Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-08143-6

Gibbs-ringing artifact suppression with knowledge transfer from natural images to MR images

Xiaole Zhao[1] · Huali Zhang[1] · Yuliang Zhou[1] · Wei Bian[1] · Tao Zhang[1,2,3] · Xueming Zou[1,2,3]

Received: 3 January 2019 / Revised: 14 March 2019 / Accepted: 29 August 2019 /
Published online: 21 November 2019

Abstract
Gibbs-ringing is a common artifact in magnetic resonance imaging (MRI), which is mainly caused by the finite *k*-space sampling and the truncation of high frequency (HF) information at the sampling border. It is especially visible for imaging acquisitions at low resolution and can be typically suppressed by filtering at the expense of further loss of HF components. As a classic image restoration problem in MRI, Gibbs-ringing artifact suppression can be viewed as a typical ill-posed inverse problem of image generation in computer vision community, such as image super-resolution and inpainting. Inspired by this, the present work presents a novel method to suppress the Gibbs-ringing artifacts with knowledge transfer from natural images to MR images. The highly nonlinear relation between the artifact-degraded image and the corresponding artifact-free counterpart is modeled with a typical convolutional neural network (CNN), which is first trained with natural images and then fine-tuned with MR images. Unlike many other works, we use transfer learning between different types of images to deal with regression problems, rather than classification problems. The experimental results exhibit that there exists information sharing between natural images and MR images with regard to the same problem, and the knowledge learned from natural images can indeed improve the performance of regression models on MR images.

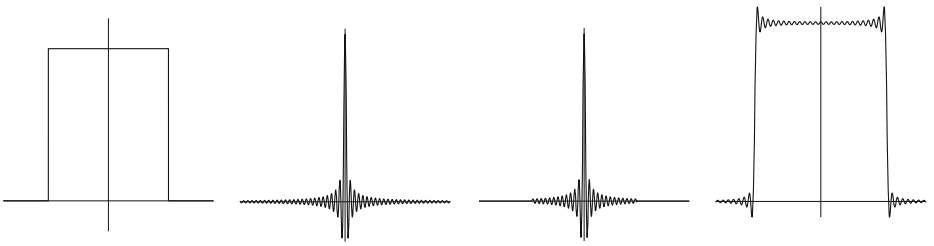
Keywords Convolutional neural networks · Gibbs-ringing artifacts · Knowledge transfer · Magnetic resonance imaging

1 Introduction

Magnetic resonance imaging (MRI) is a primary medical imaging technique based on the physical principles of nuclear magnetic resonance and mainly used to produce high quality images for both clinical diagnosis and research purposes. In the commonly applied two-dimensional Discrete Fourier Transform (2D DFT)-based MR imaging methods, the reconstructed MR images are usually affected by Gibbs-ringing artifacts, which manifest

✉ Xiaole Zhao
zxlation@foxmail.com

Extended author information available on the last page of the article.



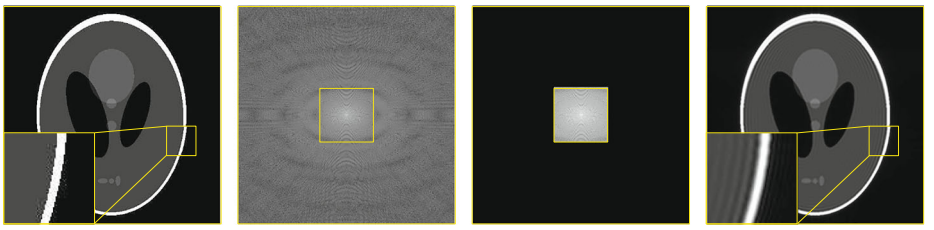
(a) Rectangle function (b) Fourier transform (c) Frequency truncation (d) Inverse recovery

Fig. 1 Gibbs-ringing artifacts generated by the truncation of HF information on a 1D continuous rectangle function. It can be seen that there are conspicuous parasitic oscillations in the vicinity of the jump discontinuity

themselves as spurious oscillations (Fig. 1) in the vicinity of high frequency (HF) region, e.g., edges and ridges at tissue boundaries [24]. In practice, these oscillatory artifacts could seriously degrade the quality of images and mislead clinical diagnosis.

In the conventional case, a 2D DFT of the finitely measured k-space data is performed to reconstruct a MR image. This strategy implicitly assumes that the Fourier transform is zero everywhere outside the sampled region. Theoretically, it corresponds to a multiplication of the true object's Fourier transform with a rectangle function (in case of 2D Cartesian sampling), which results in a convolution of the true object with a sinc function in the image space [6, 24]. Therefore, Gibbs-ringing is also known as *truncation artifact* and it usually appears as oscillating overshoots of the pixel intensity near the discontinuities in the image (Figs. 1 and 2). The problem may be alleviated by expanding sampling region or increasing the sampling density (i.e., high resolution sampling). However, this will inevitably increase the imaging time and cost. Some practical applications still use relatively low-resolution acquisitions in at least one dimension for fast imaging, thus suffering from respective artifacts [6].

Deep learning techniques [30], especially convolutional neural networks (CNNs), have rapidly become a preferred choice for analyzing and understanding digital images [36]. One consensus is that the deeper network is more difficult to be well-trained because the information is gradually weakened as the network deepens [22, 58]. This problem is even more serious for medical images due to the degradation of training examples [60]. To this end, we proposed a novel method based on deep CNNs and knowledge transfer to relieve the problem of training deep networks with degraded training samples. The proposed method uses the knowledge learned from natural images to guide the training of models using MR images. Unlike many other methods that use knowledge transfer from natural images to



(a) Artifact-free image (b) Fourier transform (c) Frequency truncation (d) Inverse recovery

Fig. 2 Gibbs-ringing artifact generation in common 2D images. It is worth noting that besides the ringing artifacts, the significant blurring effect can also be observed in (**d**), indicating the loss of HF information

