Received: 28 September 2022 | Revised: 17 May 2023 | Accepted: 10 June 2023

DOI: 10.1002/mp.16591

MEDICAL PHYSICS

RESEARCH ARTICLE

Complementation-reinforced network for integrated reconstruction and segmentation of pulmonary gas MRI with high acceleration

Zimeng Li[1,2] | **Sa Xiao**[2,3] | **Cheng Wang**[2,3] | **Haidong Li**[2,3] | **Xiuchao Zhao**[2,3] | **Qian Zhou**[2] | **Qiuchen Rao**[2] | **Yuan Fang**[2] | **Junshuai Xie**[2] | **Lei Shi**[2,3] | **Chaohui Ye**[1,2,3,4] | **Xin Zhou**[2,3,4]

[1]School of Physics, Huazhong University of Science and Technology, Wuhan, China

[2]State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan National Laboratory for Optoelectronics, Wuhan, China

[3]University of Chinese Academy of Sciences, Beijing, China

[4]Key Laboratory of Biomedical Engineering of Hainan Province, School of Biomedical Engineering, Hainan University, Haikou, China

Correspondence
Xin Zhou, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, 30 West Xiaohongshan, Wuhan, 430071, China.
Email: xinzhou@wipm.ac.cn

Funding information
National key Research and Development Project of China, Grant/Award Number: 2018YFA0704000; National Natural Science Foundation of China, Grant/Award Numbers: 82127802, 21921004, 82001915; Key Research Program of Frontier Sciences, Grant/Award Number: ZDBS-Y-JSC004; Hubei Provincial Key Technology Foundation of China, Grant/Award Number: 2021ACA013; China Postdoctoral Science Foundation, Grant/Award Numbers: 2020M672454, 2020T130125ZX, 2022M713218; Tencent Foundation: the XPLORER PRIZE

Abstract

Background: Hyperpolarized (HP) gas MRI enables the clear visualization of lung structure and function. Clinically relevant biomarkers, such as ventilated defect percentage (VDP) derived from this modality can quantify lung ventilation function. However, long imaging time leads to image quality degradation and causes discomfort to the patients. Although accelerating MRI by undersampling k-space data is available, accurate reconstruction and segmentation of lung images are quite challenging at high acceleration factors.

Purpose: To simultaneously improve the performance of reconstruction and segmentation of pulmonary gas MRI at high acceleration factors by effectively utilizing the complementary information in different tasks.

Methods: A complementation-reinforced network is proposed, which takes the undersampled images as input and outputs both the reconstructed images and the segmentation results of lung ventilation defects. The proposed network comprises a reconstruction branch and a segmentation branch. To effectively exploit the complementary information, several strategies are designed in the proposed network. Firstly, both branches adopt the encoder-decoder architecture, and their encoders are designed to share convolutional weights for facilitating knowledge transfer. Secondly, a designed feature-selecting block discriminately feeds shared features into decoders of both branches, which can adaptively pick suitable features for each task. Thirdly, the segmentation branch incorporates the lung mask obtained from the reconstructed images to enhance the accuracy of the segmentation results. Lastly, the proposed network is optimized by a tailored loss function that efficiently combines and balances these two tasks, in order to achieve mutual benefits.

Results: Experimental results on the pulmonary HP ^{129}Xe MRI dataset (including 43 healthy subjects and 42 patients) show that the proposed network outperforms state-of-the-art methods at high acceleration factors (4, 5, and 6). The peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and Dice score of the proposed network are enhanced to 30.89, 0.875, and 0.892, respectively. Additionally, the VDP obtained from the proposed network has good correlations with that obtained from fully sampled images ($r = 0.984$). At the highest acceleration factor of 6, the proposed network promotes PSNR, SSIM,

Zimeng Li and Sa Xiao contributed equally to this work.

and Dice score by 7.79%, 5.39%, and 9.52%, respectively, in comparison to the single-task models.

Conclusion: The proposed method effectively enhances the reconstruction and segmentation performance at high acceleration factors up to 6. It facilitates fast and high-quality lung imaging and segmentation, and provides valuable support in the clinical diagnosis of lung diseases.

KEYWORDS

hyperpolarized gas MRI, image reconstruction, lung segmentation, multi-task learning, ventilation defects

1 | INTRODUCTION

Magnetic Resonance Imaging (MRI) is a radiation-free and noninvasive medical imaging technique, which provides images with high resolution and excellent soft tissue contrast.[1,2] Hyperpolarized (HP) gas MRI (such as ^{3}He or ^{129}Xe) enables visualization of lung structure and function.[3] Particularly, HP ^{129}Xe MRI can offer gas-gas and gas-blood exchange information, which is helpful for the clinical diagnosis of lung diseases such as coronavirus disease 2019 (COVID-19).[4–6] Automatic analysis (e.g. segmentation) of the gas MRI can obtain clinically relevant parameters, which quantitatively evaluate pulmonary pathologies. However, HP gas MRI data are generally acquired during a single breath-hold, and the acquisition time is approximately 20 s. Such long breath-holding time will cause discomfort to the patients, especially for those with compromised respiratory function. Moreover, long imaging time can degrade image quality because the longitudinal magnetization of HP noble gases decays over time.[7] Accordingly, it is important to achieve fast and accurate imaging and segmentation of gas MRI.

To shorten the imaging time, the raw k-space data are usually undersampled. The undersampling schemes cause severe aliasing artifacts in the undersampled images. Therefore, various efforts have focused on developing reconstruction methods to remove undersampling artifacts and improve image quality. Compressed sensing (CS) is one representative approach, which exploits the sparsity of signals in a specific transform domain to recover the fully sampled images from undersampled k-space data through nonlinear reconstructions.[8] Ajraoui and coworkers first applied the CS algorithm to accelerate pulmonary HP gas MRI in 2010,[9] and then they combined some prior knowledge to further improve acceleration factor (AF) and reconstruction performance.[10] Moreover, CS was used to reconstruct pulmonary dynamic HP gas MR images with high temporal and spatial resolution.[11] However, there are limitations in conventional CS techniques, such as relatively low reconstruction speed and difficulty in tuning the weighting parameters.[12]

Recently, deep learning has shown great potential in accelerating HP gas MRI and exhibits superior performance both in reconstruction quality and speed. Deep learning-based reconstruction algorithms can learn the mapping between the undersampled and fully sampled MRI data.[13–15] Convolutional neural network (CNN) is a popular deep learning-based reconstruction model.[16] Duan and coworkers first introduced deep learning into HP gas MRI reconstruction.[17] They adopted a cascaded U-Net model incorporating prior knowledge of ^{1}H images (CasNet) to fast and accurately reconstruct pulmonary ^{129}Xe MR images from undersampled images. Then, a deep cascade of residual dense network (DC-RDN) was proposed to accelerate multiple *b*-value HP gas diffusion-weighted MRI (DW-MRI).[18]

Although existing CNN-based HP gas MRI reconstruction methods outperform conventional CS algorithms, they still contain some challenges. Firstly, current gas MRI reconstruction networks require a large amount of training data to ensure the accuracy of the reconstruction, especially for high AFs. However, HP gas MRI data are limited because gas MRI has not been widely used in the clinic. This results in the low generalization capability of the current HP gas MRI reconstruction networks at high AFs. Secondly, these models only focus on the reconstruction quality and do not consider the downstream applications (such as segmentation) used to obtain clinically relevant parameters. Therefore, these reconstruction networks may lose some details that are crucial for the segmentation but less influential to overall image quality.[19]

Ventilation defect percentage (VDP) is an important quantitative metric to evaluate lung function, which is defined as the ratio of the ventilation defect volume in pulmonary HP gas MR images to the thoracic volume in ^{1}H MR images.[20] To facilitate the computation of VDP, it is necessary to accurately segment the ventilation defect region of the lung.[21] Current HP gas MR image segmentation algorithms typically take reconstructed images as input without considering the reconstruction process.[22,23] This results in the segmentation performance that highly depends on the quality of the reconstructed images. However, the loss of information in the reconstruction process is inevitable, especially at high AFs, even with advanced reconstruction methods.[24–28] Existing HP gas MRI reconstruction methods mainly perform reconstruction at an AF of 4. For higher AFs (e.g., 6), although they can restore lung structure to some extent, many fine details are lost.[17,18] This will negatively influence the segmentation results. Therefore, it is desirable to develop an efficient method to concurrently perform reconstruction and segmentation of HP gas MRI.

Multi-task learning is a type of learning algorithm that aims to leverage multiple sources of information from different tasks to enhance the generalization ability of the model. Previous studies have attempted to simultaneously perform reconstruction and segmentation on ^{1}H MRI such as cardiac and liver. Huang and coworkers proposed a joint network called Joint-FR-Net for concurrent cardiac MRI reconstruction and myocardium contour segmentation.[26] Sui and coworkers developed a multi-task learning-based framework (called RecSeg), which employed two cascaded U-Nets to handle liver MRI reconstruction and lesion segmentation at the same time.[27] Their results have demonstrated that multi-task learning can enhance the performance of ^{1}H MRI reconstruction and segmentation. Building on these previous studies, we believe that the multi-task learning strategy can address the issue of unsatisfactory performance in HP gas MRI reconstruction and segmentation at high AFs. The main challenge in establishing a multi-task learning framework for HP gas MRI lies in effectively utilizing the complementary information between the undersampling reconstruction and the segmentation of lung ventilation defects. Moreover, HP gas MRI is more susceptible to noise than ^{1}H MRI due to the nonrenewability of HP magnetization,[17] making it challenging to extract accurate features at high AFs.

In this work, we propose a complementation-reinforced network for joint reconstruction and segmentation of pulmonary gas MRI (called RS-Net) using multi-task learning. The RS-Net adopts a dual-branch encoder-decoder architecture, with one branch dedicated to the reconstruction and the other to the segmentation of lung ventilation defects. Complementary information is reinforced and utilized in multiple aspects of the proposed network. Through the weight-sharing of the two encoders, the encoding process allows for knowledge transfer between the two tasks. Moreover, a feature-selecting block is devised to maximize the benefits of feature sharing between the two tasks by picking the most relevant features for each task. Additionally, the lung mask generated from the reconstructed image is fed to the segmentation branch for further improving segmentation results. In this way, RS-Net enables mutual assistance between the two tasks, resulting in improved results for both tasks. Furthermore, RS-Net achieves segmentation of the ventilation defects directly from undersampled k-space data. This can retain more original information that contributes to segmentation, thereby improving performance at high AFs. Accordingly, RS-Net would be beneficial for the clinical diagnosis of lung diseases, especially for patients who have difficulty holding their breath for a long time.

