependent forward operators, implemented using a non-uniform fast Fourier transformation, are applied to the time-resolved reconstruction to generate a set of measurements. These are then compared to the original measurements collected from the MRI scanner to ascertain the degree of divergence. This deviation, quantified as a mean-square error, is subsequently used to train the STN, enabling it to iteratively refine its reconstruction process and improve accuracy. This reconstruction algorithm operates in a fully unsupervised manner, relying solely on the undersampled k-space data acquired for each individual subject. Importantly, the training of the network does not necessitate any fully-sampled ground-truth data, rendering the method adaptable and free from pre-existing biases.

2. Methods

2.1. Background on 3D time-resolved reconstruction

Time-resolved MRI, also known as real-time MRI, aims to capture the temporal evolution of anatomical structures and physiological processes by acquiring a series of images over time. This approach enables the visualization and analysis of dynamic events, such as blood flow, cardiac motion,

and respiratory motion, which can provide valuable insights into the underlying pathophysiology of various diseases. The main challenges in 3D time-resolved MRI reconstruction are the inherently long acquisition times, motion artifacts and limited spatiotemporal resolution. These challenges arise from the need to acquire a large number of k-space data points to achieve high spatial resolution while also capturing the dynamic changes in the imaged structures. Traditional approaches, such as Cartesian sampling and breath-holding, can lead to sub-optimal image quality and restricted temporal resolution due to their sensitivity to motion artifacts and limited efficiency in sampling k-space. To overcome the challenges in 3D time-resolved MRI reconstruction, various strategies have been proposed, including novel sampling schemes and advanced reconstruction algorithms.

Mathematically speaking, we model the MRI signal as a function of the underlying spin density, relaxation times T_1, T_2 and magnetic field inhomogeneities. In the context of 3D time-resolved MRI, we consider the signal model as a function of both spatial coordinates (x, y, z) and time i:

$$S(x, y, z, i) = \rho(x, y, z) \cdot \exp(-i/T_2^*) \cdot \exp(j\phi(x, y, z, i)),$$

where $S(x, y, z, i)$ is the MRI signal, $\rho(x, y, z)$ represents the spin density, T_2^* is the effective transverse relaxation time and $\phi(x, y, z, i)$ denotes the phase term accounting for magnetic field inhomogeneities and other sources of phase variation. The k-space data can then be formulated using Fourier transform:

$$K(u, v, w, i) = \iiint S(x, y, z, i) \cdot \exp(-j2\pi(ux + vy + wz))\, dx\, dy\, dz,$$

where $K(u, v, w, i)$ represents the k-space data, and (u, v, w) are the spatial frequencies corresponding to the spatial coordinates (x, y, z). In 3D time-resolved MRI, the k-space data is acquired using various sampling trajectories, such as Cartesian, radial, or spiral, over time.

The goal of 3D time-resolved MRI reconstruction is to estimate the underlying image series $I(x, y, z, i)$ from the acquired k-space data $K(u, v, w, i)$. The image series can be obtained by applying an inverse Fourier transform to the k-space data:

$$I(x, y, z, i) = \iiint K(u, v, w, i) \cdot \exp(j2\pi(ux + vy + wz))\, du\, dv\, dw.$$

However, due to the limited spatiotemporal resolution and motion artifacts, the acquired k-space data is often incomplete or corrupted. In such cases, we usually pose the time-resolved reconstruction as a minimization problem. A general optimization framework for 3D time-resolved MRI reconstruction can be written as

$$\mathbf{I}^* = \arg\min_{\mathbf{I}} \|\mathcal{A}(\mathbf{I}) - \mathbf{K}\|^2 + \lambda \cdot \mathcal{R}(\mathbf{I}), \tag{2.1}$$

where $\mathbf{I} = [I(x, y, z, i)]$, and $\mathbf{K} = [K(u, v, w, i)]$ for all i in the time series, and $\mathcal{A}(\mathbf{I})$ represents the forward imaging model. $\mathcal{R}(\mathbf{I})$ is a regularization term that encodes prior knowledge or constraints about the image series (e.g., sparsity, smoothness, or motion consistency), and λ is a regularization parameter that balances the trade-off between data fidelity and regularization. Directly solving the above minimization problem poses a few challenges because the problem is vastly underdetermined and heavily stressing computing resources as well as memory management and storage. To overcome the challenges, [34] proposed a compressed representation using multi-scale low-rank matrix factorization for time-resolved imaging representation, and to use stochastic optimization to reduce computation. This method is termed as Extreme-MRI.

2.2. Background on STN

The STN is a differentiable network with spatial transformation capabilities. It consists of three main components: a localization network, a grid generator, and a sampler. The localization network is a sub-network that predicts the transformation parameters from the input feature map. The output of the localization network is a set of parameters that define the transformation matrix. The grid generator uses the transformation parameters output by the localization network to compute a sampling grid. This grid is a set of points that correspond to the locations from which the output will be sampled. The sampler is responsible for interpolating the input at the locations specified by the sampling grid to produce the output. The interpolation can be performed using various techniques, such as bilinear, nearest-neighbor, or bicubic interpolation. The sampler is designed to be differentiable, which allows gradients to be backpropagated through the STN module during training.

2.3. Proposed approach

In this work, we propose a time-resolved 3D cardiopulmonary MRI reconstruction framework using a modified STN. Instead of using a localization network to output the transformation parameters as the first step, we choose to use the smoothly filtered multi-channel center k-space data as the first step. The signal model $S(x, y, z, i)$ is implemented on the MR scanner and the k-space data $K(u, v, w, i)$ is encoded and obtained directly on the scanner using the receiver coils. The multi-channel center k-space data is also averaged based on the number of radial interleaves for each image frame in the time series. The processed multi-channel center k-space data is then fed into a CNN to generate the time-dependent sampling grids $\Phi_i(G)$. Based on a static image volume, which can be either trainable or fixed using some simple reconstruction (e.g., CG-SENSE [35]) from all the k-space data, the sampler produces the time-resolved reconstruction using the time-dependent sampling grids $\Phi_i(G)$. In this work, we make the static image trainable. For the training of the network, we apply the forward imaging model on the time-resolved reconstruction and compare the results with the undersampled k-space measurements. The forward model is realized using non-uniform inverse Fast Fourier Transformation. The detailed formulas for non-uniform inverse Fast Fourier Transformation can be found in [36]. We use the torchkbnufft [37] library in Python to implement the operator. The mean-square-error (MSE) loss in k-space is used for training. The whole framework is illustrated in Figure 1.