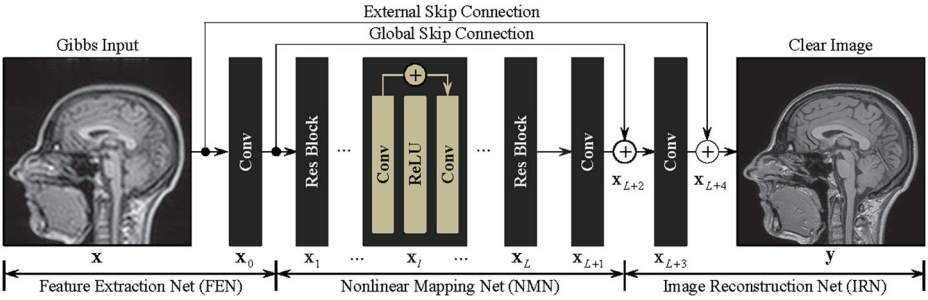


Fig. 3 The overall network structure used in this work (modified EDSR model [35]). It contains three typical phases: feature extraction, nonlinear mapping and image reconstruction. The nonlinear mapping is mainly achieved by a series of stacked residual blocks. The structure of a residual block is detailed by the lth block

medical images in classification problems, the proposed method applies it to the regression problem of dense prediction.

We construct an end-to-end model based on the EDSR model [35], which was originally designed for single (natural) image super-resolution (SISR). However, medical images are relatively simple in image structures and texture details compared with natural images (NI). This will lead to instability in model training, especially when training large scale networks [60]. Thus, we modified the original EDSR model by: (i) *introducing an external skip connection (ESC) to stabilize model training*; (2) *reducing the scale of model parameters by limiting network width (the number of feature maps)*; (3) *removing the upscale module to keep the size of the input and output the same, as shown in* Fig. 3. The model is largely downscaled and termed as Gibbs-ringing Artifact Suppression Convolutional Neural Network (GAS-CNN).

To generate training examples, we follow the pipeline shown in Fig. 2 to simulate Gibbs-ringing artifacts (4-fold truncation in k-space), where Fig. 2a and d form a pair of training sample ($\mathbf{y}$, $\mathbf{x}$). It is worth noting that the artifact-degraded image $\mathbf{x}$ is generated by zero-padding the k-space to match the size of the artifact-free image $\mathbf{y}$. However, the artifacts may appear in the images produced with or without zero-padding in real MRI scans. Zero-padding just makes the Gibbs-ringing artifacts more visible because the oscillation pattern of the artifacts is amplified by the zero-filling (i.e., sinc) interpolation [24]. In this work, we set the problem to the case of zero-padding[1] so that the artifacts in degraded images are more visible.

In summary, we present a new method for Gibbs-ringing artifact removal based on current popular deep learning and transfer learning techniques. Extensive experiments on various MR images, including proton density (PD), T1 and T2 images, verify the effectiveness of the proposed GAS-CNN model in suppressing Gibbs-ringing artifacts. More importantly, we experimentally verify that *knowledge learned from natural images can indeed improve the performance of deep regression models on medical images*. The remainder of this paper is organized as follows. In Section 2, we present some previous contributions related to this work. Next, the proposed method is detailed in Section 3 and the experimental results are presented in Section 4. We discuss some related issues and future work in Section 5. Finally, we conclude the whole work in Section 6.

[1] In the case of non-zero padding, the current problem is more like the task of SISR with the image degradation of k-space truncation.

2 Related work

2.1 Gibbs-ringing artifact suppression in MRI

Many approaches have been proposed in medical image processing community to reduce the disturbances caused by Gibbs-ringing artifacts [2, 6, 9, 24, 29, 34, 51, 56]. One of the most direct methods is image filtering, e.g., sigma or median filtering [9], Lanczos-σ approximation [17] etc. In the majority of MRI applications, these simple filters are routinely employed although many methods have been proposed [6]. However, filtering typically leads to blurring and significant reduction of HF details. More advanced methods have been proposed based on piecewise reconstruction of smooth regions using Gegenbauer polynomials [2, 17, 18]. One main drawback of these methods is the requirement for edge detection. However, it is a challenging task in itself, to detect edges accurately in artifact-degraded images. Besides, such methods are unstable in some applications due to parameter selection [7, 24]. Data extrapolation constrained with total variation (TV) regularization is another representative method [6, 40, 42]. Although these methods achieved better performance than filtering methods, they are extremely time-consuming due to the usage of iterative optimization techniques, e.g., conjugate gradient descent [6] and Euler's integration [40, 42].

2.2 Convolutional neural networks

Deep convolutional neural network (CNN) is a feed-forward artificial neural networks (ANNs) which can date back to the end of the 1980s [31]. In recent years, due to its tremendous success in computer vision community (such as image classification [20, 28, 47, 54, 57], object detection [38, 48], face recognition [45, 46], image super-resolution [12, 26, 35, 44, 59], image denoising [8, 23, 55], image retrieval [61–63] and image inpainting [4, 5] etc.), it has shown an explosive popularity in various industrial and research fields. The unprecedented success of CNNs arises mostly from the following factors [43]:

- Advanced high performance computing equipment (e.g. CPUs, GPUs and recent TPUs);
- The availability of large-scale high quality datasets (e.g. ImageNet [41], DIV2K [1] etc.);
- The development of related theories and techniques [36].

The proposed Gibbs-ringing artifact suppression approach also benefits from these advances to a large extent due to its essence of deep CNNs and transfer learning.

2.3 Knowledge transfer

Many machine learning algorithms assume that training data and testing data have the same feature space and probability distribution implicitly. However, this assumption may not hold in many real-world applications [10, 11, 32, 33]. In these cases, knowledge transfer may be an alternative way to improve model performance on the target task based on knowledge of the source task [39]. Many medical image processing applications use knowledge transfer to solve target tasks, e.g., a pre-trained model of Google's Inception v3 was fine-tuned on medical images [14, 19] and achieved the performance close to human experts [36]. However, these works are aimed at classification problems, not regression problems. An example of knowledge transfer in regression problems is presented in [35], which investigated the knowledge transfer between different scales of image super-resolution. In this work, we transfer knowledge from natural images to MR images to deal with the regression task of Gibbs-ringing artifact suppression.

