Dynamic MRI Reconstruction with Motion-Guided Network

Qiaoying Huang QH55@CS.RUTGERS.EDU
Dong Yang DON.YANG.MECH@GMAIL.COM
Hui Qu HUI.QU@CS.RUTGERS.EDU
Jingru Yi JY486@CS.RUTGERS.EDU
Pengxiang Wu PW241@CS.RUTGERS.EDU
Dimitris Metaxas DNM@CS.RUTGERS.EDU
Department of Computer Science, Rutgers University, Piscataway, NJ 08854, USA

Abstract

Temporal correlation in dynamic magnetic resonance imaging (MRI), such as cardiac MRI, is informative and important to understand motion mechanisms of body regions. Modeling such information into the MRI reconstruction process produces temporally coherent image sequence and reduces imaging artifacts and blurring. However, existing deep learning based approaches neglect motion information during the reconstruction procedure, while traditional motion-guided methods are hindered by heuristic parameter tuning and long inference time. We propose a novel dynamic MRI reconstruction approach called MODRN that unitizes deep neural networks with motion information to improve reconstruction quality. The central idea is to decompose the motion-guided optimization problem of dynamic MRI reconstruction into three components: dynamic reconstruction, motion estimation and motion compensation. Extensive experiments have demonstrated the effectiveness of our proposed approach compared to other state-of-the-art approaches.

1. Introduction

Dynamic magnetic resonance imaging (MRI) is critical in clinical applications, such as cardiovascular and pulmonary. However, high spatiotemporal resolution reconstruction from under-sampled MRI k-space data is still very challenging due to the lack of gold-standard to clinical practice and strong dependence on the parameters tuning (Yang et al., 2016). Many works have been proposed to tackle these challenges. Some of them argued that motion plays a leading role in dynamic reconstruction because correlation and redundancy exist along the temporal dimension, such as cardiac moves relatively periodically against the static background (Jung et al., 2010; Jung and Ye, 2010; Chen et al., 2018). In other computer vision problems, such as video super resolution, motion estimation and compensation are also powerful techniques to further explore the temporal correlations between video frames (Caballero et al., 2017; Makansi et al., 2017).

Traditional Compressed Sensing (CS) approaches have dominated dynamic reconstruction in the past few years. Some studies have successfully incorporated the physical motion into the CS schemes to improve reconstruction performance, either by refining the results after the image reconstruction step (Jung et al., 2009; Bai and Brady, 2011) or embedding the motion estimation into the reconstruction process (Gigengack et al., 2012; Chen et al., 2018). For both methods, displacement motion fields or optical flow are calculated to estimate motion between image pairs. According to the form of motion constraint, many algorithms were proposed, among which, Horn-Schunck (Horn and Schunck, 1981) and Lucas-Kanade (Lucas and Kanade, 1981) are widely used.

However, heuristic parameter tuning and long reconstruction time are major drawbacks of these methods. Additionally, motion estimation itself is still a very difficult task and often produces inaccurate results if object movement is large or fast.

Recent advances in deep learning technique have sparked the new research interests in MRI reconstruction. Deep convolutional neural network was proposed to learn mapping directly from k-space data to fully-sampled reconstructed image, which introduced an interesting way for MRI reconstruction (Zhu et al., 2018). Compared to CS-based methods, deep learning approaches are faster in inference and learn the implicit prior automatically based on the training data (Sun et al., 2016; Schlemper et al., 2017, 2018; Lønning et al., 2018; Qin et al., 2018; Huang et al., 2018). DC-CNN was introduced in (Schlemper et al., 2017), where a differentiable data consistency (DC) layer was added into a deep cascaded convolution neural network for both 2D or dynamic reconstruction. Qin *et al*. (Qin et al., 2018) efficiently modeled the recurrence of the iterative reconstruction stages by a recurrent network. However, most of them were derived for 2D reconstruction problem. The deep learning based dynamic MRI reconstruction problem is largely unsolved yet critical in clinical scenarios. Especially, all existing state-of-the-art methods did not take motion information into consideration that may generate blurry or temporal inconsistent results.