2 | MATERIALS AND METHODS

2.1 | Problem formulation

Let $\mathbf{x} \in \mathbb{C}^{M \times N}$ denotes the 2D fully sampled HP gas MR image, and $\mathbf{y}_u \in \mathbb{C}^{T \times N}$ ($T << M$) denotes the undersampled k-space data, which can be described as follows:

$$\mathbf{y}_u = \mathbf{u} \odot f_{2D}(\mathbf{x}) + \mathbf{n} \tag{1}$$

where $\mathbf{u}$ represents a binary undersampling mask, $\mathbf{n}$ represents the acquisition noise, $\mathbb{C}$ represents the complex-valued vector space, $\odot$ represents element-wise multiplication and f_{2D} represents 2D Fourier transform.

Our purpose is to reconstruct the fully sampled image $\mathbf{x}$ and predict the ventilation defect region $\mathbf{s}$ from the undersampled k-space data. To simultaneously solve the two issues, we construct a complementation-reinforced network (RS-Net). The network aims to find the following minimization:

$$\underset{\theta}{\operatorname{argmin}}\, L(\mathbf{x},\ \mathbf{s}, H(f_{2D}^{-1}(\mathbf{y}_u); \theta)) \tag{2}$$

where H denotes the RS-Net model, θ denotes hidden parameters in the model, L denotes the loss function of the model, and f_{2D}^{-1} denotes the 2D inverse Fourier transform.

2.2 | Proposed complementation-reinforced network RS-Net

In this section, we provide the detailed description of the proposed complementation-reinforced network (RS-Net), which combines the reconstruction and segmentation of HP gas MRI in a unified model for allowing the two tasks to benefit from each other. Figure 1 illustrates the architecture of the proposed RS-Net. It consists of two branches: Recon Branch and Seg Branch for reconstruction and segmentation of HP gas MRI, respectively.

Recon Branch and Seg Branch are both based on the U-Net architecture, which can transfer finely detailed features from the shallower layers of the encoder to the decoder via skip connections.[29–31] The real and imaginary parts of the Zero-filling complex image are

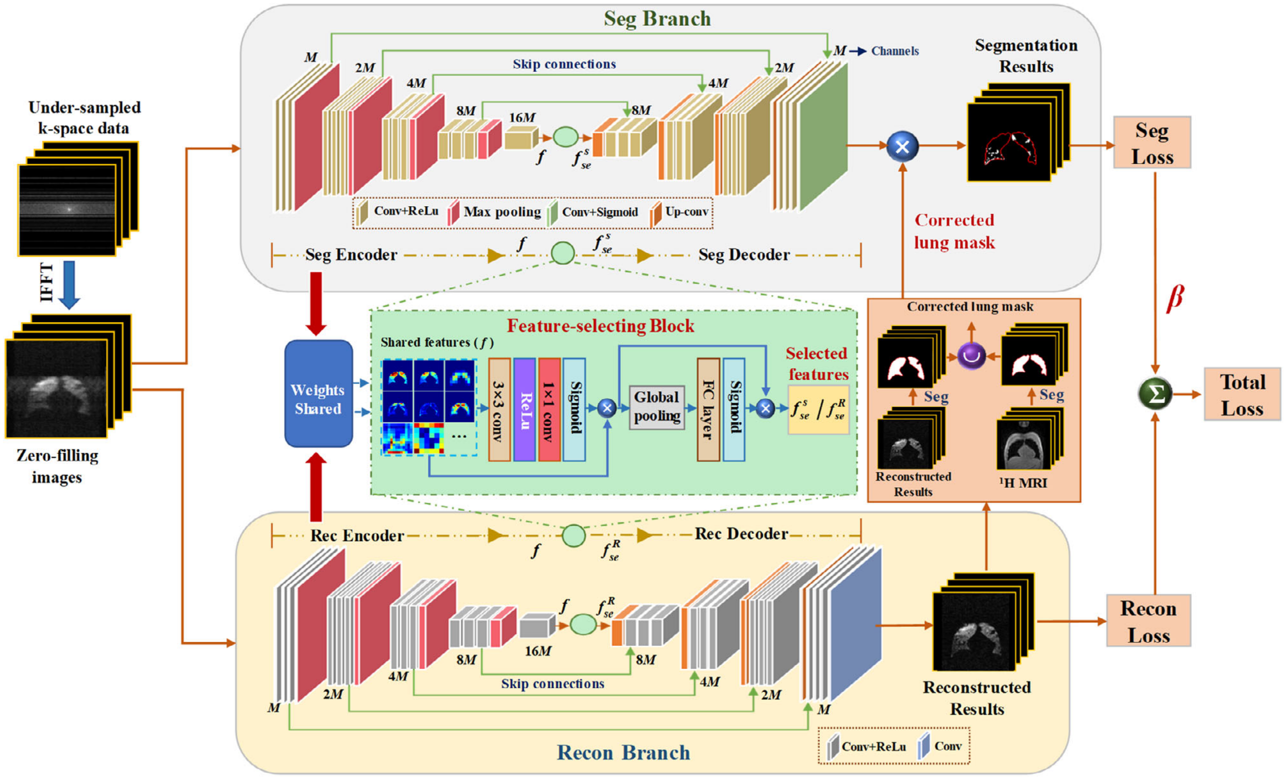

FIGURE 1 The overview of the RS-Net. The RS-Net is comprised of two branches, namely Recon Branch and Seg Branch. The two branches are both in encoder-decoder form and their encoders share weights. Feature-selecting block is designed to adaptively choose suitable features for different tasks from the shared features f. Then Recon Decoder and Seg Decoder take the selected features f_{se}^{R} and f_{se}^{S} as inputs, respectively, to generate the reconstructed images and segmentation results. To improve the accuracy of the segmentation result, the lung mask obtained from the reconstructed image is used as supplementary information to correct the lung mask obtained from the ^{1}H MRI.

concatenated to form a two-channel input map that is fed to the encoders in both branches. The encoder employs a series of contracting steps to generate multi-scale encoded features. The number of feature channels in the first contracting step is M. Then, the channel dimension of the feature is doubled and the spatial resolution is halved in each step. The feature extracted by the encoders in each branch can be formulated by:

$$f_r = H_e^r(f_{2D}^{-1}(\mathbf{y}_u); \theta_e^r) \tag{3}$$

$$f_s = H_e^s(f_{2D}^{-1}(\mathbf{y}_u); \theta_e^s) \tag{4}$$

where H_e^r and H_e^s denote Recon Encoder and Seg Encoder. θ_e^r and θ_e^s denote the parameters in the H_e^r and H_e^s, respectively. To facilitate feature sharing between the reconstruction and segmentation tasks, we enforced weight sharing between the two encoders. That is $\theta_e^r = \theta_e^s$.

The shared features contain complex information from the two tasks, and it is important to carefully select reliable information for different tasks to effectively leverage them. To this end, we propose a feature-selecting block that selects discriminative features for each task.

The detailed architecture of the feature-selecting block is illustrated in Figure 1 (the green container). The shared features are represented as $f \in \mathrm{R}^{H\times W\times C}$, where $f = f_r = f_s$, $H \times W$ means the spatial dimension, and C means the channel dimension. First, a pixel-wise weight $\alpha \in \mathrm{R}^{H\times W\times C}$ is learned by a 3×3 convolutional layer followed by rectified linear unit (ReLu) activation and a 1×1 convolutional layer followed by sigmoid activation. Then f is weighted by α, which can be formulated by:

$$f_\mathrm{p} = f \otimes \alpha, \alpha \in (0, 1) \tag{5}$$

where f_p is the pixel-wise weighted features, and $\otimes$ denotes the element-wise multiplication. Further, a channel-wise weight $\gamma \in \mathrm{R}^{1\times 1\times C}$ is learned by a global pooling layer and a fully connected layer followed by sigmoid activation. Finally, f_p is channel weighted by γ to obtain the selected features f_se, which can be formulated by:

$$f_{se} = \mathbf{F}_\mathrm{c}\left(f_\mathrm{p}, \gamma\right), \gamma \in (0, 1) \tag{6}$$

where $\mathbf{F}_\mathrm{c}(\cdot)$ is a channel-wise multiplication for feature map channels and corresponding channel weights. By learning different values of α and γ, Recon Branch and Seg Branch identify the task-specific features (f_{se}^{R} and

f_{se}^{S}) that make the most significant contributions to their respective objectives.

Subsequently, the selected features f_{se}^{R} are used by the Recon Decoder to produce the reconstructed images. Meanwhile, the selected features f_{se}^{S} are utilized by the Seg Decoder to generate the segmentation results. The architectures of the Recon Decoder and the Seg Decoder are the mirror structures of the encoder, containing multiple expansive steps and an output layer for obtaining the target results. Both decoders can reuse the features extracted by the encoders through skip connections, allowing the two tasks to regularize each other. The main difference between the two decoders is the output layer. The output layer of Recon Decoder is a 1×1 convolutional layer without activation. While Seg Decoder adds a sigmoid activation after the 1×1 convolutional layer. Moreover, to constrain the segmentation in the lung region, the output of the sigmoid activation layer is multiplied with the lung mask.

In general, the lung mask is obtained from lung ^{1}H MRI. However, due to noise and the influence of the trachea during the actual acquisition process, some lung regions in the ^{1}H MRI image may be unclear or incomplete. These missing areas of the lung in ^{1}H MRI may be contained in the ^{129}Xe MR images. Therefore, we segment the reconstructed ^{129}Xe MR images using the seeded region growing method to obtain a complementary lung mask. Then we merge this mask with the lung mask obtained from the ^{1}H MRI images via the union operation, which yields a corrected lung mask. In this way, the RS-Net can achieve more precise segmentation results for ventilation defects. This could facilitate more accurate clinical diagnoses of lung diseases.

During the training, the segmentation task enforces the network to focus more on high-level semantic features, as they are crucial for predicting correct ventilation defect regions. This benefits RS-Net to reconstruct images with more details, especially in ventilation defect regions. Moreover, under the constraint of the reconstruction task, the network is powerful in suppressing artifacts, leading to clearer and more precise segmentation results. Accordingly, in the proposed RS-Net, the two tasks will complement and reinforce each other and then achieve better results at the same time.

2.3 | Loss function

To find an optimal balance between reconstruction and segmentation tasks, we design a weighted loss to train RS-Net. The loss of the RS-Net is composed of two parts: the reconstruction loss L_r and the segmentation loss L_s. For the reconstruction task, the target objective function can be defined as,

$$\underset{\theta_e^r,\theta_d^r}{\operatorname{argmin}} L_r(\mathbf{x}, H(f_{2D}^{-1}(\mathbf{y}_u); \theta_e^r, \theta_d^r)) \tag{7}$$

TABLE 1 The basic information of the study subjects.