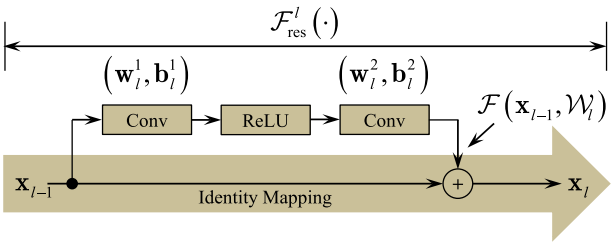


Fig. 4 The structure of a residual block proposed in [35]. It has two conv layers with a ReLU operation in the middle. We elaborate on this residual block structure in more detail

3 Proposed method

In principle, Gibbs artifact suppression can be viewed as an image generation problem where obstructing factors in the original images (Gibbs-ringing artifacts here) are required to be removed or suppressed to improve image quality. Formally, an artifact-free image $\mathbf{y} \in \mathbb{R}^{N \times M}$ is recovered from one single artifact-degraded image $\mathbf{x} \in \mathbb{R}^{N \times M}$, where M and N represent the height and width of the images, respectively. With the absence of noise and other artifacts, it can be formulated as:

$$\mathbf{y} = \mathbf{x} + \mathbf{z} \tag{1}$$

in the image space. $\mathbf{z} = \mathbf{y} - \mathbf{x}$ represents the compensation for Gibbs-ringing artifacts. Let $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z} \in \mathbb{C}^{N \times M}$ be the Fourier transform of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ respectively. Then (1) can be expressed equivalently in the k-space as follow:

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z} \implies \mathbf{Y} = \mathbf{X} \cup \mathbf{Z} \tag{2}$$

where $\cup$ indicates union or merging operation.[2] The target here is to recovery $\mathbf{y}$ from $\mathbf{x}$ by predicting the appropriate artifact compensation $\mathbf{z}$. Because there exist many nonlinear features in natural or medical images (like edge, ridge and texture etc.) and an artifact-degraded image corresponds to many possible artifact-free images, this is a highly nonlinear ill-posed inverse problem analogous to single image super resolution (SISR) [12, 13, 15, 22, 25, 26, 35, 44, 49, 50] or image inpainting [4, 5] in computer vision community.

3.1 The overall network architecture

Although the network used in this paper is derived from the EDSR model [35], we still illustrate it in details due to the structural changes described above. The overall structure of the network is shown in Fig. 3. Like many other deep learning models, it consists of three typical parts: feature extraction subnet (FE), nonlinear mapping subnet (NM) and image recovery subnet (IR). There is only one convolutional layer in both FEN and IRN, while NMN is composed of multiple residual blocks, which are the same as the residual blocks of EDSR model [35]. Specifically, it consists of two conv layers with a ReLU layer in the middle (Figs. 3 and 4).

[2] We also utilize the symbol "$\cup$" to emphasize the complementary relationship between $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$.

3.1.1 Feature extraction

This part of the network contains only one conv layer, which is equivalent to densely extract patches and then represent them by a set of pre-trained bases, e.g., PCA, DCT and Haar etc [12]. Formally, it can be represented as:

$$\mathbf{x}_0 = \mathbf{x} * \mathbf{w}_0 + \mathbf{b}_0 \tag{3}$$

where $\mathbf{x}$ denotes the artifact-degraded input image, and $\mathbf{x}_0$ is the output of FE. $\mathbf{w}_0$ and $\mathbf{b}_0$ are the filters and biases of this layer respectively, and "$*$" represents the convolution operation. Although the FE follows the structure of most previous models (e.g. [12, 13, 22, 35]) and consists of only one conv layer, it can be designed arbitrarily, e.g., [26] and [59] use two conv layers in FE, while [25] does not explicitly use FE and maps the input images to the output image directly.

3.1.2 Nonlinear mapping

The residual networks that stack multiple building blocks of the same shape are modularized [21]. This mainly refers to that the nonlinear mapping stage of the model is modularized. Our NM is also composed of many stacked residual blocks, but they have the same structure as EDSR [35] instead of the original one in [20]. Let $\mathcal{F}^l_{res}(\cdot)$ denote the function that corresponds to the lth residual block, then the output of a residual block is given by:

$$\mathbf{x}_l = \mathcal{F}^l_{res}(\mathbf{x}_{l-1}), \quad l = 1, 2, \ldots, L \tag{4}$$

where L is the total number of residual blocks. Each residual block has a *residual path* and a *shortcut connection* within it. Here, the residual path is composed of two conv layers and one ReLU operation, and the shortcut connection is a *identity mapping*, as shown in Fig. 4. Let $\mathcal{F}(\mathbf{x}_{l-1}, \mathcal{W}_l)$ represent the function corresponding to the residual path, where $\mathcal{W}_l = \{(\mathbf{w}^1_l,\ \mathbf{b}^1_l), (\mathbf{w}^2_l,\ \mathbf{b}^2_l)\}$ is the parameter set of the lth residual block (the superscripts are the indexes of the conv layers). Thus (4) can be rewritten as:

$$\mathbf{x}_l = \mathbf{x}_{l-1} + \mathcal{F}(\mathbf{x}_{l-1}, \mathcal{W}_l), \quad l = 1, 2, \ldots, L \tag{5}$$

$$\mathcal{F}(\mathbf{x}_{l-1}, \mathcal{W}_l) = \left[\max(\mathbf{x}_{l-1} * \mathbf{w}^1_l + \mathbf{b}^1_l,\ 0)\right] * \mathbf{w}^2_l + \mathbf{b}^2_l \tag{6}$$