In the whole proposed framework, only the CNN parameters in the STN, and the static image volume (if we make the static image volume trainable) need to be learned based on the minimization criteria (2.1), where the images I is obtain using the output of the STN. We should note here that we are actually solving for the motion fields (which are the outputs of the STN) in the framework, rather than the time-resolved images. This make the whole framework memory-efficient because solving the images in the time-series will require way more memory as the images contain more details than the motion fields. We also applied a total variation regularization [38] term on the images during the reconstruction.

The network and optimization was implemented using PyTorch [39]. We use ADAM optimization [40] with a batch size of one time-frame for the training. The CNN in the STN is implemented using an eight-layer network. The 3D convolutional layers have 250 features per layer. ReLU activation function [41] is used for all the convolutional layers. All the experiments in this work were done on a workstation with with an Intel Xeon CPU at 2.40 GHz and a Tesla A100 80 GB GPU.

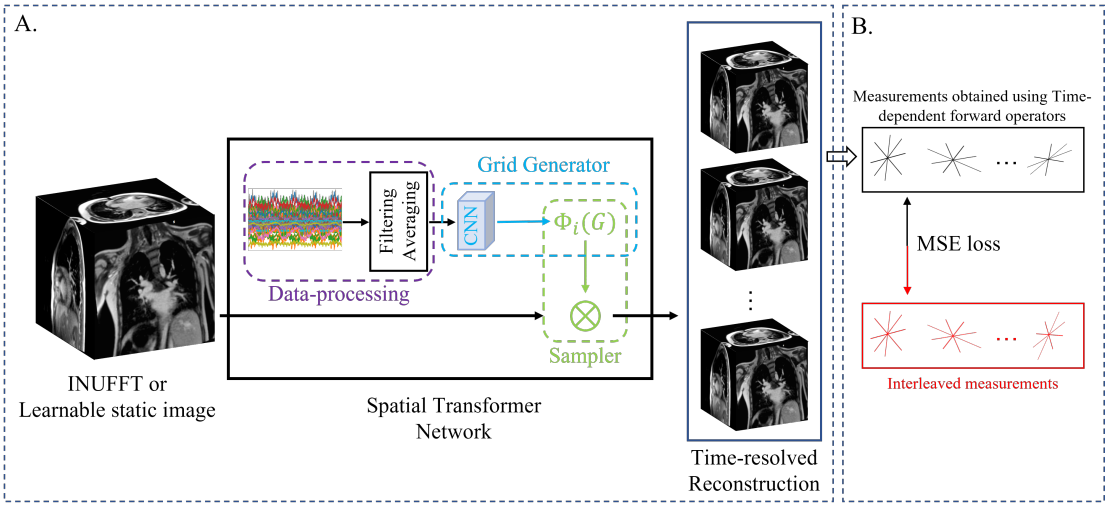


Figure 1. Illustration of the proposed scheme. We adopt the smoothly filtered multi-channel center k-space data, averaged based on radial interleaves per image frame, as the input into a CNN to generate time-dependent sampling grids, $\Phi_i(G)$. Utilizing a static image volume, obtained through INNUFT reconstruction or a trainable alternative, the sampler yields time-resolved reconstructions based on $\Phi_i(G)$. Training the network involves applying the forward imaging model to the time-resolved reconstruction, comparing outcomes with undersampled k-space measurements, and utilizing MSE loss.

The proposed framework offers multiple benefits that make it well-suited for time-resolved MRI reconstruction. First, the approach is subject-specific, as the training and reconstruction processes are tailored to each individual subject and rely exclusively on highly undersampled k-space measurements. This unsupervised nature of the scheme renders it particularly advantageous for time-resolved MRI reconstruction, where obtaining fully-sampled ground truth data for training purposes is infeasible. Second, the proposed framework demonstrates memory efficiency, as it does not necessitate the storage of all reconstructed image volumes in the time series. Instead, only a single static image volume and the STN require storage. This feature enables the high spatiotemporal resolution 3D MRI reconstruction to be performed even with limited GPU memory resources, further enhancing the practical applicability of the approach.

3. Datasets and evaluation

3.1. Data description and pre-processing

The Institutional Review Board (IRB) at the University of Wisconsin granted approval for all study procedures and protocols, adhering to the guidelines set forth by the campus Human Research Protection Program. All study processes were conducted in accordance with the Declaration of Helsinki, including the acquisition of written informed consent from each participant. A total of seven healthy volunteers participated in the study, undergoing post-ferumoxytol (4 mg/kg) free-breathing UTE acquisitions. Data were collected using a 3T GE scanner equipped with a 32-channel coil.

For post-ferumoxytol data, the scan parameters were as follows: scan duration of 5:45 minutes, flip angle of 24°, echo time (TE) of 0.25 ms and repetition time (TR) of 3.6 ms. A total of 94,957 projections

were acquired employing a 3D pseudorandom bit-reversed view ordering [42] with a readout length of 636 points per acquisition. This approach facilitated whole chest coverage with a reconstruction matrix size of $256 \times 256 \times 256$. Density compensation was normalized utilizing the maximum eigenvalue of the non-uniform fast Fourier transform (NUFFT) operator, and k-space was subsequently rescaled based on this value.

We use the principal component analysis (PCA) coil combination method to combine the 32-channel data into 8 virtual coil data for reconstruction. Specifically, PCA coil combination performs a singular value decomposition of the Hermitian matrix formed by all the coil elements. Then the approximation matrix can be found by selecting the first few number of rows of the unitary matrix. The detailed process can be found in [43]. We delete 1500 projections at the very beginning and very end of the total 94,957 projections, and bin the rest of the radial spokes into 1500 frames for time-resolved reconstruction. This corresponds to a temporal resolution of $\sim$ 200 ms for the reconstruction.