Clinical indications	Age (mean ± SD)	Subject number	
		Male	Female
Healthy	35.5 ± 13.5	25	18
Asthma	43.5 ± 23.3	0	2
COPD	58.0 ± 12.9	8	1
Bronchiectasis	60.5 ± 24.7	2	0
Chronic inflammation	59.6 ± 11.3	9	2
Pulmonary tuberculosis	66.5 ± 21.9	1	1
Pulmonary nodule	52.7 ± 11.7	7	9
Total	45.7 ± 16.5	52	33

Abbreviation: SD, standard deviation.

where θ_e^r and θ_d^r denote the parameters of the Recon Encoder and Recon Decoder, respectively. The reconstruction loss L_r is defined as a sum of L_1 and L_2 loss functions, which not only prevents over-smoothing but ensures robust convergence.[32]

For the segmentation task, the target objective function can be defined as,

$$\underset{\theta_e^s,\theta_d^s}{\operatorname{argmin}} L_s(\mathbf{s}, H(f_{2D}^{-1}(\mathbf{y}_u); \theta_e^s, \theta_d^s)) \tag{8}$$

where θ_e^s and θ_d^s denotes the parameters of the Seg Encoder and Seg Decoder, respectively. The segmentation loss L_s is the Dice loss function.[33]

The total loss of the RS-Net is the weighted sum of the two losses, which is described as:

$$L_{total} = L_r + \beta L_s \tag{9}$$

where β is a weighting parameter, which controls the influence of the segmentation loss on the total loss.

The RS-Net is trained in an end-to-end manner, so the parameters of the model are updated simultaneously. The key issue in the training process is to find the optimal hyperparameter of the loss (β) so that the two tasks are mutually beneficial. The selection of β will be discussed in section 2.7.

2.4 | Data acquisition and preprocessing

85 subjects in total were enrolled for the experimental evaluations including 43 healthy subjects and 42 patients with various pulmonary pathologies such as chronic obstructive pulmonary disease (COPD), asthma, and pulmonary nodule (see Table 1 for the details about subjects' information). All the experiments were approved by the local Institutional Review Board (IRB) and all subjects provided informed consents. MRI scans were conducted using a 1.5 Tesla whole-body MRI scanner (Avanto, Siemens Medical Solutions). Enriched

^{129}Xe gas was polarized by a commercial xenon polarizer (verImagin Healthcare, Wuhan) based on the Rb-^{129}Xe spin-exchange optical pumping (Rb-^{129}Xe SEOP) method. Then 500 mL hyperpolarized ^{129}Xe thawed into a Tedlar bag was mixed with 500 mL medical-grade N_2 gas to generate 1 L gas mixture. The available polarization of ^{129}Xe in the Tedlar bag was approximately 25%. All subjects inhaled the gas mixture from functional residual capacity and then held their breath for data acquisition.

A home-built transmit-receive vest radiofrequency (RF) coil was used for HP ^{129}Xe imaging and a volume coil was used for ^{1}H imaging. The MRI parameters for pulmonary HP ^{129}Xe imaging were as follows: 3D bSSFP sequence, matrix size = 96 × 84, repetition time (TR) = 4.2 ms, echo time (TE) = 1.9 ms, number of slices = 24, slice thickness = 8 mm, flip angle = 10°, scan time = 8.4 s. In the same breath-hold, anatomic ^{1}H images were acquired after pulmonary ^{129}Xe MR images. The parameters for ^{1}H imaging were as follows: 3D FLASH sequence, matrix size = 96 × 84, TR = 2.4 ms, TE = 0.7 ms, number of slices = 24, slice thickness = 8 mm, flip angle = 5°, scan time = 2 s. ^{129}Xe and ^{1}H MR images were aligned through affine registration before further experiments.[21]

To confirm the image quality, we excluded slices outside the lung region and slices with signal-to-noise ratio (SNR) lower than 6.6. Since the noise distribution of an MR image is Rician, SNR was calculated by[34]:

$$\mathrm{SNR} = \frac{\mathrm{Mean_{signal}} - \mathrm{Mean_{noise}}}{\mathrm{Std_{noise}}} \times \sqrt{2 - \frac{\pi}{2}} \qquad (10)$$

where $\mathrm{Mean_{signal}}$ is the average signal within the lung, $\mathrm{Mean_{noise}}$ and $\mathrm{Std_{noise}}$ are the mean value and standard deviation of the background noise outside the lung. In this way, we excluded 1111 images and finally obtained a total of 929 HP ^{129}Xe MR images for experiments. The images were padded to 96 × 96 to fit the U-Net architecture. We randomly selected 68 subjects for training, 8 subjects for validation, and another 9 subjects for testing. Training images were augmented with horizontal flips and the 90°, 180°, and 270° rotations. In this way, the number of training images was augmented 4 folds. Cartesian undersampling scheme was used in this work and the AFs were 4, 5, and 6. Specifically, the fully sampled k-space data was undersampled in the phase-encoding direction according to a variable-density Cartesian random undersampling matrix. The matrix was generated by the Monte Carlo algorithm.[11]

The ground truth of segmentation of the ventilation defect region was obtained by classifying fully sampled HP ^{129}Xe MR images with the K-means clustering algorithm (the commonly used method for segmentation of the ventilation defect region).[21] The lung mask from ^{1}H MRI was initially segmented by the seeded region growing method,[35] and manually corrected by two experts (with more than 5 years of experience in pulmonary imaging). Then the lung mask obtained from the fully sampled HP ^{129}Xe MR images was used to supplement the lung mask obtained in the previous step.

2.5 | Implementation details

RS-Net was implemented using Tensorflow2.0 on a workstation with 3 GPU NVIDIA RTX 2080Ti, and Intel Xeon (R) Gold 6128 central processing unit (CPU). The number of initial feature channels (M) was set to 64. Due to the complexity of training a multi-task network, we employed different learning rates for the various blocks of RS-Net for better convergence. An initial learning rate of 0.0005 was applied for both encoders and the Recon Decoder, which exponentially decayed with the increasing training epochs. The initial learning rate of Seg Decoder was set to 0.0002 and it also exponentially decayed with the increasing training epochs. Adam optimizer was used to optimize RS-Net. The batch size was set to 16 and the training was stopped after 200 epochs.

2.6 | Performance evaluation

To evaluate the effectiveness of the multi-task learning for the reconstruction task, we compared RS-Net with a plain U-Net for HP gas MRI reconstruction (called Single-Recon), Zero filling, CS method,[9] and three state-of-the-art deep learning-based methods, including CasNet,[17] KIKI-net,[36] and a deep cascade CNNs (referred to as Deep-Cascade).[37] To evaluate the segmentation performance of RS-Net, we compared RS-Net with a plain U-Net for segmentation of ventilation defect region from undersampled k-space data (called Single-Seg), and reconstruction-based segmentation methods, namely Reconstruction-k (including Zero filling-k, CS-k, CasNet-k, KIKI-net-k, and Deep-Cascade-k). The Reconstruction-k method means that K-means clustering method is used to segment reconstructed images of various state-of-the-art reconstruction methods. Additionally, as an upper bound, we employed a U-Net for the direct segmentation of fully sampled images (called FS_U-Net).

Quantitative results of reconstruction were evaluated in terms of peak signal-to-noise ratio (PSNR) and structure similarity (SSIM). PSNR is a widely used image quality metric, which is related to the mean squared error (MSE).[38] SSIM is used to evaluate the structural similarity and detailed features of two images.[39] In a pulmonary gas MR image, the background and the lung region have almost the same area, but only the lung region is helpful for clinical diagnosis. Therefore, both PSNR and SSIM were only computed over the lung region for better evaluation of reconstruction performance.[17] Higher PSNR and SSIM indicate better reconstruction results.

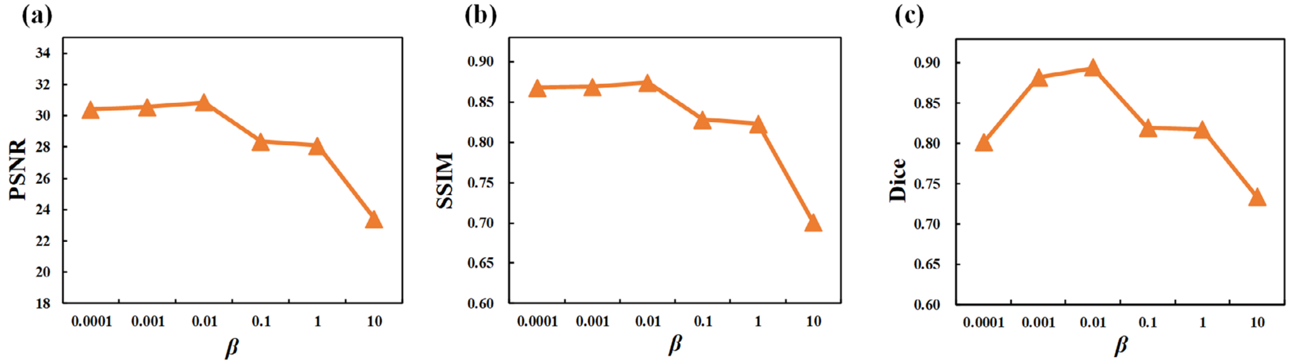

FIGURE 2 The average PSNR, SSIM, and Dice scores of RS-Net with different β. β goes from 0.0001 to 10.

Dice score was adopted to evaluate the segmentation performance, which is defined as:

$$\text{Dice}\,(S, G) = \frac{2|S \cap G|}{|S| + |G|} \tag{11}$$

where S is the segmentation result and G is the ground truth. The higher Dice score indicates better segmentation results. In addition, to validate the clinical potential of the RS-Net, we compared the VDP values calculated from various segmentation methods with that calculated from fully sampled images.

2.7 | The selection of β

To find an appropriate β, we trained a series of RS-Net with different β and evaluated their performance on the same datasets at an AF of 4. The average PSNR, SSIM, and Dice scores are shown in Figure 2. We can observe that the network performance is heavily influenced by the weights of segmentation loss (β). When $\beta = 0.01$, the RS-Net achieves optimal performance in both the reconstruction and segmentation tasks. For our datasets, the value of the segmentation loss is 100 times greater than that of the reconstruction loss. Upon setting β to 0.01, the value of the segmentation loss becomes comparable to that of the reconstruction loss, enabling the network to attain a better balance between the segmentation and reconstruction tasks. As such, we set β to 0.01 for all the other investigations in this work.