Iteratively, we have (i.e. $\mathbf{x}_{l+1} = \mathbf{x}_l + \mathcal{F}(\mathbf{x}_l, \mathcal{W}_{l+1}) = \mathbf{x}_{l-1} + \mathcal{F}(\mathbf{x}_{l-1}, \mathcal{W}_l) + \mathcal{F}(\mathbf{x}_l, \mathcal{W}_{l+1}) = \cdots$):

$$\mathbf{x}_N = \mathbf{x}_n + \sum_{i=n}^{N-1} \mathcal{F}(\mathbf{x}_i, \mathcal{W}_{i+1}), \quad 0 \leqslant n < N \leqslant L \tag{7}$$

In a residual block, (7) always hold if there is no additional operation after the addition and the shortcut connection is a identical mapping [21]. It exhibits some very nice properties that are helpful for forward and backward propagation. This is usually beneficial for low-level computer vision problems. At the end of the NM part, there is another conv layer that is used to collect the final nonlinear features from all the preceding residual blocks:

$$\mathbf{x}_{L+1} = \mathbf{x}_L * \mathbf{w}_{L+1} + \mathbf{b}_{L+1} \tag{8}$$

Finally, a global skip connection (GSC) between NM's input $\mathbf{x}_0$ and $\mathbf{x}_{L+1}$ is used to further stabilize model training. The final output of the whole NM is as follows:

$$\mathbf{x}_{L+2} = \mathbf{x}_{L+1} + \mathbf{x}_0 \tag{9}$$

3.1.3 Image reconstruction

The artifact-degraded image $\mathbf{x}$ and the clean image $\mathbf{y}$ are of the same size in our settings (zero-padding). Therefore, a single conv layer is simply applied to produce the estimated image (see Fig. 3). The output of this final conv layer is:

$$\mathbf{x}_{L+3} = \mathbf{x}_{L+2} * \mathbf{w}_{L+3} + \mathbf{b}_{L+3} \tag{10}$$

Besides, we set an external skip connection (ESC) between the input of the model and $\mathbf{x}_{L+3}$, which means that the whole model learns the residual between $\mathbf{x}$ and $\mathbf{y}$. Thus, the final output of our network can be represented as:

$$\mathbf{y} = \mathbf{x}_{L+4} = \mathbf{x}_{L+3} + \mathbf{x} \tag{11}$$

It is worth noting that the model training becomes gradually unstable as the network depth increases. The ESC can also stabilize model training and improve the performance of the model slightly, just like constant scaling [35].

3.2 Model training

3.2.1 Optimization objective

Given a training dataset $\mathcal{D} = \left\{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\right\}_{i=1}^{T}$, where T is the total number of the dataset, the target of learning an end-to-end mapping f that predicts $\hat{\mathbf{y}} = f(\mathbf{x}; \boldsymbol{\theta})$ is to decide the model parameters $\boldsymbol{\theta}$, where $\hat{\mathbf{y}}$ is an estimate of the artifacts-free image $\mathbf{y}$. This is achieved through minimizing the L_1 loss between $\hat{\mathbf{y}}$ and $\mathbf{y}$:

$$L(\boldsymbol{\theta}) = \frac{1}{T}\sum_{i=1}^{T} \| f(\mathbf{x}^{(i)}; \boldsymbol{\theta}) - \mathbf{y}^{(i)} \|_1 \tag{12}$$

Despite that minimizing L_2 is generally preferred since it maximizes peak signal to noise ratio (PSNR), L_1 loss provides better convergence during model training [35]. This is especially helpful in the case of training the model with MR images due to the degradation of training samples [60].

3.2.2 Multilevel residual mechanism

In general, the artifact-degraded images and their clear counterparts share same information to a large extent. This indicates that a large part of the topological structures of their manifolds in high dimensional space are similar to each other. Therefore, it is more advantageous to explicitly learning the residual image, instead of the original image [25].

Because the structure of MR training samples is much simpler than that of natural images and the number of the model parameters is relatively large (about 2.7M), we use the multilevel residual mechanism to further stabilize the model training. As shown in Fig. 3, in addition to the stacked residual blocks and GSC, we also add an ESC between the outermost layers to enforce the model learn the residual between $\mathbf{x}$ and $\mathbf{y}$. Our experiments confirm the effect of this multilevel residual mechanism on the stabilization and the convergence of the model training.

3.2.3 Knowledge transfer

In this work, we not only study training the network directly using MR images, but also the effect of knowledge transfer from natural images to MR images. Concretely, given a natural image dataset $\mathcal{D}_N = \left\{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\right\}_{i=1}^{T_N}$ and a MR image dataset $\mathcal{D}_M = \left\{\mathbf{x}^{(j)}, \mathbf{y}^{(j)}\right\}_{j=1}^{T_M}$, we first train the model using $\mathcal{D}_N$ with respect to:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{T_N} \sum_{i=1}^{T_N} \| f\left(\mathbf{x}^{(i)}; \boldsymbol{\theta}\right) - \mathbf{y}^{(i)} \|_1, \ \boldsymbol{\theta}_0 \sim \mathcal{N}\left(\mu, \sigma^2\right) \tag{13}$$

where $\boldsymbol{\theta}_0$ is the initial value of model parameters, and $\hat{\boldsymbol{\theta}}$ denotes the final value obtained by training the model using natural image dataset $\mathcal{D}_N$. $\mathcal{N}(\mu, \sigma^2)$ is a normal distribution with mean μ and standard deviation σ. After the model converges, we retrain it using the second dataset $\mathcal{D}_M$ with model parameters initialized by $\hat{\boldsymbol{\theta}}$, that is,

$$\tilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{T_M} \sum_{j=1}^{T_M} \| f(\mathbf{x}^{(j)}; \boldsymbol{\theta}) - \mathbf{y}^{(j)} \|_1, \ \boldsymbol{\theta}_0 = \hat{\boldsymbol{\theta}} \tag{14}$$

where $\tilde{\boldsymbol{\theta}}$ represents the final model parameters. This process is termed as *knowledge transfer of parameters* [39] that implicitly assumes that individual models for related tasks (Gibbs-ringing artifact suppression over natural images and MR images here) should share some parameters or prior distributions of hyperparameters. We experimentally show that the assumption is actually effective.