3.2. Metrics for quantitative evaluation

In this work, we compare the proposed method with the state-of-the-art time-resolved 3D MRI reconstruction method termed as "Extreme-MRI". To quantitatively compare the reconstructed image quality, we use the following metrics in this work.

1) Signal-to-noise ratio (SNR):

$$\text{SNR} = 20\log(\frac{\mu_s}{\sigma_n}),$$

where μ_s represents the mean intensity of a selected region of interest, while σ_n denotes the standard deviation of the intensity within a designated noise region. A higher SNR typically correlates with improved image quality. For our investigation, regions of interest were manually selected.

2) Contrast-to-noise ratio (CNR):

$$\text{CNR} = 20\log(\frac{|\mu_A - \mu_B|}{\sigma_n}),$$

where μ_A and μ_B correspond to the mean intensities of two distinct regions (one within the region of interest and one within the background), while σ_n signifies the standard deviation of the intensity for a selected noise region. A higher CNR generally indicates superior image quality. In our research, regions of interest were determined manually.

4. Results

4.1. Time-resolved cardiopulmonary MRI reconstruction using STN

In this section, we showcase the proposed time-resolved MRI reconstruction using STN. The proposed method is used to reconstruct the seven datasets that we collected.

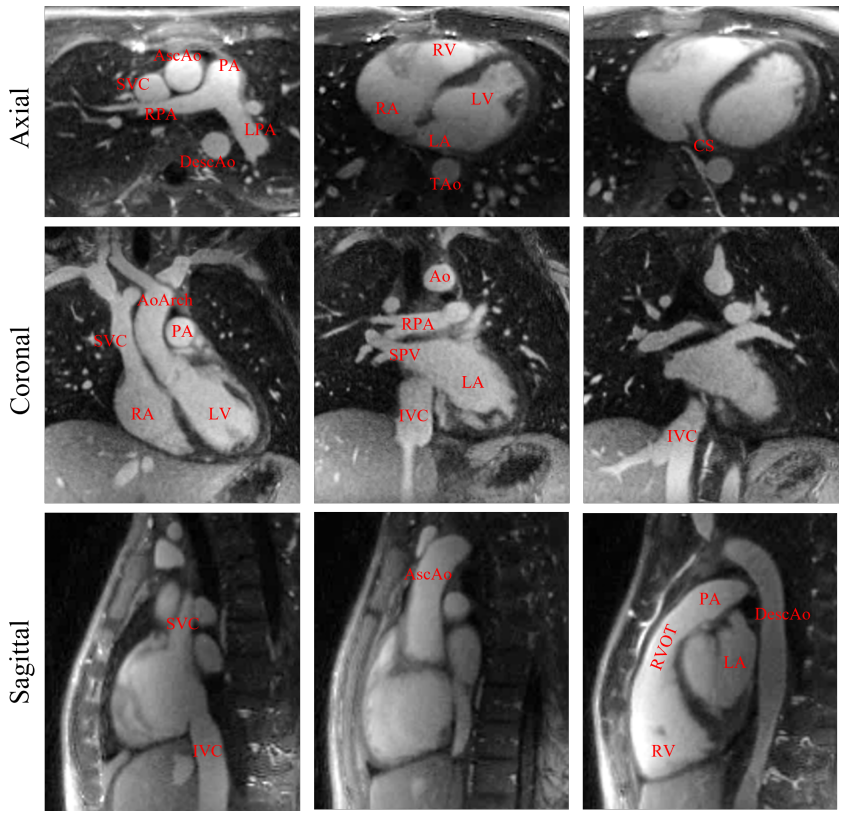


Figure 2. Showcase of the proposed reconstruction method on one of the subjects. To illustrate the reconstruction method, three representative slices were chosen for each view, namely axial, coronal, and sagittal. These views are commonly utilized in cardiac imaging to capture different aspects of the heart and surrounding structures. By selecting slices from the end-diastolic cardiac phase, which corresponds to the phase of maximum relaxation and ventricular filling, the reconstruction method aimed to capture the anatomical details during a critical point in the cardiac cycle. In the showcased images, important cardiovascular structures were labeled to provide clear visual guidance and enhance the interpretability of the results. These labels highlight anatomical landmarks and structures of clinical significance, such as the left ventricle, right ventricle, aorta, and pulmonary artery. By accurately depicting these structures in the reconstructed images, the proposed method demonstrates its ability to faithfully capture the intricate details of the cardiovascular system.

In Figure 2, we use the reconstructed images of one of the datasets to show the results. A comprehensive evaluation was conducted to demonstrate the capabilities and effectiveness of the reconstruction technique. The goal was to provide visual evidence of the reconstruction results and highlight the key cardiovascular structures that are accurately depicted in the images. We chose three representative slices from each of the three views for illustration. Important structures of the cardiovascular system, including ascending aorta (AscAo), descending aorta (DescAo), thoracic aorta (TAo), aortic arch (AoArch), main pulmonary artery (PA), right pulmonary artery (RPA), left pulmonary artery (LPA), superior pulmonary vein (SPV), superior vena cava (SVC), inferior vena cava (IVC), left ventricle (LV), right ventricle (RV), left atrium (LA), right atrium (RA), right ventricular outflow tract (RVOT) and coronary sinus (CS) are labeled in the images.

