JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ACCELERATING MAGNETIC RESONANCE IMAGING VIA DEEP LEARNING

Prof. Jyoti Ishwar Nandalwar

Assistant Professor, CSE Department, SSWCOE, Solapur

Prof. Tukaram Appasaheb Chavan

Professor & Principal, SSWCOE, Solapur

Abstract: This paper proposes a deep learning approach for accelerating magnetic resonance imaging (MRI) using a large number of existing high quality MR images as the training datasets. An off-line convolutional neural network is designed and trained to identify the mapping relationship between the MR images obtained from zero-filled and fully-sampled k-space data. The network is not only capable of restoring fine structures and details but is also compatible with online constrained reconstruction methods. Experimental results on real MR data have shown encouraging performance of the proposed method for efficient and effective imaging. Under sampling in k-space violates Nyquist Sampling and creates artifacts in the image domain. In the proposed method, we consider the reconstruction problem as a de-aliasing problem in complex spatial domain. To test the proposed method, from fully sampled k-space data under sampling in k-space was performed in the phase-encode direction based on a probability density function which ensures maximum rate of sampling in low frequency regions. For the deep convolutional neural network (CNN) we chose the U-net architecture.
Keywords: Deep learning, magnetic resonance imaging, prior knowledge, convolutional neural network

1. INTRODUCTION

Magnetic resonance imaging (MRI) is an indispensable tool for medical diagnosis, disease staging and clinical research due to its strong capability in providing rich anatomical and functional information and its non-radiation and non-ionizing nature. However, most of advanced applications such as cardiovascular imaging, functional MRI, magnetic resonance spectroscopy and parameter mapping are not yet widely used in clinic due to the long scanning time of MRI [1]. To accelerate MR scans, efforts are mainly in three directions 1) physics based fast imaging sequences, 2) hardware based parallel imaging and 3) signal processing based MR image reconstruction from reduced samples. The combination of these techniques have also shown their appearance in a great number of publications [6]. The first two categories and a few techniques of the third category (e.g. partial Fourier) have already been applied in commercial scanners as a routine protocol for shortening the total scan time [5].

1.1 Background

Phase Contrast MRI is an imaging method which can noninvasively derive hemodynamic information inside the human body. Moving spins accumulate phase in a linearly spatially variant magnetic field direction with the phase offset being proportional to velocity in that direction. The velocity information can be collected from complex valued MR images by subtracting flow encoded image from a flow compensated image. However with the advent of 4D flow imaging, scan efficiency has become a significant issue. To reduce scan time, in the past, various methods have been proposed. These includes k-t SENSE [1] , k-t BLAST [2] and k-t GRAPPA[3], non-Cartesian trajectory acquisitions namely radial or spiral acquisitions [4,5, 19] and

Compressed sensing [6,7,20] methods. In compressive sensing method, scan time is reduced by taking fewer sample in k-space (reduced number of phase-encoding steps) and image is reconstructed from under sampled k-space data by exploiting image sparsity in pixel domain or a transform domain.

1.2 Motivation

Deep learning, a technique attempting to model high-level abstractions in data with multiple processing layers, has shown explosive popularity in recent years with the availability of powerful GPUs. Especially, convolutional neural network (CNN) has exhibited its significance in addressing large-scale vision tasks such as action recognition [15], image classification [16], super-resolution [17] and demising [18]. CNNs have quite a few merits, such as the lack of dependence on prior-knowledge, no need to design hand-engineered features and strong ability to capture image structures, which motivated us to employ it for MR image reconstruction from under sampled k-space data. What is more important is that the relationship between the zero-filled MR image and the ground-truth image can be interpreted in a convolutional way, which will be explained in the theory parts. In this paper we propose an off-line convolutional neural network to learn an end-to-end mapping between zero-filled and fully-sampled MR images. This network is not only capable of restoring the details and fine structures of the MR images, but is also compatible with any online reconstruction algorithm for more efficient and effective imaging. We have tested the proposed method on a set of in-vivo MR data and the results have shown promising.