4 Experiments

In this section, the generation of training datasets and the related model configuration are firstly introduced. Then, the impact of ESC on model training and performance is studied and knowledge transferred from natural images to MR images is evaluated in different training settings. Finally, the superiority of the proposed model is demonstrated by comparing with other conventional methods both quantitatively and qualitatively. We adopt PSNR and SSIM [52] as quantitative evaluation indexes.

4.1 Data generation

To boost the performance of the proposed model for Gibbs-ringing artifact suppression, we prepared two datasets with natural images and MR images respectively. The detailed information of the datasets is outlined in Table 1.

Table 1 The generated datasets used for Evaluation experiments. Note the validation sets are small for quick validation. $|\cdot|$ denotes the size of the corresponding dataset

	Training	Validation	Testing
MRI	$\mathcal{D}_M$	$\mathcal{V}_M$	$\mathcal{T}_M$
	$\lvert\mathcal{D}_M\rvert = 10535$	$\lvert\mathcal{V}_M\rvert = 10$	$\lvert\mathcal{T}_M\rvert = 2633$
NI	$\mathcal{D}_N$	$\mathcal{V}_N$	$\mathcal{T}_N$
	$\lvert\mathcal{D}_N\rvert = 600$	$\lvert\mathcal{V}_N\rvert = 10$	$\lvert\mathcal{T}_N\rvert = 175$

4.1.1 NI dataset

The training set of DIV2K [1] dataset is applied to generate NI dataset. To make the sampling range of training samples relatively concentrated, we exclude the images that have more than 2K pixels in both horizontal and vertical directions. This leads to 785 images left. These images are then converted to grayscale and resized to 512×512. Lastly, 600 images are used to construct training set, and 175 images are included in testing set and the rest 10 images are used for quick validation.

We follow the process shown in Fig. 2 to simulate the ringing artifacts for both natural and MR images. All the degraded images are produced by 4-fold truncation in k-space so that the image degradation is relatively serious. Therefore, the network recovers the artifact-free image from only $1/16$ of the total information.

4.1.2 MRI dataset

The IXI dataset[3] is used to construct MR dataset. It contains 581 T1, 578 T2 and 578 PD volumes. Firstly, we take the intersection of these subsets according to imaging subjects and obtain 577 volumes. By the default order, the first 400 volumes and the next 100 volumes are used for generating training set and testing set respectively, and the rest is used for producing validation data. To reduce the redundancy and preserve the example diversity, 25 slices at both ends of the slice-selection direction are discarded and one slice at the interlaced interval of 10 slices is extracted to generate 2D samples. Finally, we obtain 10535 training slices and 2633 testing slices. Only 10 randomly selected MR slices are used for quick validation.

4.2 Training setting

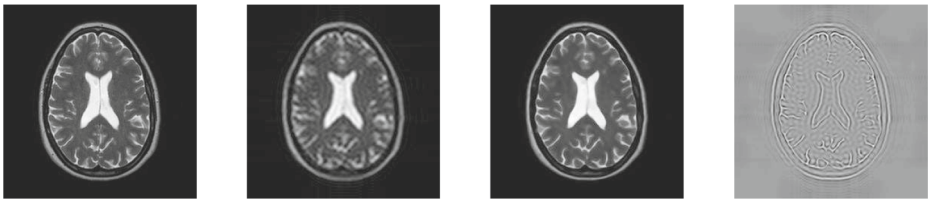
In our settings, the artifact-degraded image $\mathbf{x}$ is 4-fold truncated and zero-padded in k-space so that it has the same size as the output image $\hat{\mathbf{y}}$. All training images are split into sub images of size 48×48 before feeding into the model. This is achieved by random extraction of paired artifact-degraded and artifact-free image patches. The data augmentation is realized by random horizontal flips and 90° rotations.

The number of residual blocks and feature maps is set to 32 and 64 respectively, and kernel size is 3×3 for all conv layers. The proposed models are implemented and trained in TensorFlow 1.7.0, and the Adam optimizer [27] is used to minimize the loss function by setting $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-8}$. Batch size is set to 16. We use piecewise constant decay for learning rate, i.e., it is initialized as 10^{-4} for all layers and halved at every 2×10^5 iterations. All models are trained for one million iterations. The Xavier's method [16] is used to initialize network parameters when the model is trained from scratch.

4.3 External skip connection

We denote REF-CNN as the network with the same structure as GAS-CNN except that it does not have a ESC. Both REF-CNN and GAS-CNN are trained with $\mathcal{D}_N$ and $\mathcal{D}_M$. The training dataset is attached to model name to indicate which dataset is used to train the model, e.g., REF-CNN ($\mathcal{D}_N$) indicates the model without ESC is trained with $\mathcal{D}_N$ and GAS-CNN ($\mathcal{D}_M$) refers to the proposed model trained with $\mathcal{D}_M$.

[3] http://brain-development.org/ixi-dataset/



(a) Ground truth (b) Gibbs input (c) REF-CNN (d) GAS-CNN

Fig. 5 The external skip connection (ESC) enables the model to learn the residual $\mathbf{z}$ between the original input and the corresponding label, which is considered to have a simpler manifold structure in high-dimensional space [3]. However, the reference network (REF-CNN) directly learns to approximate the ground truth

4.3.1 What the models learn

Figure 5 presents the difference in what REF-CNN and GAS-CNN have learned, where Fig. 5d is normalized for visualization purpose). As can be seen, REF-CNN tends to learn $\mathbf{y}$ directly while GAS-CNN learns the residual of the ground truth and the input image $\mathbf{x}$, i.e., $\mathbf{y} - \mathbf{x}$. The residual $\mathbf{y} - \mathbf{x}$ has been considered to have a simpler manifold structure, thus relieving the representational burden of the model [3]. Therefore, manifold simplification is helpful to improve the performance of the model with certain representational capacity. This is still an open question that may explain why residual learning [20, 21] works.