3 | RESULTS

3.1 | Single-Recon versus RS-Net

Figure 3 presents the quantitative comparisons between Single-Recon and RS-Net at various AFs. PSNR and SSIM values obtained by the two methods are shown in Figure 3a and Figure 3b, respectively. From the results, we can find that the proposed RS-Net has higher PSNR and SSIM values than Single-Recon at each AF. Particularly, at the highest AF of 6, RS-Net leads to the increase of PSNR by 7.79% and the increase of SSIM by 5.39%, compared with Single-Recon. Moreover, when the AF increases from 4 to 6, although it is anticipated that the reconstruction performance degrades, the decrease in PSNR and SSIM values of RS-Net is less than that of Single-Recon.

Figure 4 shows the representative reconstruction results of Single-Recon and RS-Net at AFs of 4–6. The reconstructed images of the two methods and corresponding error maps are displayed on the left and right of the figure, respectively. It can be seen that both Single-Recon and RS-Net can successfully remove undersampling artifacts at each AF. However, some detailed structures are blurred in the reconstructed images of Single-Recon, especially at high AF (indicated by red arrows). These structures are recovered clearly and the edge sharpness is well preserved in the results of the RS-Net. In summary, the proposed RS-Net yields better visual results with more details and fewer errors than Single-Recon.

3.2 | Comparison with the state-of-the-art reconstruction algorithms

Table 2 tabulates the quantitative results of all comparison reconstruction methods at different AFs for the healthy and patient test data. In the last five rows, the total average PSNR and SSIM values are listed. It can be observed that the proposed RS-Net has the highest PSNR and SSIM values among comparison methods for all clinical indications at each AF. In particular, compared with Deep-Cascade (the highest PSNR value of the other methods), RS-Net leads to the increase of PSNR by 2.83%, 3.73%, and 4.71% at AF of 4–6 for all test data, respectively.

Figures 5 and 6 show the representative results from healthy subjects and patients obtained by all comparison reconstruction methods at AF of 4 and 6, respectively. The representative results at an AF of 5 are

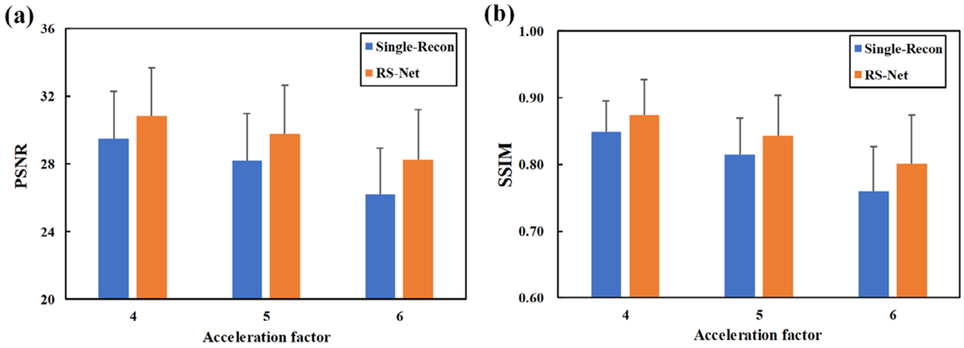

FIGURE 3 Quantitative comparisons between Single-Recon and RS-Net at various AFs. (a) The bar plot shows the average and standard deviation of PSNR values of images reconstructed by Single-Recon and RS-Net at various AFs. (b) The bar plot shows the average and standard deviation of SSIM values of images reconstructed by Single-Recon and RS-Net at various AFs.

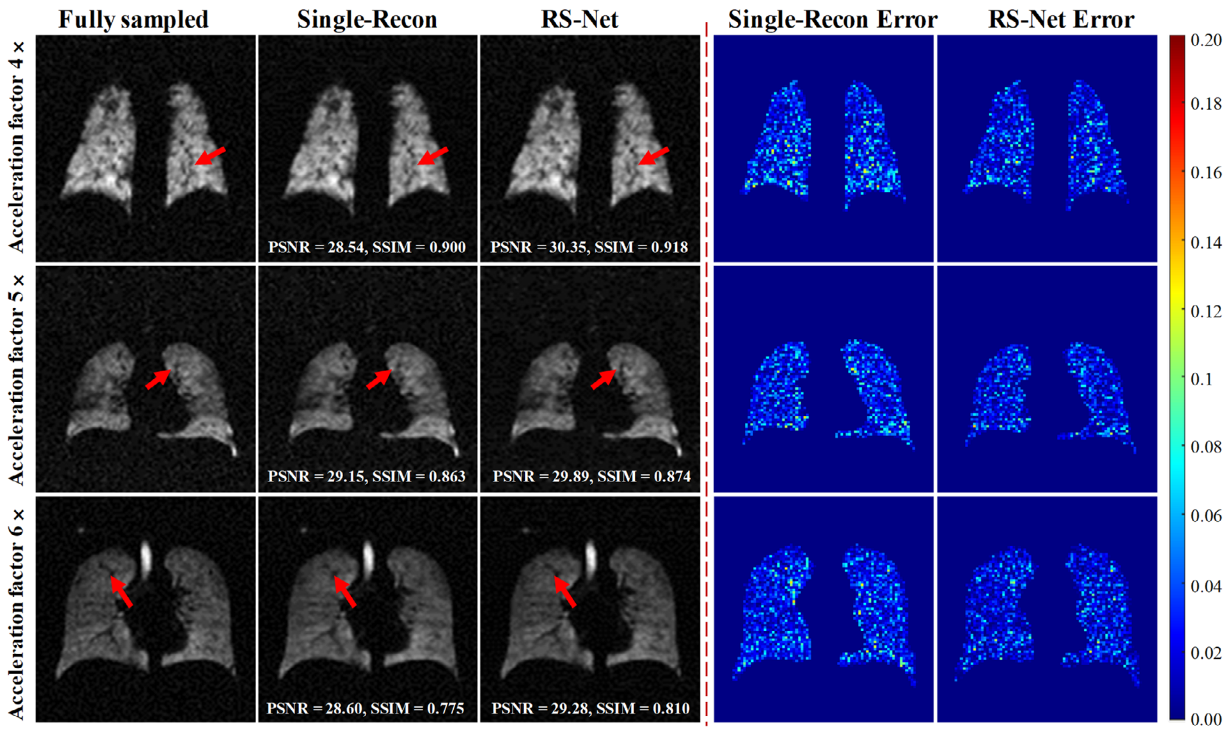

FIGURE 4 Visual comparison of Single-Recon and RS-Net at different AFs. The error maps show the differences in the lung regions between the reconstructed images and fully sampled images.

shown in Supporting Information Figure S-1. The top two rows of the figure are the results from a healthy subject and the bottom two rows of the figure are the results from a patient. Overall, RS-Net achieves more accurate reconstruction than other reconstruction methods both for healthy subjects and patients at all AFs. Specifically, we can observe that deep learning-based reconstruction methods exhibit better performance in removing undersampling artifacts than CS method. Nevertheless, the reconstructed images of CasNet are oversmoothed and miss lots of details in ventilation defect regions (indicated by red arrows), especially at a high AF of 6. These details become clearer in the results of KIKI-net and Deep-Cascade, but are still insufficient. However, the reconstructed images of RS-Net are almost the same as the fully sampled images at relatively low AFs (4 and 5), and preserve most of the fine details even at a high AF of 6. In addition, the reconstructed images of the RS-Net have the minimum pixel-value differences with fully sampled images among all comparison methods.

3.3 | Comparison of segmentation performance

Table 3 lists the Dice scores calculated between the ventilation defect regions obtained by all comparison

TABLE 2 Quantitative results of all comparison methods at different Afs.

Clinical indications	Model	AF = 4 PSNR	AF = 4 SSIM	AF = 5 PSNR	AF = 5 SSIM	AF = 6 PSNR	AF = 6 SSIM
Healthy	Zero filling	16.11 (2.14)	0.536 (0.050)	15.92 (2.20)	0.513 (0.053)	15.10 (2.14)	0.458 (0.054)
	CS	20.73 (1.28)	0.632 (0.077)	20.43 (1.16)	0.617 (0.077)	19.40 (1.58)	0.574 (0.078)
	CasNet	25.73 (1.82)	0.808 (0.042)	23.92 (1.77)	0.755 (0.040)	23.90 (2.06)	0.732 (0.050)
	KIKI-net	28.34 (1.89)	0.873 (0.035)	26.93 (1.87)	0.835 (0.042)	25.05 (1.98)	0.795 (0.051)
	Deep-Cascade	28.27 (1.92)	0.861 (0.033)	26.88 (2.12)	0.831 (0.040)	25.29 (2.00)	0.797 (0.051)
	RS-Net	**29.39 (1.72)**	**0.880 (0.034)**	**28.21 (2.09)**	**0.850 (0.038)**	**26.74 (2.07)**	**0.808 (0.052)**
Patient	Zero filling	17.32 (2.16)	0.534 (0.087)	17.18 (2.08)	0.514 (0.086)	16.26 (2.01)	0.462 (0.088)
	CS	22.69 (2.54)	0.624 (0.106)	22.35 (2.59)	0.608 (0.106)	21.15 (2.69)	0.566 (0.103)
	CasNet	27.32 (3.01)	0.783 (0.073)	25.58 (2.88)	0.734 (0.079)	25.45 (2.80)	0.719 (0.083)
	KIKI-net	30.54 (2.98)	0.866 (0.055)	29.15 (2.93)	0.827 (0.065)	27.48 (3.29)	0.780 (0.071)
	Deep-Cascade	30.80 (3.25)	0.870 (0.048)	29.50 (3.14)	0.840 (0.056)	27.69 (3.13)	0.783(0.069)
	RS-Net	**31.52 (2.82)**	**0.873 (0.059)**	**30.45 (2.89)**	**0.841 (0.067)**	**28.87 (3.09)**	**0.798 (0.079)**
Total	Zero filling	16.96 (2.22)	0.534 (0.078)	16.81 (2.19)	0.514 (0.078)	15.91 (2.11)	0.461 (0.080)
	CS	22.11 (2.41)	0.626 (0.098)	21.78 (2.43)	0.611 (0.098)	20.63 (2.54)	0.569 (0.096)
	CasNet	26.85 (2.81)	0.790 (0.064)	25.09 (2.69)	0.740 (0.069)	24.99 (2.77)	0.723 (0.079)
	KIKI-net	29.89 (2.88)	0.868 (0.050)	28.49 (2.84)	0.830 (0.059)	26.76 (3.15)	0.787 (0.066)
	Deep-Cascade	30.04 (3.13)	0.869 (0.044)	28.72 (3.11)	0.838 (0.052)	26.97 (3.04)	0.789 (0.064)
	RS-Net	**30.89 (2.72)**	**0.875 (0.053)**	**29.79 (2.87)**	**0.844 (0.061)**	**28.24 (2.99)**	**0.801 (0.072)**

[a]Results (mean (standard deviation)) were computed only over the lung region, and the best results are shown in bold.