1.3 Review of Literature

Signal processing based methods, explores prior information on MR images and utilize them to regularize the reconstruction from under sampled K-space measurements with the advantage of no physical, physiological and hardware restrictions. Sparsity is one commonly used prior information due to the emergence of Compressed sensing (C-S) and there are also other priors being considered, such as low-rank [7], statistics distribution regularization [8], manifold fitting [9], generalized series (GS) model [10] and so on. The prior information used can be roughly categorized into adaptive and non-adaptive ones. For example, total variation and Wavelet transform, singular value decomposition (SVD) are non-adaptive ones [4]; dictionary learning and data driven tight frames are adaptive [1].

Generally, adaptive priors can capture more structures while non-adaptive ones are more computationally efficient. Nevertheless, despite all the successes achieved by the aforementioned methods, it is easy to discover that they only explore the prior information either directly from the image to be reconstructed or with very few reference images involved. Considering the similarity on the anatomic information of the same organ/tissue between different people and the enormous images acquired every day, it is straightforward to collect many reference images and learn an off-line prior model to aid online fast imaging [6].

In machine learning one develops and studies methods that give computers the ability to solve problems by learning from experiences. The goal is to create mathematical models that can be trained to produce useful outputs when fed input data. Machine learning models are provided experiences in the form of training data, and are tuned to produce accurate predictions for the training data by an optimization algorithm. The main goal of the models are to be able to generalize their learned expertise, and deliver correct predictions for new, unseen data[7].

A model's generalization ability is typically estimated during training using a separate data set, the validation set, and used as feedback for further tuning of the model. After several iterations of training and tuning, the final model is evaluated on a test set, used to simulate how the model will perform when faced with new, unseen data [8].

There are several kinds of machine learning, loosely categorized according to how the models utilize its input data during training. In reinforcement learning one constructs agents that learn from their environments through trial and error while optimizing some objective function. Clustering is a prime example. Most of today's machine learning systems belong to the class of supervised learning [9].

Here, the computer is given a set of already labeled or annotated data, and asked to produce correct labels on new, previously unseen data sets based on the rules discovered in the labeled data set. From a set of input-output examples, the whole model is trained to perform specific data-processing tasks. Image annotation using human-labeled data, e.g. classifying skin lesions according to malignancy [6] or discovering cardiovascular risk factors from retinal fundus photographs [7], are two examples of the multitude of medical imaging related problems attacked using supervised learning [10].

Machine learning has a long history and is split into many sub-fields, of which deep learning is the one currently receiving the bulk of attention. There are many excellent, openly available overviews and surveys of deep learning. For short general introductions to deep learning. Artificial neural networks (ANNs) is one of the most famous machine learning models, introduced already in the 1950s, and actively studied [12].

Roughly, a neural network consists of a number of connected computational units, called neurons, arranged in layers. There's an input layer where data enters the network, followed by one or more hidden layers transforming the data as it flows through, before ending at an output layer that produces the neural network's predictions [13].

The network is trained to output useful predictions by identifying patterns in a set of labeled training data, fed through the network while the outputs are compared with the actual labels by an objective function. During training the network's parameters – the strength of each neuron – is tuned until the patterns identified by the network result in good predictions for the training data. Once the patterns are learned, the network can be used to make predictions on new, unseen data, i.e. generalize to new data [14].

It has long been known that ANNs are very flexible, able to model and solve complicated problems, but also that they are difficult and very computationally expensive to train. This has lowered their practical utility and led people to, until recently, focus on other machine learning models. But by now, artificial neural networks form one of the dominant methods in machine learning, and the most intensively studied. This change is thanks to the growth of big data, powerful processors for parallel computations (in particular, GPUs), some important tweaks to the algorithms used to construct and train the networks, and the development of easy-to-use software frameworks. The surge of interest in ANNs leads to an incredible pace of developments, which also drives other parts of machine learning with it [16].

The nonlinear functions σk are typically sigmoid functions or ReLUs, as discussed below, and the θk are matrices of numbers, called the model's weights. During the training phase, the network is fed training data and tasked with making predictions at the output layer that match the known labels, each component of the network producing an expedient representation of its input. It has to learn how to best utilize the intermediate representations to form a complex hierarchical representation of the data, ending in correct predictions at the output layer [17].

Training a neural network means changing its weights to optimize the outputs of the network. This is done using an optimization algorithm, called gradient descent, on a function measuring the correctness of the outputs, called a cost function or loss function. The basic ideas behind training neural networks are simple: as training data is fed through the network, compute the gradient of the loss function with respect to every weight using the chain rule, and reduce the loss by changing these weights using gradient descent. But one quickly meets huge computational challenges when faced with complicated networks with thousands or millions of parameters and an exponential number of paths between the nodes and the network output. The techniques designed to overcome these challenges gets quite complicated [18].