To tackle the aforementioned problems of both traditional and deep learning methods, in this study, we develop a Motion-guided Dynamic Reconstruction Network (MODRN) that utilizes motion estimation and motion compensation (ME/MC) to improve the reconstruction quality for spatiotemporal imaging. Different from traditional motion estimation algorithms which may fail in low resolution and weak contrast, we utilize the unsupervised deep learning based optical flow estimation (Ren et al., 2017; Meister et al., 2017), which is more robust and accurate in different scenarios. To the best of our knowledge, this is the first work that embeds motion information into deep neural network for dynamic MRI reconstruction. The contribution of this work are three folds. Firstly, we derive a recurrent neural network from the optimization procedures of model-based dynamic reconstruction, which simultaneously links the relationship of data over time and iterations. Secondly, we introduce an unsupervised deep learning based motion estimation method to learn the motion between the reference image and the reconstructed image by using the combination of forward, backward and neighboring loss. Finally, we present a motion compensation component for refining reconstructed image guided by the learned motion.

2. Methodology

In this section, we start with the formulation of dynamic MRI reconstruction problem, followed by detailed description of our proposed method called Motion-guided Dynamic Reconstruction Network (MODRN).

2.1. Problem Formulation

Given a sequence of under-sampled k-space data $\{y_t\}_{t\in[T]}$ of T frames, the dynamic MRI reconstruction problem is to predict reconstructed images $\{z_t\}_{t\in[T]}$ from $\{y_t\}$, which can be formalized as an optimization problem: $\text{argmin}_{\{z_t\}}\ \mathscr{L}_{rec}(\{z_t\})$, where

$$\mathscr{L}_{rec}(\{z_t\}) = \sum_{t=1}^{T} \frac{1}{2}\|F_u(z_t) - y_t\|_2^2 + \lambda R(z_t). \tag{1}$$

The term $\|F_u(z_t) - y_t\|_2^2$ is used to guarantee data consistency by restricting the reconstructed image z_t to be close to the input measurement y_t. $F_u(\cdot)$ is an operator that transforms image-domain z_t into Fourier domain followed by undersampling. $R(\cdot)$ is a regularization function that depends on the prior knowledge of the input $\{y_t\}$. Common choices include sparsity in transformed domain (Lingala et al., 2011), total variation (TV) penalties (Knoll et al., 2012) and low-rank property (Trzasko et al., 2011). λ is a weighting factor.

In order to capture anatomical motion in the dynamic MRI acquisition, it is natural to incorporate motion estimation/motion compensation (ME/MC) technique in the reconstruction process (Jung et al., 2010). Specifically, based on the brightness constancy assumption (Horn and Schunck, 1981), for a temporal 2D image $z_t(x,y,t)$ with small movement $(\Delta x, \Delta y, \Delta t)$ with respect to the next frame, we add the following motion estimation constraint to the objective function (1):

$$\mathscr{L}_{me}(\{v_t\}) = \sum_{t=1}^{T-1} \left\| \bigtriangledown z_t^\top v_t + \frac{\partial z_t}{\partial t} \right\|_1 + \delta \|v_t\|_1, \tag{2}$$

where $\bigtriangledown z_t(x,y) = \left(\frac{\partial z_t}{\partial x}, \frac{\partial z_t}{\partial y}\right)$ are the derivatives of image z_t at position (x,y), and $v_t(x,y) = \left(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}\right)$ is the estimated displacement motion fields or optical flow.