4.3.2 Training efficiency and stability

Since ESC connects the input and output of the model directly, it also promotes the information flow in the network and thus helps to improve the training efficiency. This is similar to GSC and the constant scaling within a residual block [35]. Figure 7 shows the efficiency of training the models with $\mathcal{D}_M$. It can be observed that the training process of GAS-CNN is more stable and more efficient than that of REF-CNN. It is worth noting that the other configurations of these two models are the same except for the ESC, which means learning the residual $\mathbf{y} - \mathbf{x}$ really helps to improve the efficiency of model training. Actually, Zhao

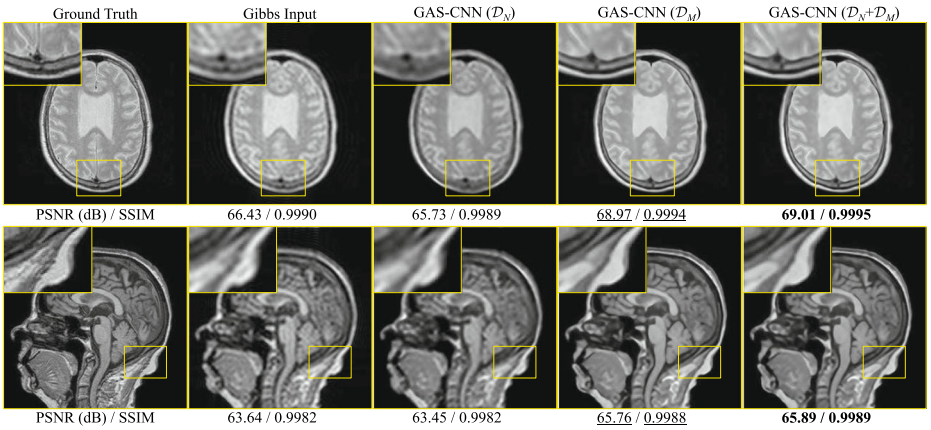


Fig. 6 The visual effect of GAS-CNN ($\mathcal{D}_N$), GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) on MR images. The maximal PSNR and SSIM of each row are in bold, and the second ones are underlined

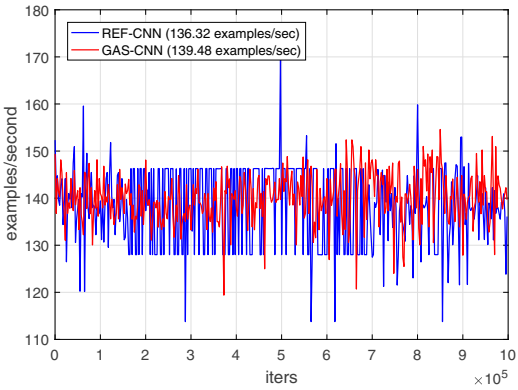


Fig. 7 Comparison of training efficiency between REF-CNN and GAS-CNN. Both models are evaluated on $\mathcal{D}_M$, and the averages are shown in parentheses. Note that the oscillation amplitude of GAS-CNN is smaller than that of REF-CNN

et al. [60] also validated the role of the ESC in stabilizing model training according to the validation performance of model training.

4.3.3 Quantitative comparison

Another advantage of the ESC is that it can improve the performance of the model. Figure 8 shows the validation results of REF-CNN ($\mathcal{D}_N$) and GAS-CNN ($\mathcal{D}_N$). It can be observed that GAS-CNN ($\mathcal{D}_N$) performs better than REF-CNN ($\mathcal{D}_N$) although both models have the same model parameters and network depth. We also trained the models with MR training set $\mathcal{D}_M$ and the quantitative evaluation results are shown in Table 2, from which we can observe that the performance of GAS-CNN is slightly better than that of REF-CNN on both $\mathcal{T}_N$ and $\mathcal{T}_M$. This further verifies the effectiveness of the ESC in improving model performance.

Note that we also tested the models with different testing images, i.e., trained with $\mathcal{D}_M$ but tested with $\mathcal{T}_N$, vice versa. The results show that the models trained on MR images have certain effect on natural images and the models trained on natural images also have some effect on MR images. This indicates that *there may exist information sharing between natural images and MR images, which may explain why knowledge transferred from natural images to medical or MR images can be effective*.

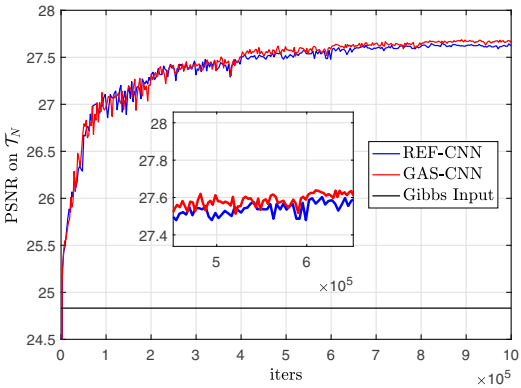


Fig. 8 The validation performance of REF-CNN and GAS-CNN on natural images. The evaluation of the input images is also plotted for reference (Gibbs Input)

Table 2 Quantitative comparison between REF-CNN and GAS-CNN on both $\mathcal{T}_N$ and $\mathcal{T}_M$. The maximum of each row is in bold, and the second one is underlined

		REF-CNN		GAS-CNN	
		$\mathcal{D}_N$	$\mathcal{D}_M$	$\mathcal{D}_N$	$\mathcal{D}_M$
$\mathcal{T}_N$	PSNR	<u>26.43</u>	24.95	**26.46**	25.27
	SSIM	<u>0.7590</u>	0.6970	**0.7596**	0.7108
$\mathcal{T}_M$	PSNR	64.08	<u>66.69</u>	64.14	**66.70**
	SSIM	0.9975	<u>0.9978</u>	0.9975	**0.9983**