2. PROPOSED METHODOLOGY

This section provides the concept and theoretical foundation for learning the convolutional neural network for under sampled MR image reconstruction. Fig. 1 presents the flowchart of the proposed method.

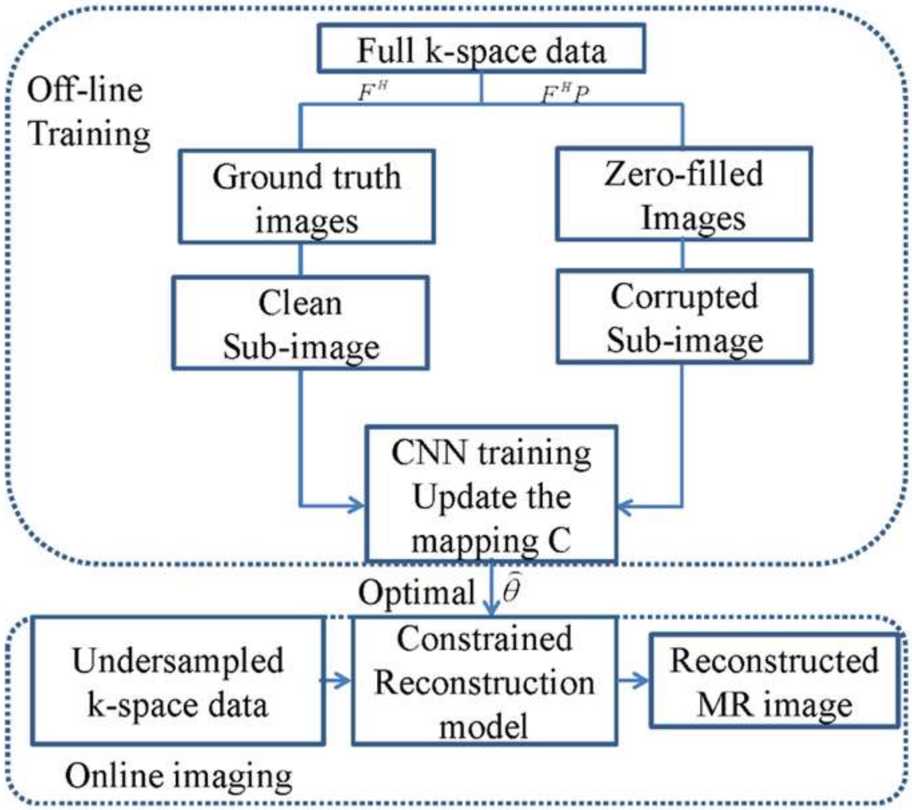

Fig. 1. The flowchart of the proposed method

2.1. Overall Training Formulation

Consider the under sampled raw K-space data as

f=PFu ……………………………………………………………………………………..……. (1)

where P is a diagonal matrix representing the under sampling mask, F denotes the full Fourier encoding matrix normalized as $F^H F = I$, u is the original (ground truth) image and therefore Fu represents the full k-space data. H represents the Hermitian transpose operation. The zero-filled MR image z is generated as the direct inverse transform of the observed data like

z=FHPFu ……………………………………………………………………………….…… (2)

As stated in [19], in terms of linear algebra, the circular convolution of a signal u with a pulse p can be written as F^H PFu, where P is a diagonal matrix whose non-zero entries are the Fourier transform of p.

We try to learn a fully convolutional neural network to restore accurate MR images from undersampled Fourier data. Given a pre-acquired dataset of MR corrupted/ground truth images, we try to minimize the following objective

∑t=1T‖C(zt;Θ)−ut‖22}……………………………………………………………………..…. (3)

where C means the end-to-end mapping function with its hidden parameters $\Theta =\{(W_1, b_1), \ldots, (W_l, b_l), \ldots, (W_L, b_L)\}$ to be estimated and T is the total number of training samples. In order to increase the robustness of the network, we generate more training samples by separating the whole image pairs into overlapping subimage pairs $x_{t,n}$ and $y_{t,n}$ and minimize

$\sum_{t=1}^{T}\sum_{n=1}^{N}\|C(x_{t,n};\Theta)-y_{t,n}\|_2^2\}$ ……………………………………………………..……….(4)

For the simplicity of explanation, we only consider one pair x and y in the following demonstration.