Furthermore, given the estimated motion field v_t, the reconstructed image z_t can be refined through MC process, *i.e.* $c_t = \mathrm{MC}(z_t, z_1, z_T) + r_t$, where c_t is the motion-compensated reconstructed image and r_t is a residual term for better exploiting temporal redundancy. Therefore, we can derive the motion compensation constraint as follows.

$$\mathscr{L}_{mc}(\{r_t\}) = \sum_{t=1}^{T-1} \frac{1}{2} \|F_u(c_t) - y_t\|_2^2. \tag{3}$$

By combining with two motion-based constraints of Equations (2) and (3), the motion-guided dynamic MRI reconstruction problem is defined as:

$$\underset{\{z_t, v_t, r_t\}}{\operatorname{argmin}} \ \mathscr{L}_{rec}(\{z_t\}) + \eta \mathscr{L}_{me}(\{v_t\}) + \zeta \mathscr{L}_{mc}(\{r_t\}). \tag{4}$$

The solution to problem (4) is non-trivial and traditional CS-based algorithms are usually computationally expensive and require long running time for hyper-parameter tuning. Recent advances in deep learning provide an alternative way for efficient MRI reconstruction, but very few works focused on the dynamic reconstruction problem and they only targeted for the simpler problem (1) without considering motion information. To this end, we propose a deep learning based method called Motion-guided Dynamic Reconstruction Network (MODRN) to solve the problem (4).

2.2. Motion-guided Dynamic Reconstruction Network

Our method dissects the motion-guided dynamic reconstruction problem into three closely-connected parts: (i) Dynamic Reconstruction (DR) component for estimating initial reconstructed image from Equation (1), (ii) Motion Estimation (ME) component for generating motion information through Equation (2), and (iii) Motion Compensation (MC) component for refining reconstructed image guided by learned motion based on Equation (3).

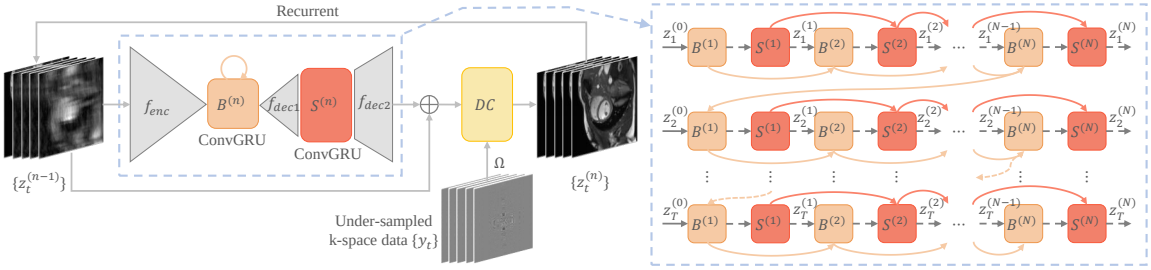

Figure 1: Architecture of DR component and the workflow of ConvGRU units B and S.

2.2.1. DYNAMIC RECONSTRUCTION

Instead of directly solving Equation (1), we adopt an iterative process through DR component to approximate reconstructed images z_t. Formally, given under-sampled k-space data $\{y_t\}_{t\in[T]}$ with sampled mask Ω, DR component learns to reconstruct images in N iterations:

$$z_t^{(n)} = DR\left(z_t^{(n-1)}, y_t, \Omega\right) \Leftrightarrow \begin{cases} x_{bt}^{(n)}, b_t^{(n)} = B\left(f_{enc}(z_t^{(n-1)}), b_t^{(n-1)}\right) \\ x_{st}^{(n)}, s_t^{(n)} = S\left(f_{dec1}(x_{bt}^{(n)}), s_t^{(n-1)}\right) \\ z_t^{(n)} = DC\left(f_{dec2}(x_{st}^{(n)}) + z_t^{(n-1)}, y_t, \Omega\right) \end{cases}, n \in [N]. \tag{5}$$

where $z_t^{(0)}$ is zero-filling image and $z_t^{(n)}$ is the reconstructed image of y_t after iteration n. B and S are two ConvGRU (Ballas et al., 2015) units that respectively output features $x_{bt}^{(n)}$ and $x_{st}^{(n)}$ together with hidden states $b_t^{(n)}$ and $s_t^{(n)}$. f_{enc} and f_{dec1}, f_{dec2} are convolutional encoder and decoders in the U-Net (Ronneberger et al., 2015), which is used as the backbone of the DR component to capture course-to-fine features of reconstructed images. Equation (5) is visualized in figure 1 for better understanding. One benefit here is that regularization function $R(\cdot)$ in Equation (1) is now built upon the convolutional network for automated feature learning and hence avoid the requirements of prior knowledge on the selection of R. $DC(\cdot)$ is the differentiable DC layer (Schlemper et al., 2017) that takes the same effect as the data consistency term $\|F_u(z_t) - y_t\|_2^2$ in Equation (1) to force the reconstructed image to be consistent with the input data. It fills the reconstructed image z_t with the original values of input data y_t in the Fourier domain by the sampled mask Ω.