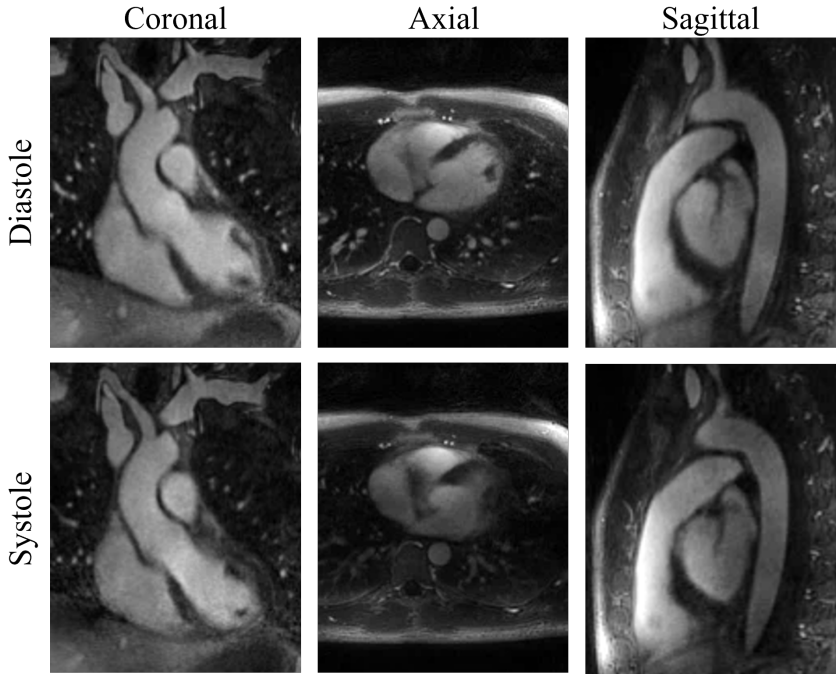


Figure 3. The proposed time-resolved reconstruction scheme showcases the capability in resolving cardiac motion. By leveraging the proposed method, it successfully captures the dynamic changes that occur during different phases of the cardiac cycle. The reconstructed diastolic and systolic phases from one subject, obtained using the proposed method, are presented in the figure.

In Figure 3, we showcase the capacity of the method to accurately capture and represent different phases of the cardiac cycle, including both diastolic and systolic phases. To achieve this, one slice was selected from each of the three views (axial, coronal and sagittal) to provide a comprehensive representation of the cardiac motion in different anatomical planes. By carefully choosing slices that span the entire cardiac cycle, the reconstruction method aimed to showcase its effectiveness in capturing the dynamic changes that occur during the heart's contraction and relaxation phases. The selected slices corresponded to both diastolic and systolic phases of the cardiac cycle. The reconstructed images, presented in the figure, clearly demonstrate the effectiveness of the proposed method in capturing various cardiac phases. This capacity to resolve cardiac motion in the reconstructed images has significant clinical implications. Accurate depiction of cardiac motion enables clinicians to assess the function of the heart, evaluate cardiac performance and identify abnormalities or irregularities in the cardiac cycle. The showcased results validate the proposed scheme's ability to faithfully capture the dynamic nature of the heart, empowering medical experts to make more informed clinical decisions based on comprehensive and accurate assessments of cardiac motion.

We further demonstrate the proposed method's ability to resolve respiratory motion in the reconstructed images, as illustrated in Figure 4. Two coronal slices at distinct time points are presented, exhibiting identical cardiac phases but differing respiratory phases. This showcases the effectiveness of the proposed scheme in capturing and distinguishing respiratory motion in the reconstructed images.

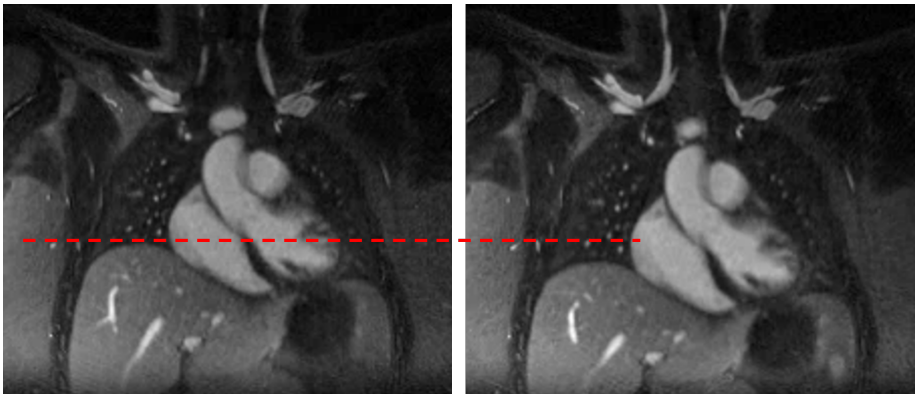


Figure 4. The proposed time-resolved reconstruction scheme is able to resolve the respiratory motion. The end-inhalation and end-exhalation phase from one subject, which are reconstructed using the proposed method, are shown in the figure. The dashed red line indicates the respiratory motion.

4.2. Comparsion with the state-of-the-art method

We compare the proposed reconstruction scheme with the state-of-the-art 3D time-resolved cardiopulmonary MRI reconstruction method termed as Extreme-MRI. Extreme-MRI aims to reconstruct time-resolved 3D cardiopulmonary MRI from highly undersampled k-space data acquired using non-Cartesian methods. Extreme-MRI employs multi-scale low-rank matrix factorization to establish a compressed representation, which simultaneously constrains the reconstruction problem and diminishes its memory requirements. It also utilizes stochastic optimization techniques to decrease computational demands, enhance memory locality and minimize communication between threads and processors.

In this work, we quantitatively and visually compared the reconstructions from the two methods. In Figure 5, we showed the visual comparison between the two method. One slice from each of the three views from one dataset are used for the comparison. From the images, we can see that the proposed method shows its capability of capturing small details such as coronary. Also, we can see from the images that the proposed method is able to catch more details such as the papillary muscle, the lung vessels and the lung-liver interfaces.

We've also invited two cardiologists with special training on advanced cardiac imaging to assess the reconstructed image quality. The cardiologists are asked to look at the reconstructed image volumes and rate the image quality based on certain important structures in the cardiovascular system. Specifically, the cardiologists are invited to looked at the right superior pulmonary vein (RSPV), right inferior pulmonary vein (RIPV), left superior pulmonary vein (LSPV), left inferior pulmonary vein (LIPV), left atrium (LA), left atrium appendage (LAA), Coronary Sinus (CS) and PA segmental branch in the reconstructed images. They are then required to rate the image quality of these structures based on a 4-scale criterion: 1—anatomy or structure NOT seen or NOT presented in the image; 2—anatomy or structure can be barely seen in the image; 3—anatomy or structure can be seen, but are blurred, or noisy; 4—anatomy or structure can be seen clearly in the image. The results are shown in Table 1. From the scores, we can see that the proposed method is able to provide state-of-the-art reconstructions, especially for the tiny details such as PA segmental branches.