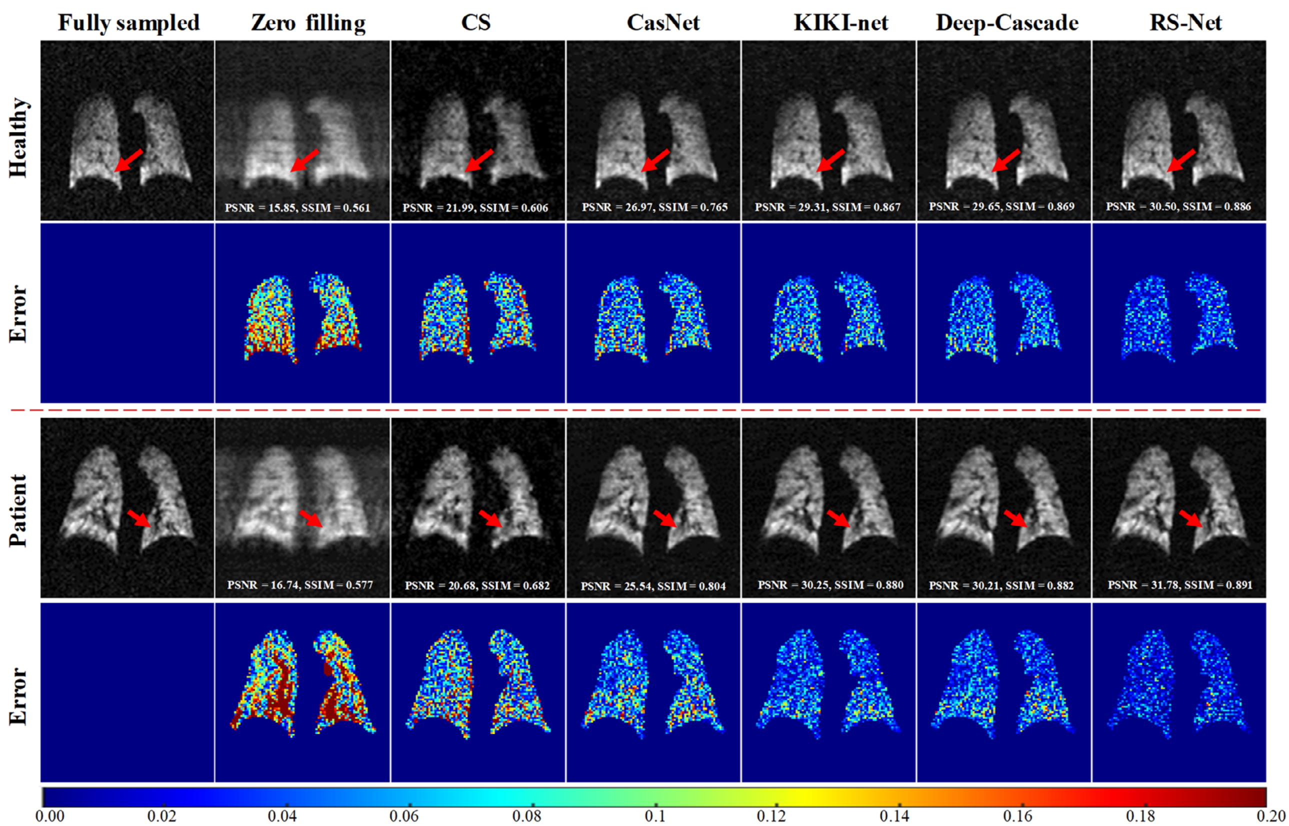

FIGURE 5 Reconstruction results from a healthy subject (top two rows) and a patient (bottom two rows) at an AF of 4. Corresponding error maps show the differences in the lung regions between the reconstructed images and fully sampled images.

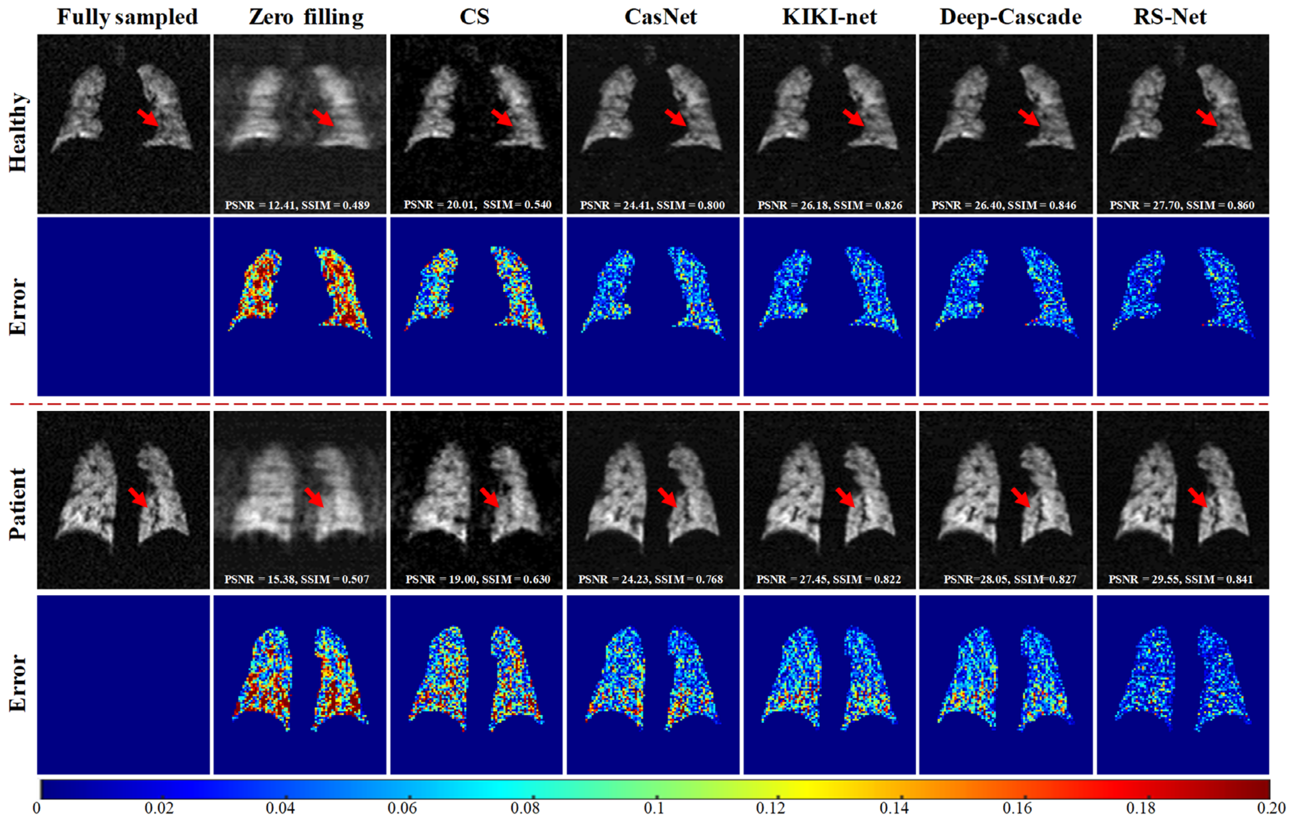

FIGURE 6 Reconstruction results from a healthy subject (top two rows) and a patient (bottom two rows) at an AF of 6. Corresponding error maps show the differences in the lung regions between the reconstructed images and fully sampled images.

TABLE 3 Dice scores (mean (standard deviation)) obtained by all comparison methods at various AFs.

Clinical indications	AF	Single-Seg	Zero filling-k	CS-k	CasNet-k	KIKI-net-k	Deep-Cascade-k	RS-Net
Healthy	4	0.802 (0.072)	0.232 (0.130)	0.616 (0.105)	0.739 (0.075)	0.814 (0.056)	0.820 (0.054)	**0.868 (0.034)**
	5	0.735 (0.087)	0.223 (0.137)	0.608 (0.102)	0.716 (0.074)	0.789 (0.056)	0.800 (0.056)	**0.833 (0.047)**
	6	0.702 (0.101)	0.196 (0.132)	0.572 (0.096)	0.700 (0.096)	0.750 (0.069)	0.764 (0.068)	**0.815 (0.057)**
Patient	4	0.862 (0.068)	0.416 (0.211)	0.736 (0.130)	0.822 (0.105)	0.880 (0.068)	0.883 (0.063)	**0.901 (0.052)**
	5	0.827 (0.090)	0.406 (0.216)	0.729 (0.132)	0.786 (0.150)	0.856 (0.085)	0.866 (0.072)	**0.876 (0.070)**
	6	0.801 (0.100)	0.377 (0.222)	0.709 (0.144)	0.771 (0.163)	0.832 (0.107)	0.834 (0.099)	**0.862 (0.070)**
Total	4	0.846 (0.074)	0.378 (0.220)	0.705 (0.137)	0.797 (0.108)	0.861 (0.071)	0.866 (0.067)	**0.892 (0.050)**
	5	0.801 (0.098)	0.372 (0.227)	0.696 (0.135)	0.765 (0.134)	0.837 (0.084)	0.847 (0.074)	**0.864 (0.067)**
	6	0.773 (0.109)	0.346 (0.226)	0.671 (0.146)	0.750 (0.149)	0.809 (0.104)	0.815 (0.096)	**0.849 (0.070)**

[a]The best results are shown in bold.

methods and fully sampled images at various AFs. At all AFs, RS-Net has the closest performance to the upper bound FS_U-Net (Dice score of 0.932), and consistently outperforms the other methods in terms of the mean and standard deviation of Dice scores for all clinical indications. Compared with the single segmentation network (Single-Seg), RS-Net leads to the increase of the Dice score by 5.44%, 7.87%, and 9.52% at the AF of 4, 5, and 6 for all test data, respectively. In comparison with the highest Dice score of the Reconstruction-k methods (Deep-Cascade-k), RS-Net leads to the increase of Dice score by 3.00%, 2.01%, and 4.17% at AF of 4, 5, and 6 for all test data, respectively.