More importantly, in order to capture dynamic information of image sequence during each iteration, we introduce two kinds of ConvGRU units, *i.e.* B and S, inspired by the work (Qin et al., 2018) in Equation (5). The difference between B and S is that GRU unit S is used to improve the performance of image z_t over N iterations while the role of B is to connect dynamic information of neighboring images z_{t-1} and z_t, which is implemented by initializing hidden state $b_t^{(0)}$ as $b_{t-1}^{(N)}$. Finally, we impose l_1 loss on the reconstructed images $\{z_t^N\}$ with respect to ground truth for penalizing.

2.2.2. MOTION ESTIMATION

In analogy to Equation (2), the Motion Estimation (ME) component takes as input the sequence of reconstructed images $\{z_t\}_{t\in[T]}$ and learn to predict displacement motion fields $\{v_t\}_{t\in[T]}$. As shown in figure 2, our proposed ME component embraces two parts. One is a FlowNet backboned by

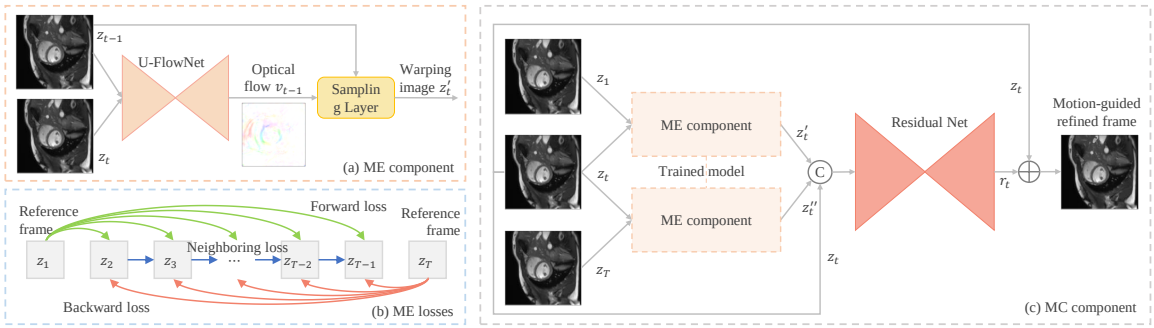

Figure 2: The network architecture of ME/MC components and ME losses.

convolutional U-Net (U-FlowNet) for motion field estimation. The other is a differentiable sampling layer based on Spatial Transformer Network (STN) (Jaderberg et al., 2015), which endows convolutional network with the ability to warp the spatial deformation between images. Unlike traditional optimization algorithms for motion estimation that depend on a strong assumption that the brightness of two frames should be consistent and the movement of the foreground object is small, our method does not succumb to any assumption and hence is more applicable in practical dynamic MRI reconstruction. The performance of ME is heavily affected by noisy input, therefore it is pre-trained with two fully sampled images z_{t-1} and z_t. The image pair is first fed to the U-FlowNet, which produces two-channel displacement v_{t-1} along the x and y directions. Then, the sampling layer warps z_{t-1} towards z_t by using v_{t-1} and yields a warping image denoted by z'_t through differentiable bilinear interpolation. This leads to a natural re-formalization of motion estimation (ME) loss ℓ_{me} between z_{t-1} and z_t from Equation (2):

$$\ell_{me}(z_{t-1}, z_t) = \left\| z'_t - z_t \right\|_1 + \beta \left\| v_{t-1} \right\|_1 + \gamma \| v_{t-1} \|_{TV} . \tag{6}$$

The first term is an image reconstruction loss used to keep the majority of high-frequency parts on images. Two additional regularization terms reinforce constraints on the motion field v_{t-1}, where l_1 regularization is to suppress unreal large magnitude of displacement and total-variation (TV) regularization is to make the displacement locally smooth.