Table 1. In the case of the proposed reconstruction scheme, the image quality was assessed through a systematic process involving the evaluation of seven datasets. The assessment was conducted by two cardiologists who possessed specialized training and expertise in advanced cardiac imaging. To ensure consistency and reliability in the assessment, the cardiologists followed a standardized protocol and rating system. This involved rating each dataset independently and assigning scores based on specific criteria established beforehand. The scores were then averaged to obtain an overall assessment score for each dataset.

	Proposed	Extreme-MRI	p-value
RSPV	3.86 ± 0.36	3.43 ± 0.65	0.04
RIPV	3.86 ± 0.36	3.43 ± 0.65	0.04
LSPV	3.86 ± 0.36	3.43 ± 0.65	0.04
LIPV	3.86 ± 0.36	3.43 ± 0.65	0.04
LA	4 ± 0	3.50 ± 0.52	0.001
LAA	3.64 ± 0.50	3.00 ± 0.78	0.016
CS	2.85 ± 0.77	2.78 ± 0.70	0.79
PA segmental branch	3.93 ± 0.27	3.07 ± 0.62	0.0001

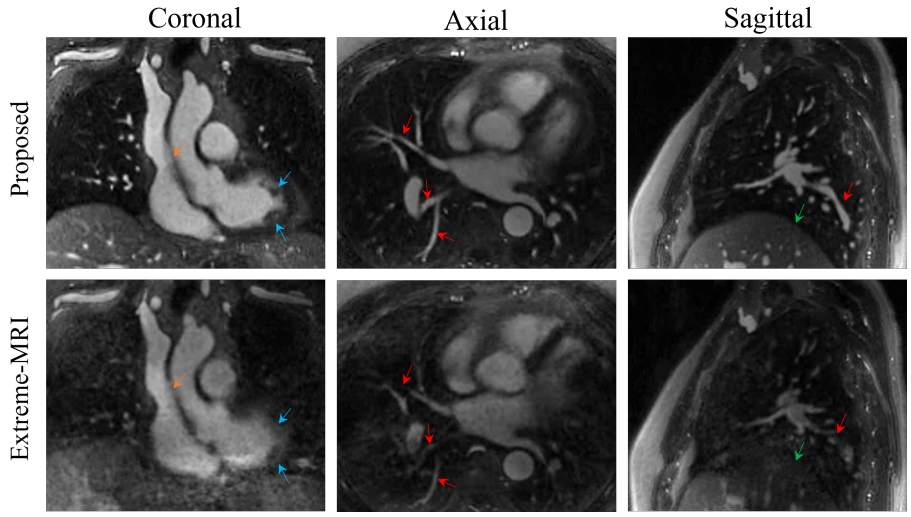
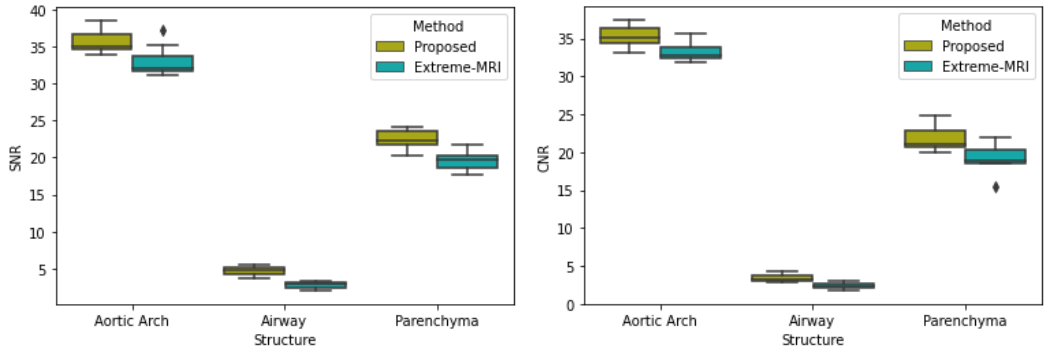


Figure 5. Visual comparison between the proposed reconstruction and Extreme-MRI reconstruction. The comparison is based on one slice of end-diastolic phase from one of the subjects. From the figure, we can see that the proposed method is able to catch more details about vascular (highlighted by the red arrows). Also, we can see that the proposed method is able to catch tiny structures such as coronary (highlighted by the orange arrows) and the papillary muscle (highlighted by blue arrows). Also, we can see that the liver-lung interface are shown more clear in the proposed reconstruction (highlighted by green arrows).

The comparison of the image quality of the reconstructions between the two methods is also done quantitatively. Specifically, we segment a section of the major airway, a section of the aortic arch and a section of lung parenchyma in all the cases, following the instructions in [44]. We then compare the SNR



(a) SNR comparison

(b) CNR comparison

Figure 6. SNR and CNR comparison. In (a), we compare the SNR of the airway, aortic arch, and lung parenchyma of all the datasets. From the comparison, we can see that the proposed method is able to provide higher SNR, corresponding to improved image quality. In (b), we show the comparison of the CNR. From the results, we can see that the proposed method is also able to provide improved reconstructions.

and CNR for the three structures between the two methods. The results are shown in Figure 6. From which we can see that the proposed method is able to provide comparable results as the state-of-the-art Extreme-MRI reconstruction.

4.3. Maximum intensity projections of the reconstructions

In this section, we present the results obtained from the proposed reconstruction method. Figure 7 displays the reconstructions derived from a single subject, generated using maximum intensity projection (MIP) [45]. MIP is recognized for its ability to effectively visualize vascular structures as tubular and branching formations within images [46]. By presenting three views of the reconstruction employing MIP, the heart structure and vascular features for the subject can be directly observed, facilitating their practical application in clinical settings. In Figure 7, each view consists of 60 slices utilized for MIP images.

5. Discussion

This work proposed a reconstruction framework for time-resolved 3D cardiopulmonary MRI, which has emerged as a valuable tool in medical imaging, offering dynamic visualization of cardiac and pulmonary structures with high spatial and temporal resolution. The implementation of a STN for reconstructing time-resolved 3D cardiopulmonary MRI offers several advantages and presents new opportunities for improving diagnostic accuracy and clinical utility.