2.2. Forward-pass training sub problems

2.2.1. Feature generation

Unlike sparse representation, where each extracted image patch is approximated by a set of pre-trained bases, we use the equivalent convolution operation [17] and transfer the optimization of the bases into the network learning process. Therefore, the first layer of network can be described as follows

$C_1=\sigma(W_1 * x+b_1)$………………………………………………………………………………. (5)

Where W_1 denotes the convolution operator of size $c \times M_1 \times M_1 \times n_1$ and b_1 is the n_1 dimensional bias with its element associated with a filter. Here, c is the number of the image channels, M_1 means the filtered size and n_1 is the number of filters. We adopt the rectified linear unit (ReLu, max(0, x)) here for the nonlinear responses, which can be computed very efficiently [17].

2.2.2. Nonlinear mapping

We further perform non-linear mapping to project the n_{l-1} dimensional vectors into an n_l one, which is conceptually the refined feature and structure to represent the full-data-reconstructed image

$C_l=\sigma(W_l * C_{l-1}+b_l)$……………………………………………………………………………..(6)

Where W_l is of a size $n_{l-1} \times M_l \times M_l \times n_l$.

2.2.3. Last Layer convolution

To produce the final predicted image from CNN, we explore another layer of convolution and hope to learn a set of linear filters W_L which are capable of projecting the coefficients onto the image domain

$C_L=\sigma(W_L * C_{L-1}+b_L)$ ……………………………………………………………………..… (7)

Where W_L is of a size $n_{L-1} \times M_L \times M_L \times c$. To sum up, we have designed an L-layer convolutional neural network to learn the mapping relationship:

3. RESULTS AND DISCUSSION

3.1 Datasets:

The training data consists of over 500 fully sampled MR brain images we collected from a 3T scanner (SIEMENS MAGNETOM). The images are of a great diversity including axial, sagittal, horizontal ones, and different contrast ones such as T1, T2 and PD-weighted images and of a variety of sizes. Informed consent was obtained from the imaging subject in compliance with the Institutional Review Board policy. Under sampled measurements were retrospectively obtained using the 1D low-frequency sampling mask and the 2D Poisson disk sampling mask. The large amount of corrupted/ground truth sub image pairs are then generated with the size of 33 × 33. Finally we use 90% of the sub image pairs as the training dataset and the rest 10% for validating the training process.

3.2 Implementation details:

We use three layers of convolution for the network. The parameters are respectively set as $n_1 = 64$, $n_2 = 32$, $M_1 = 9$, $M_2 = 5$ and $M_3 = 5$. The filter weights of each layers are initialized by random values from a Gaussian distribution with zero mean and standard deviation 0.001. The bias are all initialized as 0. The training takes about three days, on a workstation equipped with 128G memory and a processor of 16 cores (Intel Xeon (R) CPU E5–2680 V3 @2.5GHz).

Figure 2 shows a set of reconstruction results of a transversal brain image. The brain dataset was obtained fully-sampled with 12-channel head coil and T2-weighted turbo spin-echo (TSE) sequence (TE = 91.0ms, TR = 5000ms, FOV = 20 × 20cm, matrix = 256 × 270, slice thickness = 3mm) via 3T scanner. And the data was then under sampled retrospectively with 1D low-frequency sampling mask at an acceleration factor of 3 and the 2D Poisson disk at an acceleration factor of 5. We also tested the proposed method on a sagittal brain image which was acquired on a GE 3T scanner (GE Healthcare, Waukesha, WI) with a 32-channel head coil and 3D T1-weighted spoiled gradient echo sequence (TE=minimum full, TR= 7.5ms, FOV=24 × 24 cm, matrix = 256 × 256, slice thickness=1.7mm). We can observe from the images that there are quite a few details and structures captured by the network. Furthermore, the image generated by the simple reconstruction model is quite close to the original image. According to Fig. 3f, we can see the difference image is noise-like and consists only the contour information. There are no obvious details and structures lost. It demonstrates that the proposed network is capable of restoring the details and fine structures which are discarded in the zero-filled MR image. Furthermore, although the offline training takes roughly three days, under the same GPU configurations, it takes far less than 1 second for each online MR reconstruction case.