In addition, the above loss only enforces temporal consistency between consecutive frames, but there is no guarantee for long-term coherence. Therefore, we consider to train the U-FlowNet with three sets of ME losses to capture long-term motion information, as illustrated in figure 2.

$$\mathscr{L}_{me}(\{z_t\}) = \sum_{t=2}^{T-1} \ell_{me}(z_1, z_t) + \sum_{t=2}^{T-1} \ell_{me}(z_t, z_T) + \sum_{t=2}^{T-2} \ell_{me}(z_t, z_{t+1}), \tag{7}$$

where three terms on the right-hand-side are respectively forward ME loss, backward ME loss and neighboring ME loss.

2.2.3. MOTION COMPENSATION

Motion Compensated (MC) component is used to refine reconstructed images $\{z_t\}_{t\in[T]}$ through motion information and to generate motion compensated image $\{c_t\}_{t\in[T]}$. By following the work (Jung et al., 2010), during the MC stage, we also add two additional fully sampled reference frames

to learn more accurate displacement motion fields. The pre-trained U-FlowNet is fixed and directly used as an operator in the MC component. As shown in figure 2, the MC component takes a reconstructed image z_t from the DR component and two reference frame z_1 and z_T as input. It first retrieves two warping images z_t' and z_t'' from the ME component by feeding z_1, z_t and z_t, z_T respectively. These two images represent forward and backward motion information, which is then concatenated and fed to a residual network to generate residual information r_t, as described in Equation (3). Finally, the reconstructed image z_t together with the residual r_t are summed up to generate the motion-guided refined image c_t, which is penalized by l_1 loss with respect to the ground truth image.

3. Experiment

Evaluation Dataset: We experiment with a short-axis (SAX) cardiac dataset composed of 15 patients. Each subject contains around 12 SAX planes and each plane includes 24 phases (2D images) that form a whole cardiac cycle. The image resolution is normalized to $1.25mm$ and image size is cropped to 152×152 pixels. In order to simulate k-space data, we adopt the same Cartesian under-sampling method as introduced in (Jung et al., 2007), which assumes that sampled mask Ω follows a zero-mean Gaussian distribution and keeps 8 center spatial frequencies. We consider two different settings on the dataset respectively with under-sampling rates of 20% (or acceleration rate $5\times$) and 12.5% ($8\times$). For convenience, we refer to these two cases as *Rate* $5\times$ and *Rate* $8\times$. We perform 3-fold cross-validation in the following experiments that each fold contains 10 training subjects and 5 test subjects.

Implementation Details: We implement all the deep learning models with PyTorch and train them on NVIDIA K80. All models are trained for total 80 epochs using Adam optimizer, with initialized learning rate of 5×10^{-4} and decreasing rate of 0.5 for every 20 epochs. Due to hardware limitation, the number of iterations is set to be $N = 3$ and the length of image sequence is $T = 12$.

3.1. Comparison to State-of-the-Art

In this experiment, we evaluate the dynamic reconstruction performance of our proposed methods quantitatively and qualitatively in both cases of *Rate* $5\times$ and *Rate* $8\times$. We consider three variants of our models: DRN w/o GRU (the one without GRU hidden unit), DRN (the one with DR component

Table 1: Average performance of dynamic MRI reconstruction on the test subjects in both cases of *Rate* $5\times$ and *Rate* $8\times$. The best results are highlighted in bold font.