STNs have the potential to enhance the reconstruction quality of time-resolved 3D cardiopulmonary MRI. By employing CNN, STNs can adaptively learn and apply complex, non-linear transformations to the input data, resulting in improved image quality and reduced artifacts compared to conventional reconstruction methods. Furthermore, one of the challenges in time-resolved 3D cardiopulmonary MRI is the presence of motion, which can lead to artifacts and decreased image quality. STNs inherently provide robustness to motion by learning to account for and correct motion-related inconsistencies during the reconstruction

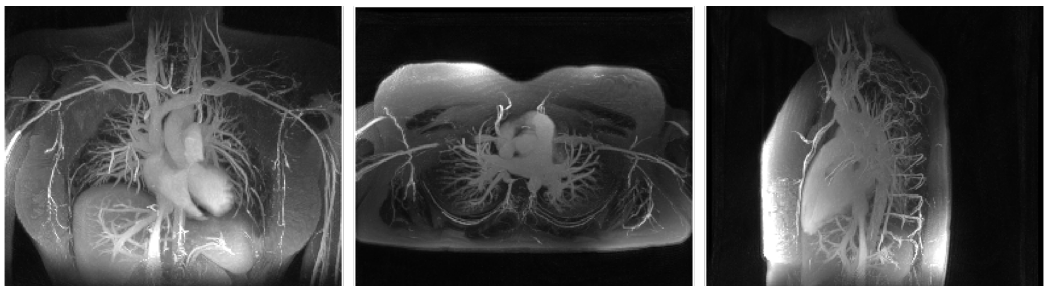


Figure 7. Showcase of the proposed method with the aim of providing visual evidence of the reconstruction results. In the showcased figures, the MIP images for each of the three views are presented, providing a holistic and comprehensive visualization of the reconstructed data. The axial MIP view allows for a comprehensive assessment of structures in a transverse plane, such as the heart, lungs, and major vessels. The coronal MIP view provides a frontal perspective, enabling a detailed examination of structures from a different angle, including the heart chambers, great vessels, and pulmonary vasculature. Finally, the sagittal MIP view offers a lateral perspective, facilitating the evaluation of structures along the longitudinal axis, such as the cardiac apex, ventricles, and associated vessels.

process. This results in clearer images that enable better visualization of anatomical structures and functional information. STNs are also capable of unsupervised learning, which means that they can be trained and perform reconstructions without requiring fully-sampled ground truth data. This characteristic is especially advantageous for time-resolved MRI, where acquiring fully-sampled data for training is often impractical or impossible. The use of STNs for time-resolved 3D cardiopulmonary MRI reconstruction also allows for memory-efficient processing. Instead of storing all reconstructed image volumes in the time series, only a single static image volume and the STN need to be stored, enabling high spatiotemporal resolution 3D MRI reconstruction even with limited GPU memory.

Experimental results have shown improved reconstruction results from the proposed method compared to the state-of-the-art reconstruction algorithm. Specifically, the proposed method is able to provide more detailed reconstructions, such as the ability to capture small vessels. Also, the the proposed method is able to provide sharper images. These improvements are also confirmed by the quantitative results (SNR and CNR). Visual comparison by the medical experts also indicates better performance.

While the application of STNs to time-resolved 3D cardiopulmonary MRI reconstruction shows promising results, several challenges and opportunities for future research remain. First of all, further exploration and optimization of STN architectures hold promising potential for enhancing reconstruction performance in the field of time-resolved 3D cardiopulmonary MRI. STN architectures are designed to leverage both spatial and temporal information in the data, enabling more accurate reconstructions. To achieve improvements in reconstruction performance, one area of exploration involves the development of novel network architectures that can effectively capture the complex spatiotemporal patterns present in cardiopulmonary MRI data. This includes investigating different layer configurations, connectivity patterns and fusion strategies to better integrate spatial and temporal information. Additionally, efforts can be made to optimize the hyperparameters and training strategies of STN models, such as exploring different loss functions, regularization techniques and data augmentation approaches. Furthermore, to ensure the clinical utility of STN-based time-resolved 3D cardiopulmonary MRI reconstruction, thor-

ough validation with real-world clinical data and evaluation by medical experts are crucial. Validation involves assessing the performance of the STN models using diverse datasets comprising patient data, including both MRI and/or CT images. This allows us to evaluate the generalizability and robustness of the reconstruction algorithms across different patient populations, pathologies, and imaging protocols.

6. Conclusions

In this study, we proposed a time-resolved 3D cardiopulmonary MRI reconstruction using spatial transformer network. The use of spatial transformer networks in time-resolved 3D cardiopulmonary MRI reconstruction presents a promising avenue for enhancing image quality, robustness to motion, and memory efficiency. Future research will focus on optimizing network architectures, integrating with other advanced imaging techniques and validating the clinical applicability of this approach.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. A. Hansell, J. Walk, J. Soriano, What do chronic obstructive pulmonary disease patients die from? Amultiple cause coding analysis, *Eur. Respir. J.*, **22** (2003), 809–814. https://doi.org/10.1183/09031936.03.00031403
2. R. Rubinsztajn, R. Chazan, An analysis of the causes of mortality and co-morbidities in hospitalised patients with chronic obstructive pulmonary disease, *Adv. Respir. Med.*, **79** (2011), 343–346. https://doi.org/10.5603/ARM.27637
3. P. Sicard, Y. O. Khaniabadi, S. Perez, M. Gualtieri, A. De Marco, Effect of O_3, PM_{10} and $PM_{2.5}$ on cardiovascular and respiratory diseases in cities of France, Iran and Italy, *Environ. Sci. Pollut. Res.*, **26** (2019), 32645–32665. https://doi.org/10.1007/s11356-019-06445-8
4. L. C. Saunders, P. J. Hughes, S. Alabed, D. J. Capener, H. Marshall, J. Vogel-Claussen, et al., Integrated cardiopulmonary MRI assessment of pulmonary hypertension, *J. Magn. Reson. Imaging*, **55** (2022), 633–652. https://doi.org/10.1002/jmri.27849
5. O. Bane, S. J. Shah, M. J. Cuttica, J. D. Collins, S. Selvaraj, N. R. Chatterjee, et al., A non-invasive assessment of cardiopulmonary hemodynamics with MRI in pulmonary hypertension, *Magn. Reson. Imaging*, **33** (2015), 1224–1235. https://doi.org/10.1016/j.mri.2015.08.005
6. J. A. Tkach, N. S. Higano, M. D. Taylor, R. A. Moore, M. Hossain, G. Huang, et al., Quantitative cardiopulmonary magnetic resonance imaging in neonatal congenital diaphragmatic hernia, *Pediatr. Radiol.*, **52** (2022), 2306–2318. https://doi.org/10.1007/s00247-022-05384-w