4.4 Evaluation on knowledge transfer

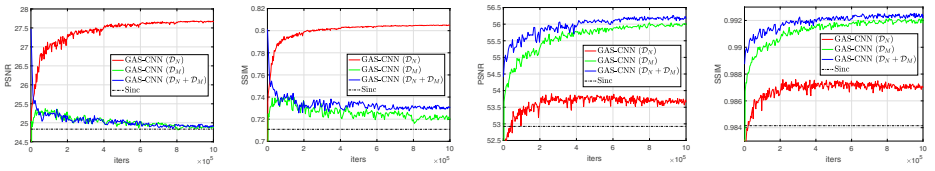
The impact of knowledge transfer from natural images to MR images on the model performance is investigated in this section. We train the GAS-CNN model with $\mathcal{D}_N$, $\mathcal{D}_M$ and $\mathcal{D}_N+\mathcal{D}_M$ respectively. Here $\mathcal{D}_N+\mathcal{D}_M$ represents the transfer learning version of the model, i.e., $\mathcal{D}_N$ is first employed to pre-train the model and the obtained parameters are used as the initial values of retraining with $\mathcal{D}_M$ (refer to Section 3.2.3). For convenience, both $\mathcal{V}_N$ and $\mathcal{V}_M$ are evaluated during quick validation.

4.4.1 Performance on Natural Images

Figure 9a and b show the performance of the compared models on $\mathcal{V}_N$. It is obvious that the model GAS-CNN ($\mathcal{D}_N$) gives the best performance. This is unsurprising because the model is trained and evaluated both with natural images. However, although GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) are mainly aimed at MR images, they do work for natural images to a certain extent. This can be visually verified according to Fig. 10, of which the last two columns display the results of GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) to process natural images. Compared with the original inputs, we can still observe artifact reduction in the results of these two models. This also indicates that the models share some information between natural images and MR images.

4.4.2 Performance on MR images

Figure 9c and d display the performance of the models on $\mathcal{V}_M$. Again, they illustrate the information sharing mechanism between the models trained with natural images and MR images since the information learned from natural images can be directly applied to suppress the Gibbs-ringing artifacts in MR images.



(a) PSNR on $\mathcal{V}_N$ (b) SSIM on $\mathcal{V}_N$ (c) PSNR on $\mathcal{V}_M$ (d) SSIM on $\mathcal{V}_M$

Fig. 9 The validation performance of the models. For all models, the validation results on both $\mathcal{V}_N$ and $\mathcal{V}_M$ are shown for comparison purposes. **a** and **b** are the PSNR and SSIM values on $\mathcal{V}_N$, while **c** and **d** are that on $\mathcal{V}_M$. Sinc indicates the artifact-degraded inputs

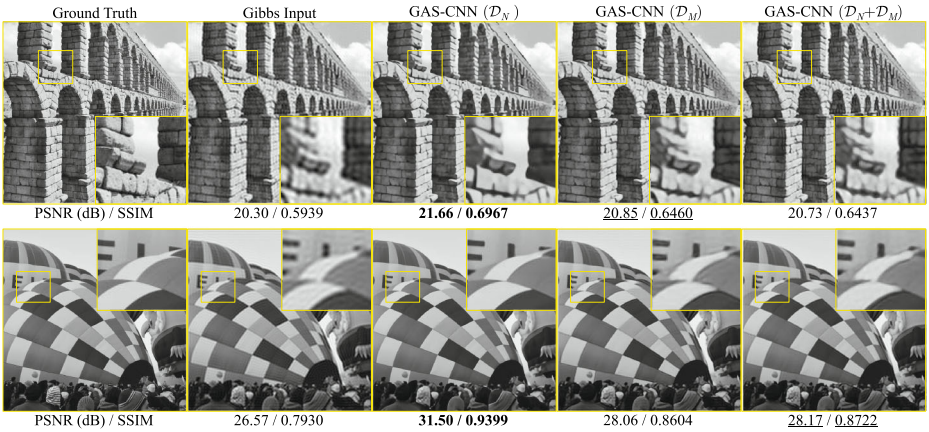


Fig. 10 The visual effect of GAS-CNN ($\mathcal{D}_N$), GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) on natural images. The maximal PSNR and SSIM of each row are in bold, and the second ones are underlined

This information sharing of the models trained with different datasets motivates us to use transfer learning to solve the problem of Gibbs-ringing artifact suppression. More important observation in Fig. 9c and d is that GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) significantly performs better than GAS-CNN ($\mathcal{D}_M$), which means that the knowledge transfer from natural images to the MR images is indeed capable of further boosting the performance of the model. In Fig. 6, we can also observe that knowledge transfer provides the best quantitative and qualitative results over both PD (top) and T1 (bottom) images.

4.5 Comparison with other methods

Filtering techniques are routinely adopted to suppress ringing artifacts in various MRI applications and commercially available MRI systems. According to our observation, bilateral filtering[4] gives better performance than other filtering methods in case of zero-padding. In addition, we also include NLM [37] and a recent method based on CNN [53]. Strictly speaking, NLM [37] is not a method specifically for Gibbs-ringing artifact suppression, but a typical MR image upsampling technique. However, it can improve the quality of upsampled images to some extent, so we use it as a comparison method to illustrate the effectiveness of the proposed method in improving image resolution. Note that the input images of NLM [37] is not zero-padded in k-space. Wang et al. [53] proposed a CNN-based method similar to SRCNN [12]. We reimplemented this model as a comparison method according to the configuration described in the paper, and termed it as Gibbs-ringing Artifact Reduction Convolutional Neural Network (GARCNN).