Method	NRMSE↓	PSNR↑	SSIM↑	NRMSE↓	PSNR↑	SSIM↑
	$5\times$			$8\times$		
k-t SLR	0.0934	21.0858	0.6794	0.1054	19.9504	0.6193
k-t FOCUSS	0.0766	22.7471	0.6581	0.0879	21.4063	0.5920
k-t FOCUSS+ME/MC	0.0758	22.8139	0.6701	0.0854	21.6547	0.6131
DC-CNN(3D)	0.0360	29.1292	0.8449	0.0513	25.9709	0.7441
DRN w/o GRU	0.0381	28.7187	0.8286	0.0519	25.9120	0.7448
DRN	0.0349	29.5394	0.8502	0.0485	26.5275	0.7687
MODRN	**0.0274**	**32.0403**	**0.9104**	**0.0364**	**29.4774**	**0.8702**

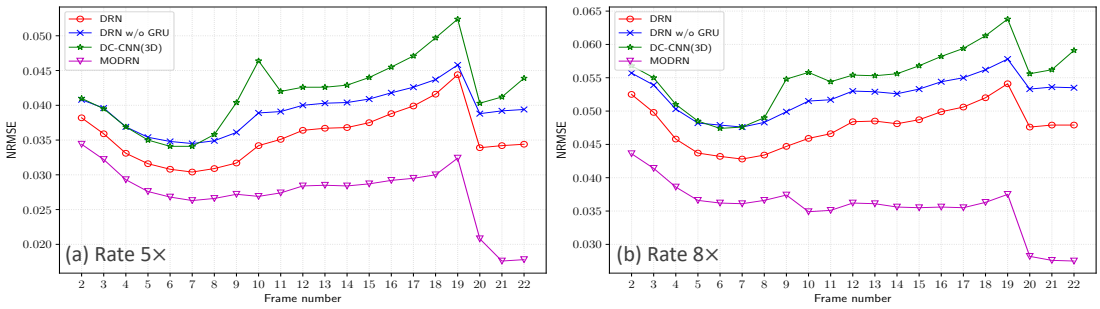

Figure 3: NRMSE curves of deep learning methods within one cardiac cycle in two cases.

only) and MODRN (the complete version). We compare with four state-of-the-art approaches including k-t SLR (Lingala et al., 2011), k-t FOCUSS (Jung et al., 2009), k-t FOCUSS+ME/MC (Jung et al., 2010) and DC-CNN (3D) (Schlemper et al., 2017). The first three are traditional CS-based methods and only k-t FOCUSS+ME/MC includes ME/MC procedures. The last one is also a deep learning based method that explores spatio-temporal information using 3D convolution. Three common quantitative metrics are used: root square mean error (NRMSE), peak signal-to-noise ration (PSNR) and structural similarity index measure (SSIM).

Quantitative Results: The results of all methods are reported in Table 1. We observe that all our methods consistently outperform four state-of-the-art approaches in both *Rate* 5× and *Rate* 8× cases. In particular, MODRN achieves the best performance for all metrics, mainly attributing to the motion information exploited by ME/MC components. We also find that DRN outperforms DRN w/o GRU by a large margin, which indicates the importance of utilizing dynamic sequence of image.

To further investigate the performance of four deep learning methods, we plot NRMSE values within a complete cardiac cycle of one example in Figure 3. It shows that our method MODRN consistently achieves the smallest error of dynamic reconstruction for the sequence of images. In contrast, the models without ME/MC are unstable along the temporal dimension, especially in the case of DC-CNN(3D). For example, in the case of *Rate* 8×, the gap between DRN and MODRN model become larger, which implies the significance of using motion information.

Qualitative Results: We visualize the reconstructed images and error with respect to ground truth of all methods in Figure 4. It is obvious that all CS-based methods have streaking artifacts and larger

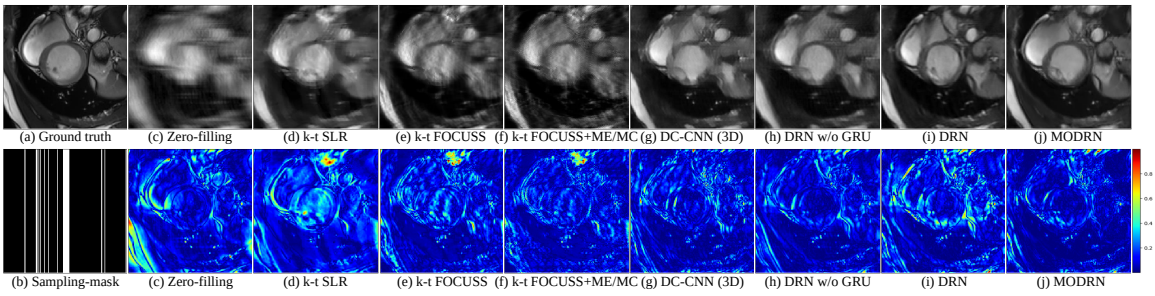

Figure 4: Visualization of reconstructed images and errors in the case of *Rate* 8×.