7. J. Y. Cheng, T. Zhang, M. T. Alley, M. Uecker, M. Lustig, J. M. Pauly, et al., Comprehensive multi-dimensional MRI for the simultaneous assessment of cardiopulmonary anatomy and physiology, *Sci. Rep.*, **7** (2017), 1–15. https://doi.org/10.1038/s41598-017-04676-8
8. A. Vonk Noordegraaf, F. Haddad, H. J. Bogaard, P. M. Hassoun, Noninvasive imaging in the assessment of the cardiopulmonary vascular unit, *Circulation*, **131** (2015), 899–913. https://doi.org/10.1161/CIRCULATIONAHA.114.006972
9. R. A. Lafountain, J. S. da Silveira, J. Varghese, G. Mihai, D. Scandling, J. Craft, et al., Cardiopulmonary exercise testing in the MRI environment, *Physiol. Meas.*, **37** (2016), N11. https://doi.org/10.1088/0967-3334/37/4/N11
10. R. Jogiya, A. Schuster, A. Zaman, M. Motwani, M. Kouwenhoven, E. Nagel, et al., Three-dimensional balanced steady state free precession myocardial perfusion cardiovascular magnetic resonance at 3T using dual-source parallel RF transmission: initial experience, *J. Cardiovasc. Magn. Reson.*, **16** (2014), 1–10. https://doi.org/10.1186/s12968-014-0090-0
11. C. Lin, M. A. Bernstein, 3D magnetization prepared elliptical centric fast gradient echo imaging, *Magn. Reson. Med.*, **59** (2008), 434–439.
12. S. Potthast, L. Mitsumori, L. A. Stanescu, M. L. Richardson, K. Branch, T. J. Dubinsky, et al., Measuring aortic diameter with different MR techniques: Comparison of three-dimensional (3D) navigated steady-state free-precession (SSFP), 3D contrast-enhanced magnetic resonance angiography (CE-MRA), 2D T2 black blood, and 2D cine SSFP, *J. Magn. Reson. Imaging*, **31** (2010), 177–184. https://doi.org/10.1002/jmri.22016
13. M. J. Kim, D. G. Mitchell, K. Ito, P. N. Kim, Hepatic MR imaging: comparison of 2D and 3D gradient echo techniques, *Abdom. Imaging*, **26** (2001), 269–276. https://doi.org/10.1007/s002610000177
14. S. J. Kruger, S. B. Fain, K. M. Johnson, R. V. Cadman, S. K. Nagle, Oxygen-enhanced 3D radial ultrashort echo time magnetic resonance imaging in the healthy human lung, *NMR Biomed.*, **27** (2014), 1535–1541. https://doi.org/10.1002/nbm.3158
15. W. Zha, S. J. Kruger, K. M. Johnson, R. V. Cadman, L. C. Bell, F. Liu, et al., Pulmonary ventilation imaging in asthma and cystic fibrosis using oxygen-enhanced 3D radial ultrashort echo time MRI, *J. Magn. Reson. Imaging*, **47** (2018), 1287–1297. https://doi.org/10.1002/jmri.25877
16. K. S. Sodhi, A. Bhatia, P. Rana, J. L. Mathew, Impact of radial percentage k-space filling and signal averaging on native lung MRI image quality in 3D radial UTE acquisition: A pilot study, *Acad. Radiol.*, **2023** (2023). https://doi.org/10.1016/j.acra.2023.01.029
17. M. D. Robson, P. D. Gatehouse, M. Bydder, G. M. Bydder, Magnetic resonance: an introduction to ultrashort TE (UTE) imaging, *J. Comput. Assisted Tomogr.*, **27** (2003), 825–846. https://doi.org/10.1097/00004728-200311000-00001
18. R. Shekhar, V. Zagrodsky, Cine MPR: interactive multiplanar reformatting of four-dimensional cardiac data using hardware-accelerated texture mapping, *IEEE Trans. Inf. Technol. Biomed.*, **7** (2003), 384–393. https://doi.org/10.1109/TITB.2003.821320
19. K. T. Block, H. Chandarana, S. Milla, M. Bruno, T. Mulholland, G. Fatterpekar, et al., Towards routine clinical use of radial stack-of-stars 3D gradient-echo sequences for reducing motion sensitivity, *J. Korean Soc. Magn. Reson. Med.*, **18** (2014), 87–106. https://doi.org/10.13104/jksmrm.2014.18.2.87