Figures 7 and 8 display the ventilation defect region segmentation results from healthy subjects and patients of all comparison methods at different AFs, respectively. From Figure 7, it can be found that Single-Seg, Zero filling-k, CS-k, CasNet-k, KIKI-net-k, and Deep-Cascade-k fail to restore the small ventilation defect regions of the healthy subjects (indicated by the blue arrows), while

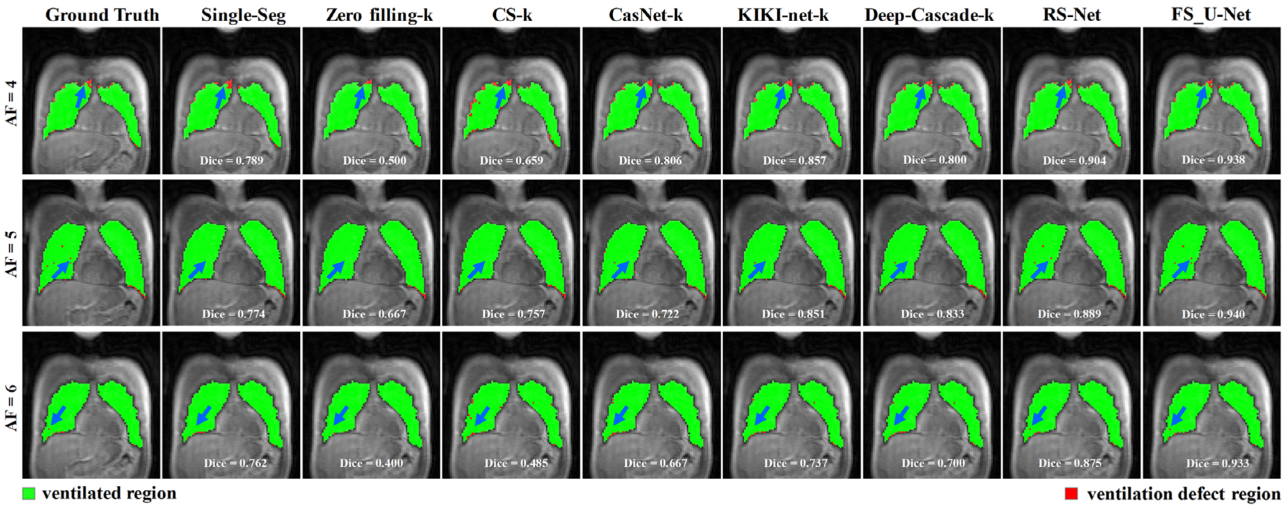

FIGURE 7 Representative segmentation results from healthy subjects at different AFs. The green color denotes the ventilated region and the red color denotes the ventilation defect region.

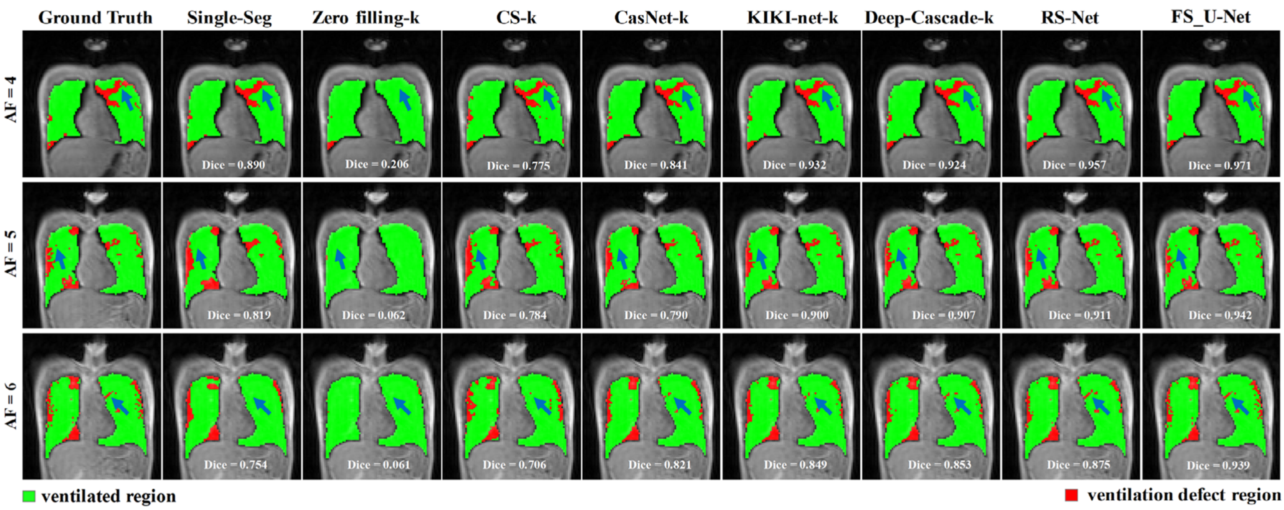

FIGURE 8 Representative segmentation results from patients at different AFs. The green color denotes the ventilated region and the red color denotes the ventilation defect region.

RS-Net can predict these regions with high fidelity at all AFs. Similarly, from Figure 8, we can see that RS-Net exhibits more satisfactory segmentation results of patients than the other comparison methods at each AF. Specifically, the segmentation results of Zero filling-k are entirely inaccurate. Single-Seg, CS-k and CasNet-k only obtain the coarse segmentation results, which cannot predict the fine structures (indicated by the blue arrows), especially at a high AF of 6. The segmentation results of KIKI-net-k, and Deep-Cascade-k become finer to some extent but are still limited. In the segmentation results of RS-Net, these fine structures are accurately predicted. Additionally, the segmentation results from the RS-Net and FS_U-Net trained on fully sampled images are comparable qualitatively even at high AFs, which are very close to the ground truth.

3.4 | Comparison of VDP performance

Figure 9 and Supporting Information Figures S-2and S-3 depict the scatter plots of VDP calculated from the fully sampled images and various methods for all test data at AF of 4, 5, and 6, respectively. Pearson correlation coefficients (r values) are calculated to evaluate the correlation between VDP values obtained from fully sampled images and various methods, which are listed in the scatter plots.

It can be observed that the r values of RS-Net are the highest among all comparison methods at each AF. In comparison with the Single-Seg, RS-Net leads to the increase of r values by 2.82%, 6.22%, and 9.16% at AF of 4, 5, and 6, respectively. Additionally, RS-Net leads to the increase of r values by 2.18%, 1.89%, and 2.62% at

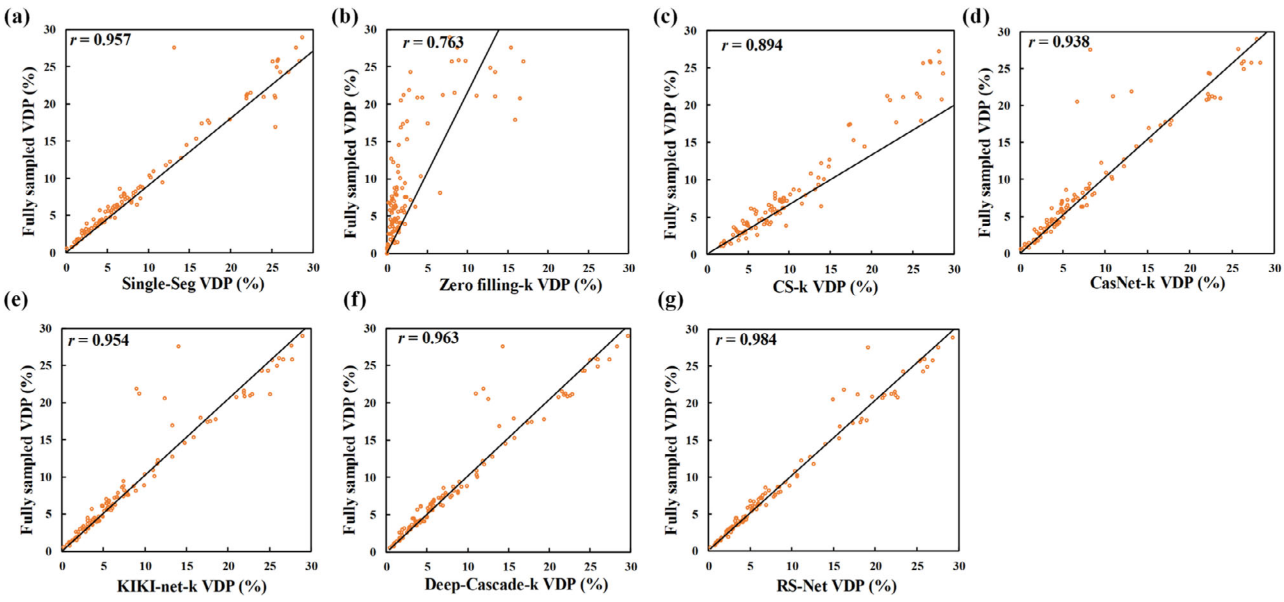

FIGURE 9 VDP comparisons of fully sampled images and different methods at an AF of 4. (a) Scatter plots of VDP calculated from Single-Seg and fully sampled images. (b) Scatter plots of VDP calculated from Zero filling-k and fully sampled images. (c) Scatter plots of VDP calculated from CS-k and fully sampled images. (d) Scatter plots of VDP calculated from CasNet-k and fully sampled images. (e) Scatter plots of VDP calculated from KIKI-net-k and fully sampled images. (f) Scatter plots of VDP calculated from Deep-Cascade-k and fully sampled images. (g) Scatter plots of VDP calculated from RS-Net and fully sampled images. Pearson correlation coefficients (*r* values) are listed in the scatter plots.

AF of 4, 5, and 6, respectively, compared with the Deep-Cascade-k (the highest *r* values of the Reconstruction-k methods). Moreover, there are no significant differences between the VDP values obtained by RS-Net and the VDP values obtained by fully sampled images ($p > 0.05$).

3.5 | Comparison with uncoupled models

To investigate the strength of the proposed multi-task architecture, we compared it with two uncoupled models that perform reconstruction and segmentation in a sequential way. In the first model (called CS_U-Net), reconstructed images are obtained using the CS method and then reconstructed images are utilized to train a U-Net for the segmentation of ventilation defects. Subsequently, segmentation results are employed to guide the optimization of CS, including weights for total variation norm and L1 norm, for better reconstruction. The second model (called Cascade U-Net) adopts a cascaded U-Net architecture, where the first U-Net is used for reconstruction and the second U-Net is used for segmentation on the reconstructed images from the first U-Net. Cascade_U-Net is trained in an end-to-end manner.

Table 4 and Figure 10 present the performance comparison of different multi-task architectures at an AF of 4. The results show that RS-Net achieves the best reconstruction and segmentation performance

TABLE 4 Quantitative metrics obtained by different multi-task architectures at an AF of 4.