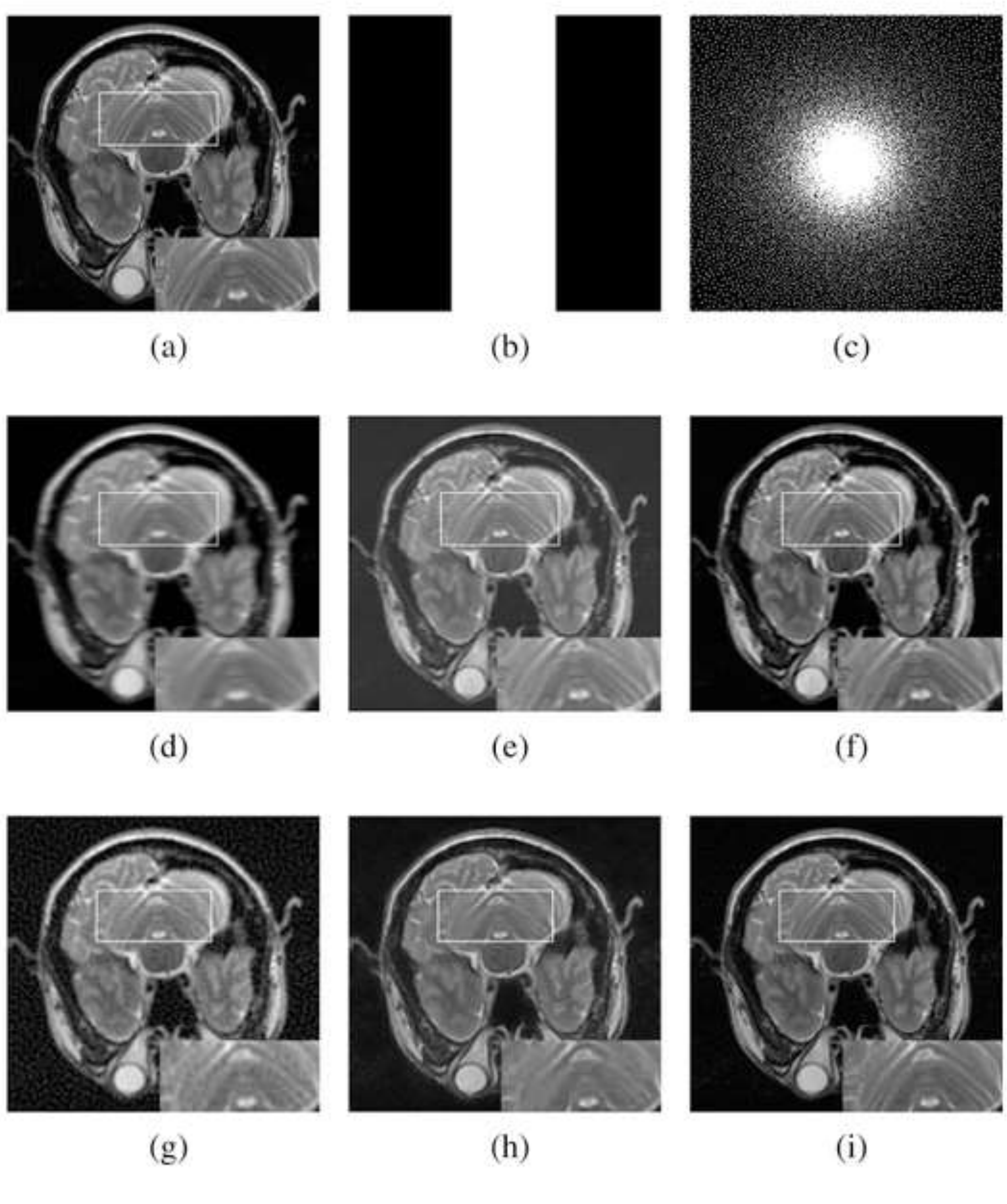

Fig. 2. (a) Ground truth image; (b) 1 D central low-frequency sampling mask with acceleration factor of 3; (c) 2D poison under sampling mask with acceleration factor of 5; (d)(e)(f) the zero-filled MR image, network output and reconstruction result from 1D under sampled data; (g)(h)(i) the zero-filled M-R image, network output and reconstruction result from 2D under sampled data.