20. J. E. Park, Y. H. Choi, J. E. Cheon, W. S. Kim, I. O. Kim, Y. J. Ryu, et al., Three-dimensional radial vibe sequence for contrast-enhanced brain imaging: an alternative for reducing motion artifacts in restless children, *AJR Am. J. Roentgenol.*, **210** (2018), 876–882. https://doi.org/10.2214/AJR.17.18490
21. J. Cao, D. Zhao, C. Tian, T. Jin, F. Song, Adopting improved adam optimizer to train dendritic neuron model for water quality prediction, *Math. Biosci. Eng.*, **20** (2023), 9489–9510. https://doi.org/10.3934/mbe.2023417
22. T. Jin, F. Li, H. Peng, B. Li, D. Jiang, Uncertain barrier swaption pricing problem based on the fractional differential equation in caputo sense, *Soft Comput.*, **2023** (2023), 1–16. https://doi.org/10.1007/s00500-023-08153-5
23. Q. Zou, A. H. Ahmed, P. Nagpal, S. Kruger, M. Jacob, Dynamic imaging using a deep generative storm (gen-storm) model, *IEEE Trans. Med. Imaging*, **40** (2021), 3102–3112. https://doi.org/10.1109/TMI.2021.3065948
24. R. Hou, F. Li, IDPCNN: Iterative denoising and projecting CNN for MRI reconstruction, *J. Comput. Appl. Math.*, **406** (2022), 113973.
25. Z. Wang, R. Wu, Y. Xu, Y. Liu, R. Chai, H. Ma, A two-stage CNN method for MRI image segmentation of prostate with lesion, *Biomed. Signal Process. Control*, **82** (2023), 104610. https://doi.org/10.1016/j.bspc.2023.104610
26. D. I. Zaridis, E. Mylona, N. Tachos, V. C. Pezoulas, G. Grigoriadis, N. Tsiknakis, et al., Region-adaptive magnetic resonance image enhancement for improving CNN-based segmentation of the prostate and prostatic zones, *Sci. Rep.*, **13** (2023), 714. https://doi.org/10.1038/s41598-023-27671-8
27. S. Shojaei, M. S. Abadeh, Z. Momeni, An evolutionary explainable deep learning approach for Alzheimer's MRI classification, *Expert Syst. Appl.*, **220** (2023), 119709. https://doi.org/10.1016/j.eswa.2023.119709
28. O. Özkaraca, O. İ. Bağrıaçık, H. Gürüler, F. Khan, J. Hussain, J. Khan, et al., Multiple brain tumor classification with dense CNN architecture using brain MRI images, *Life*, **13** (2023), 349. https://doi.org/10.3390/life13020349
29. S. Hardaha, D. R. Edla, S. R. Parne, A survey on convolutional neural networks for MRI analysis, *Wireless Pers. Commun.*, **128** (2023), 1065–1085. https://doi.org/10.1007/s11277-022-09989-0
30. N. R. Huttinga, C. A. Van den Berg, P. R. Luijten, A. Sbrizzi, MR-MOTUS: model-based non-rigid motion estimation for MR-guided radiotherapy using a reference image and minimal k-space data, *Phys. Med. Biol.*, **65** (2020), 015004. https://doi.org/10.1088/1361-6560/ab554a
31. Z. Miller, L. Torres, S. Fain, K. Johnson, Motion compensated extreme MRI: Multi-scale low rank reconstructions for highly accelerated 3D dynamic acquisitions (MoCo-MSLR), preprint, arXiv:2205.00131.
32. Q. Zou, L. A. Torres, S. B. Fain, N. S. Higano, A. J. Bates, M. Jacob, Dynamic imaging using motion-compensated smoothness regularization on manifolds (MoCo-SToRM), *Phys. Med. Biol.*, **67** (2022), 144001. https://doi.org/10.1109/ISBI52829.2022.9761440
33. M. Jaderberg, K. Simonyan, A. Zisserman, Spatial transformer networks, in *Advances in Neural Information Processing Systems*, **28** (2015).

34. F. Ong, X. Zhu, J. Y. Cheng, K. M. Johnson, P. E. Larson, S. S. Vasanawala, et al., Extreme MRI: Large-scale volumetric dynamic imaging from continuous non-gated acquisitions, *Magn. Reson. Med.*, **84** (2020), 1763–1780. https://doi.org/10.1002/mrm.28235
35. O. Maier, S. H. Baete, A. Fyrdahl, K. Hammernik, S. Harrevelt, L. Kasper, et al., CG-SENSE revisited: Results from the first ISMRM reproducibility challenge, *Magn. Reson. Med.*, **85** (2021), 1821–1839. https://doi.org/10.1002/mrm.28569
36. N. A. Gumerov, R. Duraiswami, Fast multipole method based filtering of non-uniformly sampled data, preprint, arXiv:1611.09379.
37. M. J. Muckley, R. Stern, T. Murrell, F. Knoll, Torchkbnufft: A high-level, hardware-agnostic non-uniform fast fourier transform, in *ISMRM Workshop on Data Sampling & Image Reconstruction*, (2020), 22.
38. D. Strong, T. Chan, Edge-preserving and scale-dependent properties of total variation regularization, *Inverse Probl.*, **19** (2003), S165. https://doi.org/10.1088/0266-5611/19/6/059
39. S. Imambi, K. B. Prakash, G. Kanagachidambaresan, Pytorch, in *Programming with TensorFlow: Solution for Edge Computing Applications*, (2021), 87–104. https://doi.org/10.1007/978-3-030-57077-4
40. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, preprint, arXiv:1412.6980.
41. A. F. Agarap, Deep learning using rectified linear units (RELU), preprint, arXiv:1803.08375.
42. K. M. Johnson, S. B. Fain, M. L. Schiebler, S. Nagle, Optimized 3D ultrashort echo time pulmonary MRI, *Magn. Reson. Med.*, **70** (2013), 1241–1250. https://doi.org/10.1002/mrm.24570
43. M. Buehrer, K. P. Pruessmann, P. Boesiger, S. Kozerke, Array compression for MRI with large coil arrays, *Magn. Reson. Med.*, **57** (2007), 1131–1139. https://doi.org/10.1002/mrm.21237
44. X. Zhu, M. Chan, M. Lustig, K. M. Johnson, P. E. Larson, Iterative motion-compensation reconstruction ultra-short TE (iMoCo UTE) for high-resolution free-breathing pulmonary MRI, *Magn. Reson. Med.*, **83** (2020), 1208–1221. https://doi.org/10.1002/mrm.27998
45. S. Napel, M. P. Marks, G. D. Rubin, M. D. Dake, C. H. McDonnell, S. M. Song, et al., CT angiography with spiral CT and maximum intensity projection, *Radiology*, **185** (1992), 607–610. https://doi.org/10.1148/radiology.185.2.1410382
46. J. F. Gruden, S. Ouanounou, S. Tigges, S. D. Norris, T. S. Klausner, Incremental benefit of maximum-intensity-projection images on observer detection of small pulmonary nodules revealed by multidetector CT, *Am. J. Roentgenol.*, **179** (2002), 149–157. https://doi.org/10.2214/ajr.179.1.1790149