Method	Dice↑	HD↓
Reference	0.8130	1.9254
Lucas-Kanade	0.8125	1.9577
U-FlowNet-A	0.8297	1.8755
U-FlowNet-B	**0.8306**	**1.8584**

Table 2: Motion estimation results.

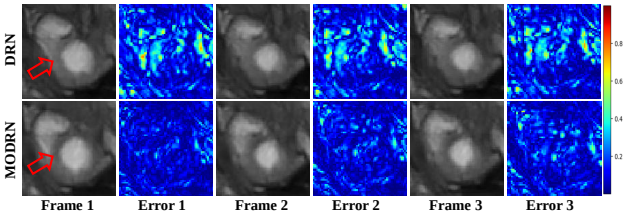

Figure 5: Motion compensation results.

reconstruction error while our MODRN model eliminates the most blurring artifacts and recovers more high-frequency details.

3.2. Motion Estimation/Compensation Analysis

First, we consider to evaluate the motion estimation results generated by the U-FlowNet from ME component. Two baseline methods, Reference and Lucas-Kanade, are compared with our U-FlowNet-A (trained with only neighboring loss) and U-FlowNet-B (trained with combined loss). Reference method directly calculates metrics using the segmentation of the target phase and the reference phase. Since it is impractical to obtain the ground truth of optical flow from cardiac MRI, we compute the overlapped area of the myocardium between the targeting image and the warping image. In particular, we calculate the average Dice's score and Hausdorff Distance between z_1 and other frames, z_T and other frames and also neighboring frames. The results of 3-fold cross-validataion are reported in Table 2. We observe that U-FlowNet-B method achieves the best performance, which indicates that compared with neighboring loss, our combined loss contributes more to accurate motion estimation with large movement between frames.

Second, we compare the quality of motion-guided refined image by MC component of MODRN with that of reconstructed image by DRN alone. The results of three consecutive frames are visualized in Figure 5. We can observe clear improvements of MODRN that its reconstruction error is reduced around cardiac region and no noticeable artifact is generated.

4. Conclusion

We present a novel deep learning based approach called MODRN for motion-guided dynamic MRI reconstruction problem. It is featured by a dynamic reconstruction (DR) component for preliminary image reconstruction from under-sampled k-space data, and a motion estimation (ME) component to predict motion of image sequence, which is further exploited by a motion compensation (MC) component to refine the motion-guided reconstructed images. We extensively evaluate our approach on a short-axis cardiac dataset in two settings. The experimental results show the effectiveness of MODRN compared to state-of-the-art methods and prove the significance of motion information from ME/MC components.

References

Wenjia Bai and Michael Brady. Motion correction and attenuation correction for respiratory gated pet images. *IEEE transactions on medical imaging*, 30(2):351–365, 2011.

Nicolas Ballas, Li Yao, Chris Pal, and Aaron Courville. Delving deeper into convolutional networks for learning video representations. *arXiv preprint arXiv:1511.06432*, 2015.

Jose Caballero, Christian Ledig, Andrew Aitken, Alejandro Acosta, Johannes Totz, Zehan Wang, and Wenzhe Shi. Real-time video super-resolution with spatio-temporal networks and motion compensation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4778–4787, 2017.

Chong Chen, Barbara Gris, and Ozan Öktem. A new variational model for joint image reconstruction and motion estimation in spatiotemporal imaging. *arXiv preprint arXiv:1812.03446*, 2018.