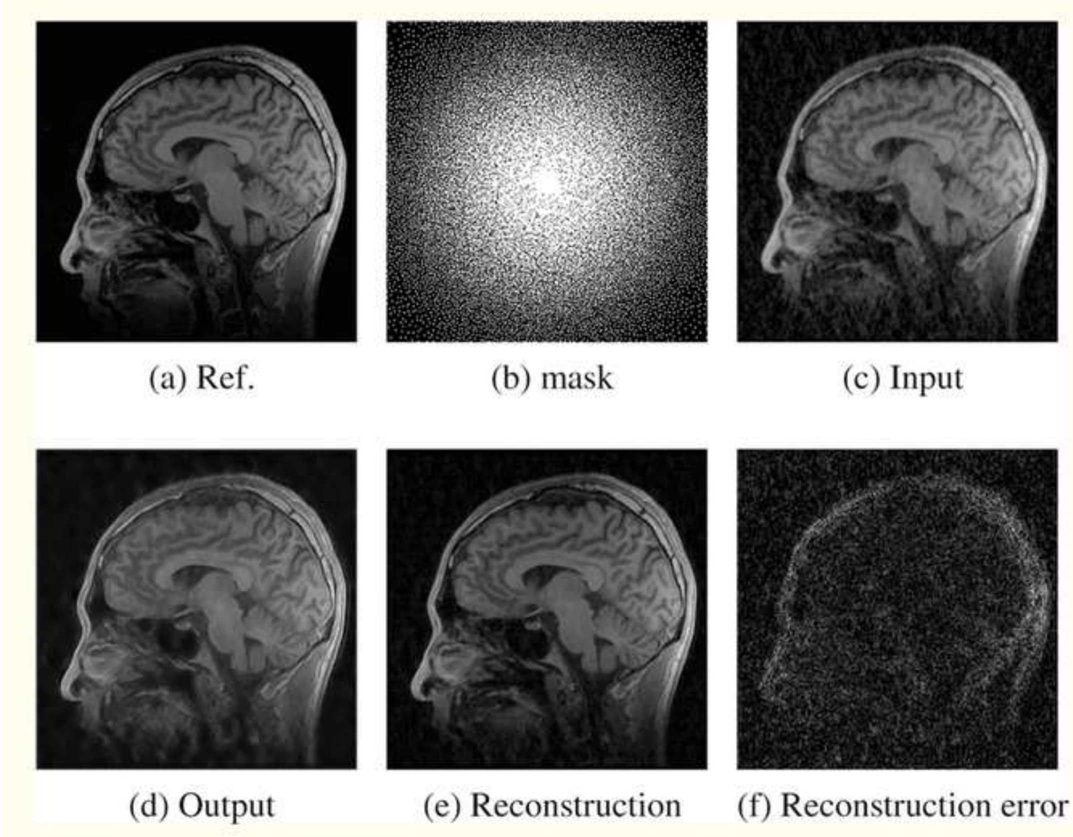

Fig. 3. The test results on another sagittal brain image at an acceleration factor of 3

CONCLUSIONS

An off-line convolutional neural network for accelerating M-R imaging is proposed in this paper, which includes the brief review and discussion of the concept, theoretical foundation, implementation and application of this network for under sampled MR image reconstruction. The experimental results on in-vivo MR images have shown the proposed network is capable of restoring the details and fine structures that are lost in the zero-filled MR image. We also provide two options for combing the proposed network with online CS-MRI methods for more efficient and effective imaging. More extensive experimental results will be provided in the future journal paper.

REFERENCES

[1] D. Kim, H. Dyvorne, R. Otazo, L. Feng, D. Sodickson and V. Lee, "Accelerated phase-contrast cine MRI using k-t SPARSESENSE", Magnetic Resonance in Medicine, vol. 67, no. 4, pp. 1054-1064, 2011. Available: 10.1002/mrm.23088.

[2] M. Carlsson et al., "Quantification and visualization of cardiovascular 4D velocity mapping accelerated with parallel imaging or k-t BLAST: head to head comparison and validation at 1.5 T and 3 T", Journal of Cardiovascular Magnetic Resonance, vol. 13, no. 1, 2011.

[3] Z. Stankovic et al., "K-t GRAPPA-accelerated 4D flow MRI of liver hemodynamics: influence of different acceleration factors on qualitative and quantitative assessment of blood flow", Magnetic Resonance Materials in Physics, Biology and Medicine, vol. 28, no. 2, pp. 149-159, 2014.