Metrics	CS_U-Net	Cascade_U-Net	RS-Net
PSNR	22.11 (2.41)	29.38 (2.92)	**30.89 (2.72)**
SSIM	0.626 (0.098)	0.857 (0.054)	**0.875 (0.053)**
Dice	0.791 (0.075)	0.879 (0.056)	**0.892 (0.050)**

[a]Best results are shown in bold.

both in qualitative visualization and quantitative validation among all compared models, demonstrating the superiority of the proposed multi-task architecture.

3.6 | Ablation study

To examine the effects of the weight sharing scheme and feature-selecting blocks in RS-Net, we investigated ablation studies on RS-Net without weight sharing scheme and RS-Net without feature-selecting blocks. We summarize the reconstruction and segmentation results at an AF of 4 in Table 5. It can be seen that both the reconstruction and segmentation performance drop dramatically without weight sharing scheme, and are even inferior to that of the Single-Recon (PSNR of 29.51 and SSIM of 0.849) and Single-Seg (Dice score of 0.846) models. Moreover, RS-Net has higher PSNR, SSIM, and Dice score than that of RS-Net without feature-selecting blocks.

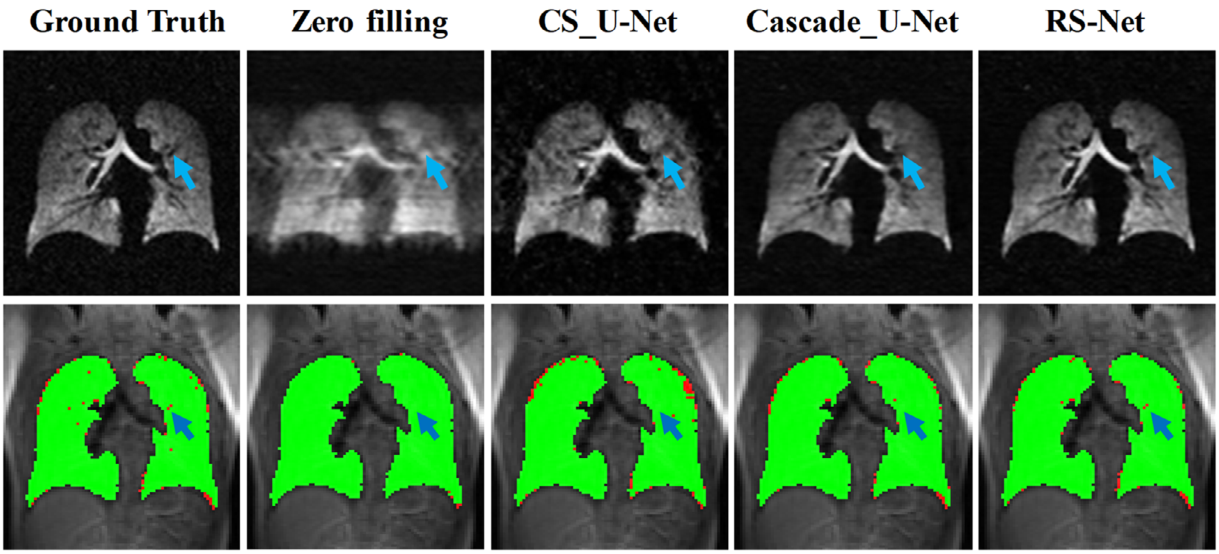

FIGURE 10 Representative results obtained by different multi-task architectures at an AF of 4.

TABLE 5 Ablation studies on the weight sharing strategy and feature-selecting block at an AF of 4.

Metrics	Without weight sharing	Without feature-selecting block	RS-Net
PSNR	29.35 (3.12)	30.55 (2.86)	**30.89 (2.72)**
SSIM	0.848 (0.051)	0.869 (0.087)	**0.875 (0.053)**
Dice	0.839 (0.111)	0.879 (0.057)	**0.892 (0.050)**

[a]Best results are shown in bold.

4 DISCUSSION

In this work, we have presented a complementation-reinforced network for simultaneous reconstruction and segmentation of HP gas MRI, called RS-Net. RS-Net employs a dual-branch encoder-decoder architecture that integrates the two tasks into a unified framework. Weight sharing scheme is applied between the two encoders to facilitate effective knowledge transfer between the tasks. The feature-selecting block is introduced to pick the most informative features for each task, further enhancing the performance of the RS-Net. Moreover, the use of the lung mask obtained from the reconstructed image as complementary information in the segmentation branch contributes to the improved segmentation results. Experimental results demonstrate the superiority of the RS-Net in the reconstruction and segmentation tasks.

A multi-task model can learn different tasks in sequential or parallel.[42] Our results (Table 4 and Figure 10) show that the proposed parallel model RS-Net performs better than simple sequential models (CS_U-Net and Cascade_U-Net). Particularly, CS_U-Net and Cascade_U-Net cannot predict the ventilation defects that are lost in the reconstructed images (see blue arrows in Figure 10). This limitation arises because they perform reconstruction before segmentation, which constrains the segmentation network to learn the features inherent in the reconstructed images. In contrast, RS-Net learns the two tasks in parallel, which enables both tasks to extract relevant features from undersampled data according to their respective needs, thereby achieving improved results.

Weight sharing and feature selection allow the multi-task model to leverage the underlying similarities and differences between the tasks, thereby facilitating performance improvement. This point has been confirmed by ablation studies. Specifically, the results (Table 5) indicate that joint optimization of two tasks without sharing underlying features is inadequate for fostering mutual regularization and synergies between tasks and even leads to performance degradation due to the inflexible nature of joint optimization. Moreover, the efficacy of feature-selecting blocks in the multi-task network has been evident from the observed performance improvements.

RS-Net outperforms its variant networks for the single task (Single-Recon and Single-Seg), indicating the presence of a synergistic effect between the two tasks in RS-Net. Compared to Single-Recon, RS-Net is more robust to high AFs and can recover more details, especially in low ventilation regions (see Figures 3 and 4). This is attributed to the fact that RS-Net focuses more on high-level semantic features of low ventilation regions under the constraint of segmentation loss,[40–42] resulting in better preservation of fine structures of low ventilation regions. While these structures will be discarded by Single-Recon because they have little impact on overall image quality (indicated by red arrows in Figure 4).[19] Moreover, RS-Net yields more precise segmentation results than Single-Seg (see Table 3,

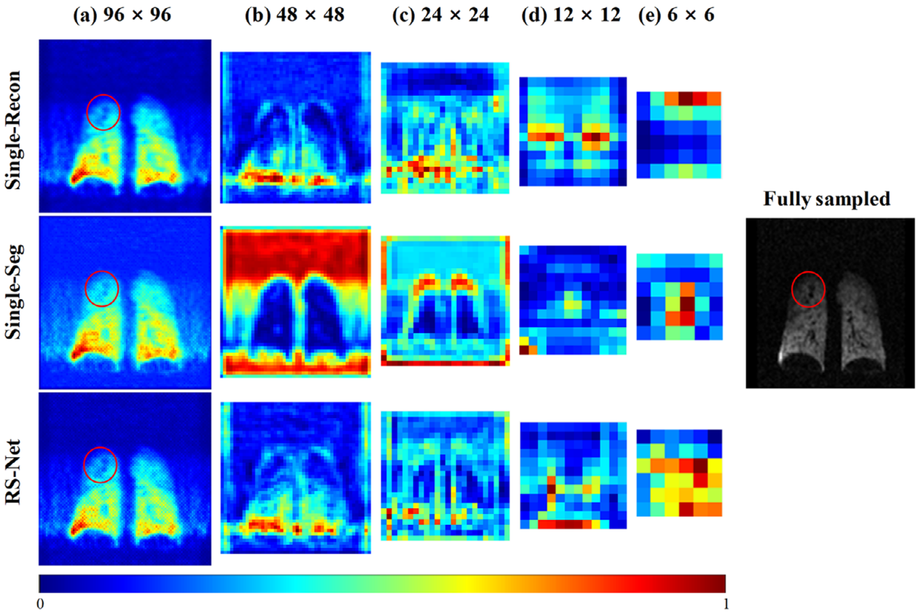

FIGURE 11 Visualization of the features extracted by the encoders in different networks at an AF of 4. (a)–(e) show the features generated by the first to fifth contracting steps of the encoders, respectively.

Figures 7 and 8). Zero-filling images are susceptible to severe undersampling artifacts, so Single-Seg is prone to the influence of artifacts.[28] As a result, some features extracted by Single-Seg are not useful for segmentation and could even result in wrong results, especially at high AFs (indicated by blue arrows in Figures 7 and 8). However, RS-Net can alleviate the influence of undersampling artifacts under the constraint of reconstruction task. Figure 11 further confirms these points. In the first stage of feature extraction, RS-Net achieves the most effective feature expression of low ventilation regions. Conversely, Single-Recon is inadequate in extracting features of these regions, and features extracted by Single-Seg are still vulnerable to undersampling artifacts, leading to information loss.

RS-Net also exhibits superior reconstruction performance compared to state-of-the-art reconstruction methods, including CS,[9] CasNet,[17] KIKI-net,[36] and Deep-Cascade,[37] as evidenced by higher PSNR and SSIM (Table 2). This further validates that effectively capitalizing on complementary information from segmentation and reconstruction tasks can enhance reconstruction performance. Moreover, RS-Net can restore sharp lung structures even at a high AF of 6 (Figure 6), while existing HP gas MRI reconstruction networks can only achieve satisfactory results at an AF of 4, with extremely blurred images at higher AFs.[17,18] This highlights the role of the multi-task learning scheme in achieving higher AF. Notably, Table S-1 indicates that for all compared methods, patients exhibit significantly higher PSNR than healthy subjects ($p < 0.01$). This is because patients have larger ventilation defect regions, leading to more low-valued pixels in their lung region. PSNR measures image similarity based on absolute pixel differences.[38] For our dataset, due to the lower pixel values of patients, they have smaller absolute pixel differences between reconstructed and fully sampled images compared to healthy subjects, resulting in higher PSNR values. However, the visual reconstruction quality between the two groups is consistent (Figures 5, 6, and S-1). This conclusion is reinforced by the results of SSIM, which show no significant difference between the two groups ($p > 0.05$) because SSIM evaluates the perceptual similarity of two images based on structural information rather than absolute pixel differences.[39]

The results of commonly used reconstruction-based segmentation methods[17] (Figures 7 and 8) confirm that the performance of segmentation is highly dependent on the quality of reconstruction. Zero filling images are too blurry to identify ventilation defect regions, and higher quality reconstructed images yield more accurate segmentation results. Moreover, reconstruction-based segmentation methods are unable to predict regions that have been lost during the reconstruction process. In contrast, RS-Net directly obtains segmentation results from the undersampled k-space data, which can exploit more original information and learn more segmentation-driven features. This effectively boosts the accuracy of segmentation. It should be noted that Table S-1 shows that healthy subjects have significantly lower Dice scores compared to patients ($p < 0.01$), mainly due to the smaller size of ventilation defects in the healthy subjects.[21] However, the qualitative segmentation quality of both groups is consistent, as shown in Figures 7 and 8.