Fabian Gigengack, Lars Ruthotto, Martin Burger, Carsten H Wolters, Xiaoyi Jiang, and Klaus P Schafers. Motion correction in dual gated cardiac pet using mass-preserving image registration. *IEEE transactions on medical imaging*, 31(3):698–712, 2012.

Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3): 185–203, 1981.

Qiaoying Huang, Dong Yang, Pengxiang Wu, Hui Qu, Jingru Yi, and Dimitris Metaxas. Mri reconstruction via cascaded channel-wise attention network. *arXiv preprint arXiv:1810.08229*, 2018.

Max Jaderberg, Karen Simonyan, Andrew Zisserman, et al. Spatial transformer networks. In *Advances in neural information processing systems*, pages 2017–2025, 2015.

Hong Jung and Jong Chul Ye. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques. *International Journal of Imaging Systems and Technology*, 20(2):81–98, 2010.

Hong Jung, Jong Chul Ye, and Eung Yeop Kim. Improved k–t blast and k–t sense using focuss. *Physics in Medicine & Biology*, 52(11):3201, 2007.

Hong Jung, Kyunghyun Sung, Krishna S Nayak, Eung Yeop Kim, and Jong Chul Ye. k-t focuss: a general compressed sensing framework for high resolution dynamic mri. *Magnetic resonance in medicine*, 61(1):103–116, 2009.

Hong Jung, Jaeseok Park, Jaeheung Yoo, and Jong Chul Ye. Radial k-t focuss for high-resolution cardiac cine mri. *Magnetic Resonance in Medicine*, 63(1):68–78, 2010.

Florian Knoll, Christian Clason, Kristian Bredies, Martin Uecker, and Rudolf Stollberger. Parallel imaging with nonlinear reconstruction using variational penalties. *Magnetic resonance in medicine*, 67(1):34–41, 2012.

Sajan Goud Lingala, Yue Hu, Edward DiBella, and Mathews Jacob. Accelerated dynamic mri exploiting sparsity and low-rank structure: kt slr. *IEEE transactions on medical imaging*, 30(5): 1042–1054, 2011.

Kai Lønning, Patrick Putzky, Matthan WA Caan, and Max Welling. Recurrent inference machines for accelerated MRI reconstruction. In *International Conference on Medical Imaging with Deep Learning (MIDL 2018)*, 2018.

Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'81, pages 674–679, San Francisco, CA, USA, 1981. Morgan Kaufmann Publishers Inc. URL http://dl.acm.org/citation.cfm?id=1623264.1623280.

Osama Makansi, Eddy Ilg, and Thomas Brox. End-to-end learning of video super-resolution with motion compensation. In *German conference on pattern recognition*, pages 203–214. Springer, 2017.

Simon Meister, Junhwa Hur, and Stefan Roth. Unflow: Unsupervised learning of optical flow with a bidirectional census loss. *arXiv preprint arXiv:1711.07837*, 2017.

Chen Qin, Joseph V Hajnal, Daniel Rueckert, Jo Schlemper, Jose Caballero, and Anthony N Price. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE transactions on medical imaging*, 2018.

Zhe Ren, Junchi Yan, Bingbing Ni, Bin Liu, Xiaokang Yang, and Hongyuan Zha. Unsupervised deep learning for optical flow estimation. In *AAAI*, volume 3, page 7, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for mr image reconstruction. In *International Conference on Information Processing in Medical Imaging*, pages 647–658. Springer, 2017.

Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE transactions on Medical Imaging*, 37(2):491–503, 2018.

Jian Sun, Huibin Li, Zongben Xu, et al. Deep admm-net for compressive sensing mri. In *Advances in Neural Information Processing Systems*, pages 10–18, 2016.

J Trzasko, A Manduca, and E Borisch. Local versus global low-rank promotion in dynamic mri series reconstruction. In *Proc. Int. Symp. Magn. Reson. Med*, page 4371, 2011.

Alice Chieh-Yu Yang, Madison Kretzler, Sonja Sudarski, Vikas Gulani, and Nicole Seiberlich. Sparse reconstruction techniques in mri: methods, applications, and challenges to clinical adoption. *Investigative radiology*, 51(6):349, 2016.

Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487, 2018.