[4] M.Kadbi et al., "4D UTE flow: A phase-contrast MRI technique for assessment and visualization of stenotic flows", Magnetic Resonance in Medicine, vol. 73, no. 3, pp. 939-950, 2014. Available: 10.1002/mrm.25188

[5] M.Negahdar, M. Kadbi, M. Kendrick, M. Stoddard and A. Amini, "4D spiral imaging of flows in stenotic phantoms and subjects with aortic stenosis ", Magnetic Resonance in Medicine, vol. 75, no. 3, pp. 1018-1029, 2015.

[6] Zhang Xiaoqun, Burger Martin, Bresson Xavier, and Osher Stanley, “Bregmanized nonlocal regularization for deconvolution and sparse reconstruction,” SIAM J. Imaging Sciences, vol. 3, no. 3, pp. 253–276, 2010.

[7] Lingala Sajan Goud, Hu Yue, DiBella Edward, and Jacob Mathews, “Accelerated dynamic mri exploiting sparsity and low-rank structure: kt slr,” IEEE Trans. Med. Imaging, vol. 30, no. 5, pp. 1042–1054, 2011.

[8] Wu Jiao, Liu Fang, Jiao LC, Wang Xiaodong, and Hou Biao, “Multivariate compressive sensing for image reconstruction in the wavelet domain: using scale mixture models,” IEEE Transactions on Image Processing, vol. 20, no. 12, pp. 3483–3494, 2011.

[9] Wachinger Christian, Yigitsoy Mehmet, Rijkhorst Erik-Jan, and Navab Nassir, “Manifold learning for image-based breathing gating in ultrasound and mri,” Medical image analysis, vol. 16, no. 4, pp. 806–818, 2012.

[10] Liang Zhi-Pei, Madore Bruno, Glover Gary H, and Pelc Norbert J, “Fast algorithms for gs-model-based image reconstruction in data-sharing fourier imaging,” IEEE Trans. Med. Imaging, vol. 22, no. 8, pp. 1026–1030, 2003.

[11] Wang Shanshan, Xia Yong, Liu Qiegen, Dong Pei, Feng David Dagan, and Luo Jianhua, “Fenchel duality based dictionary learning for restoration of noisy images,” IEEE Transactions on Image Processing, vol. 22, no. 12, pp. 5214–5225, 2013.

[12] Wang Shanshan, Liu Qiegen, Xia Yong, Dong Pei, Luo Jianhua, Huang Qiu, and Feng David Dagan, “Dictionary learning based impulse noise removal via L1-L1 minimization,” Signal Processing, vol. 93, no. 9, pp. 2696–2708, 2013.

[13] Liu Qiegen, Wang Shanshan, and Luo Jianhua, “A novel predual dictionary learning algorithm,” J. Vis. Commun. Image Represent, vol. 23, no. 1, pp. 182–193, 2012.

[14] Liu Qiegen, Wang Shanshan, Yang Kun, Luo Jianhua, Zhu Yuemin, and Liang Dong, “Highly undersampled magnetic resonance image reconstruction using two-level Bregman method with dictionary updating,” IEEE Trans. Med. Imaging, vol. 32, no. 7, pp. 1290–1301, 2013.

[15] Ji Shuiwang, Xu Wei, Yang Ming, and Yu Kai, “3d convolutional neural networks for human action recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 1, pp. 221–231, 2013.

[16] M. Lustig, D. Donoho and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging", Magnetic Resonance in Medicine, vol. 58, no. 6, pp. 1182-1195, 2007.

[17] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude", COURSERA: Neural Networks for Machine Learning, Resonance in Medicine, 2012

[18] T. Goldstein and S. Osher, "The Split Bregman Method for L1- Regularized Problems", SIAM Journal on Imaging Sciences, vol. 2, no. 2, pp. 323-343, 2009.

[19]S. Callahan et al., "Dual-V enc acquisition for 4D flow MRI in aortic stenosis with spiral readouts", Journal of Magnetic Resonance Imaging, 2019.

[20]F. Ong, J. Cheng and M. Lustig, "General phase regularized reconstruction using phase cycling", Magnetic Resonance in Medicine, vol. 80, no. 1, pp. 112-125, 2017.