We further demonstrated the clinical potential of the RS-Net by quantifying VDP values obtained from various segmentation methods. As shown in Figure 9, Figures S-2 and S-3, it can be found that there are obvious increases in Pearson correlation coefficients (*r* values) of RS-Net, compared with the other methods. Moreover, the *p*-values of RS-Net are higher than 0.05. These results mean that VDP values of RS-Net have better correlations than the comparison methods and no significant differences with VDP values of the fully sampled images. This demonstrates that RS-Net has the potential for real-time assessment of lung function.

In the present study, we mainly evaluated RS-Net on the single-coil MRI data undersampled by the Cartesian trajectory. However, RS-Net is flexible and theoretical, which is also applicable to the multi-coil and non-Cartesian undersampled MRI data, such as radial and spiral sampling patterns. In the future, we will test the performance of RS-Net on various types of HP gas MRI data.

5 CONCLUSIONS

In this work, we develop a complementation-reinforced network (RS-Net) using multi-task learning for fast and accurate reconstruction and segmentation of pulmonary gas MRI. In the RS-Net, reconstruction and segmentation tasks regularize each other by sharing underlying characteristics, which enhances the generalization capability of the network. Experimental results demonstrate that RS-Net can create more accurate reconstruction and segmentation results than state-of-the-art methods. Compared with single-task models, RS-Net leads to the increase of PSNR, SSIM by 7.79% and 5.39%, respectively, at a high AF of 6. In addition, RS-Net achieves segmentation of ventilation defect regions directly from the highly undersampled k-space data, effectively improving the segmentation performance at high AFs (Dice score increased by 9.52%). Furthermore, the VDP values obtained by the RS-Net agree well with those of the fully sampled images. Consequently, the proposed method provides a new perspective for future research where multi-task learning can be exploited for real-time and accurate imaging and segmentation of HP gas MRI, which benefits the diagnosis and prognosis of lung diseases.

ACKNOWLEDGMENTS

This work was supported by National key Research and Development Project of China (2018YFA0704000), National Natural Science Foundation of China (82127802, 21921004, 82001915), Key Research Program of Frontier Sciences (ZDBS-Y-JSC004), Hubei Provincial Key Technology Foundation of China (2021ACA013), China Postdoctoral Science Foundation (2020M672454, 2020T130125ZX, 2022M713218). Xin Zhou acknowledges the support from the Tencent Foundation through the XPLORER PRIZE.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available within the article and its Supporting Information.

REFERENCES

1. Aviles-Rivero AI, Debroux N, Williams G, Graves MJ, Schönlieb CB. Compressed sensing plus motion (CS + M): a new perspective for improving undersampled MR image reconstruction. *Med Image Anal*. 2021;68:101933.
2. Huang Q, Xian Y, Yang D, et al. Dynamic MRI reconstruction with end-to-end motion guided network. *Med Image Anal*. 2021;68:101901.
3. Mugler JP III, Altes TA. Hyperpolarized ^{129}Xe MRI of the human lung. *J Magn Reson Imaging*. 2013;37:313-331.
4. Zhou X, Graziani D, Pines A. Hyperpolarized xenon NMR and MRI signal amplification by gas extraction. *Proc Natl Acad Sci USA*. 2009;106:16903-16906.
5. Albert MS, Cates GD, Driehuys B, et al. Biological magnetic resonance imaging using laser-polarized ^{129}Xe. *Nature*. 1994;370:199-201.
6. Li H, Zhao X, Wang Y, et al. Damaged lung gas exchange function of discharged COVID-19 patients detected by hyperpolarized ^{129}Xe MRI. *Sci Adv*. 2021;7:eabc8180.
7. Xiao S, Deng H, Duan C, et al. Considering low-rank, sparse and gas-inflow effects constraints for accelerated pulmonary dynamic hyperpolarized ^{129}Xe MRI. *J Magn Reson*. 2018;290:29-37.
8. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. *Magn Reson Med*. 2009;61:103-116.
9. Ajraoui S, Lee KJ, Deppe MH, Parnell SR, Parra-Robles J, Wild JM. Compressed sensing in hyperpolarized ^{3}He Lung MRI. *Magn Reson Med*. 2010;63:1059-1069.
10. Ajraoui S, Parra-Robles J, Wild JM. Incorporation of prior knowledge in compressed sensing for faster acquisition of hyperpolarized gas images. *Magn Reson Med*. 2013;69:360-369.
11. Xiao S, Deng H, Duan C, et al. Highly and adaptively undersampling pattern for pulmonary hyperpolarized ^{129}Xe dynamic MRI. *IEEE Trans Med Imaging*. 2019;38:1240-1250.
12. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med*. 2018;79:3055-3071.
13. Ravishankar S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imaging*. 2010;30:1028-1041.
14. Yang G, Yu S, Dong H, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans Med Imaging*. 2018;37:1310-1321.
15. Aggarwal HK, Mani MP, Jacob M. MoDL: model-based deep learning architecture for inverse problems. *IEEE Trans Med Imaging*. 2018;38:394-405.
16. Wang S, Xiao T, Liu Q, Zheng H. Deep learning for fast MR imaging: A review for learning reconstruction from incomplete k-space data. *Biomed Signal Process Control*. 2021;68:102579.
17. Duan C, Deng H, Xiao S, et al. Fast and accurate reconstruction of human lung gas MRI with deep learning. *Magn Reson Med*. 2019;82:2273-2285.

18. Duan C, Deng H, Xiao S, et al. Accelerate gas diffusion-weighted MRI for lung morphometry with deep learning. *Eur Radiol*. 2022;32:702-713.
19. Huang Q, Chen X, Metaxas D, Nadar MS. Brain Segmentation from k-space with end-to-end recurrent attention network. In *Proceedings of international conference on medical image computing and computer-assisted intervention*, Shenzhen, China, 2019. pp. 275–283.
20. Ebner L, Virgincar RS, He M, et al. Multireader determination of clinically significant obstruction using hyperpolarized ^{129}Xe ventilation MRI. *Am J Roentgenology*. 2019;212:758-165.
21. Kirby M, Heydarian M, Svenningsen S, et al. Hyperpolarized ^{3}He magnetic resonance functional imaging semiautomated segmentation. *Acad Radiol*. 2012;19:141-152.
22. Hughes PJC, Horn FC, Collier GJ, et al. Spatial fuzzy c-means thresholding for semiautomated calculation of percentage lung ventilated volume from hyperpolarized gas and ^{1}H MRI. *J Magn Reson Imaging*. 2018;47:640-646.
23. Astley JR, Biancardi AM, Hughes PJC, et al. 3D deep convolutional neural network-based ventilated lung segmentation using multi-nuclear hyperpolarized gas MRI. In: *International workshop on Thoracic Image Analysis*. 2020:24-35.
24. Caballero J, Bai W, Price AN, Rueckert D, Hajnal JV. Application driven MRI: joint reconstruction and segmentation from undersampled MRI data. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, Boston, MA, 2014. pp. 106–113.
25. Sun L, Fan Z, Ding X, Huang Y, Paisley J. Joint CS-MRI reconstruction and segmentation with a unified deep network. In *Proceedings of International Conference on Information Processing in Medical Imaging*, Hong Kong, China, 2019. pp. 492–504.
26. Huang Q, Yang D, Yi J, Axel L, Metaxas D. FR-Net: Joint reconstruction and Segmentation in compressed sensing cardiac MRI. In *Proceedings of International Conference on Functional Imaging and Modeling of the Heart*, Bordeaux, France, 2019. pp. 352–360.
27. Sui B, Lv J, Tong X, Li Y, Wang C. Simultaneous image reconstruction and lesion segmentation in accelerated MRI using multitasking learning. *Med Phys*. 2021;48:7189-7198.
28. Schlemper J, Oktay O, Bai W, et al. Cardiac MR segmentation from undersampled k-space using deep latent representation learning. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 2018. pp. 259–267.
29. Jaubert O, Montalt-Tordera J, Knight D, et al. Real-time deep artifact suppression using recurrent U-nets for low-latency cardiac MRI. *Magn Reson Med*. 2021;86:1904-1916.
30. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 2015. pp. 234–241.
31. Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process*. 2017;26:4509-4522.
32. Zhou W, Du H, Mei W, Fang L. Efficient structurally-strengthened generative adversarial network for MRI reconstruction. *Neurocomputing*. 2021;422:51-61.
33. Bertels J, Robben D, Vandermeulen D, Suetens P. Theoretical analysis and experimental validation of volume bias of soft Dice optimized segmentation maps in the context of inherent uncertainty. *Med Image Anal*. 2021;67:101833.
34. He M, Zha W, Tan F, Rankine L, Fain S, and Driehuys B. A comparison of two hyperpolarized ^{129}Xe MRI ventilation quantification pipelines: The effect of signal to noise ratio. *Acad. Radiol.*, 2019;26:949-959.
35. Rebouças Filho PP, Cortez PC, da Silva Barros AC, Victor VH, Tavares RSJM. Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images. *Med Image Anal*. 2017;35:503-516.
36. Eo T, Jun Y, Kim T, Jang J, Lee HJ, Hwang D. KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. *Magn Reson Med*. 2018;80:2188-2201.
37. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans Med Imaging*. 2018;37:491-503.
38. Qu X, Zhang W, Guo D, Cai C, Cai S, Chen Z. Iterative thresholding compressed sensing MRI based on contourlet transform. *Inverse Probl Sci Eng*. 2010;18:737-758.
39. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.
40. Tomczak A, Ilic S, Marquardt G, et al. Multi-task multi-domain learning for digital staining and classification of leukocytes. *IEEE Trans Med Imaging*. 2021;40:2897-2910.
41. Zhang Q, Yang G, Zhang G. Collaborative network for super-resolution and semantic segmentation of remote sensing images. *IEEE Trans Geosci Remote Sens*. 2022;60:4404512.
42. Corona V, Aviles-Rivero A, Debroux N, Guyader CL, Schönlieb CB. Variational multi-task MRI reconstruction: joint reconstruction, registration and super-resolution. *Med Image Anal*. 2021;68:101941.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Li Z, Xiao S, Wang C, et al. Complementation-reinforced network for integrated reconstruction and segmentation of pulmonary gas MRI with high acceleration. *Med Phys*. 2024;51:378–393.
https://doi.org/10.1002/mp.16591