Table 3 collects the average PSNR and SSIM results of all the compared methods on both $\mathcal{T}_N$ and $\mathcal{T}_M$. GAS-CNN ($\mathcal{D}_N$) achieves the best performance on natural images, and GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) surpass other methods significantly in terms of MR images. This verifies the effectiveness of the proposed GAS-CNN model in dealing with the task of Gibbs-ringing artifact suppression. More importantly, GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)

[4] http://people.csail.mit.edu/jiawen/software/bilateralFilter.m

Table 3 Quantitative evaluation of the compared methods on testing sets $\mathcal{T}_N$ and $\mathcal{T}_M$. The maximal PSNR and SSIM of each row are in bold, and the second ones are underlined (4-fold truncation on both directions)

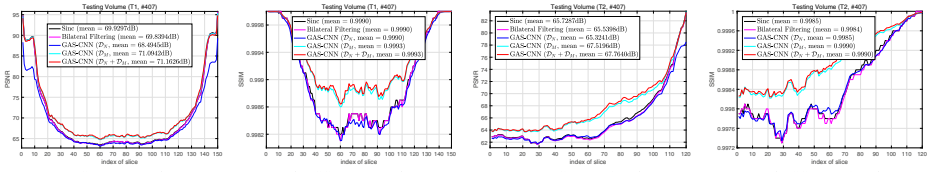
Datasets \ Methods		Sinc Filtering	Bilateral Filtering	NLM [37]	GARCNN [53]	GAS-CNN ($\mathcal{D}_N$)	GAS-CNN ($\mathcal{D}_M$)	GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)
$\mathcal{T}_N$	PSNR (dB)	25.20	25.09	25.19	<u>25.82</u>	**26.45**	25.27	25.34
	SSIM	0.7030	0.6877	0.7426	<u>0.7501</u>	**0.7596**	0.7108	0.7193
$\mathcal{T}_M$	PSNR (dB)	64.25	64.21	64.14	64.73	64.14	<u>66.70</u>	**66.94**
	SSIM	0.9971	0.9971	0.9972	0.9980	0.9974	<u>0.9983</u>	**0.9985**

presents better results than GAS-CNN ($\mathcal{D}_M$), implying that the knowledge learned from natural images is actually beneficial to improving the performance of deep models.

We also compared several methods on 3D MR volumes. Figure 11 exhibits the results of these methods on T1 and T2 volumes, from which we can observe that GAS-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$) surpass other methods over all slices. As shown in the plots, the advantage of our models on the middle slices is more obvious than that on both ends. This is because the image structure is more complex in the middle slices. Figures 12 and 13 show the visual comparison between the compared methods on PD, T1 and T2 images. We can see that bilateral filtering and GARCNN [53] can eliminate Gibbs-ringing artifacts in the input image to some extent, but the overall effect is unsatisfactory. On the contrary, the proposed methods can effectively eliminate the artifacts and improve the image quality, e.g., the positions pointed by the red arrows in Figs. 12 and 13.

4.6 Low frequency replacement

In our settings, the artifact-degraded image $\mathbf{x}$ is zero-padded in k-space to match the size of the artifact-free image $\mathbf{y}$. In this case, a reasonable assumption is to keep the low frequency information of the input image unchanged (i.e., keep $\mathcal{X}$ in (2) fixed). This can be achieved by simply replacing the low frequency part of the resulting image into the corresponding frequency components of the original input image. We term it as *low frequency replacement* (LFR). Figure 14 shows that LFR can improve the performance of bilateral filtering and the models trained and evaluated with different datasets. However, it is harmful to the well-trained models (black rectangles in Fig. 14). This may be because that LFR destroys the continuity of the learned k-space data and leads to the reappearance of these artifacts. However, as a priori information, LFR is beneficial to the models with relatively poor representational capacity, e.g., GARCNN [53].



(a) PSNR, T1 volume (b) SSIM, T1 volume (c) PSNR, T2 volume (d) SSIM, T2 volume

Fig. 11 The quantitative evaluation of the several methods on 3D volumes. As can be observed, the advantage of our method is more obvious on slices in the middle. The results of sinc filtering (inputs to the models) are also plotted for reference

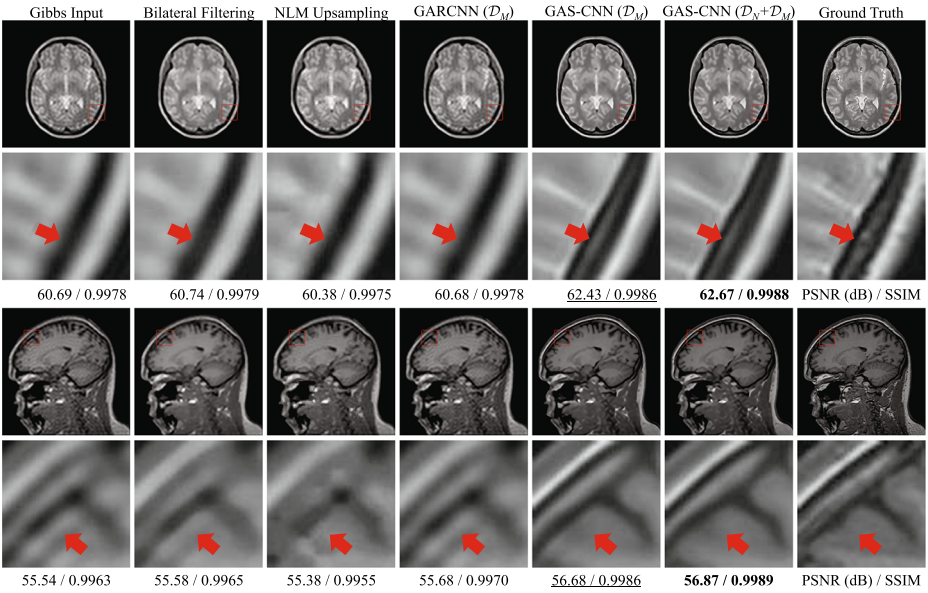


Fig. 12 The visual effects of the compared methods (top: T1 image, bottom: T2 image). Gibbs input corresponds to sinc filtering. Note that 4-fold truncation in k-space makes the image degradation very serious (Gibbs input vs. Ground truth). It can be seen that our methods can eliminate Gibbs-ringing artifacts and improve the spatial resolution of the image

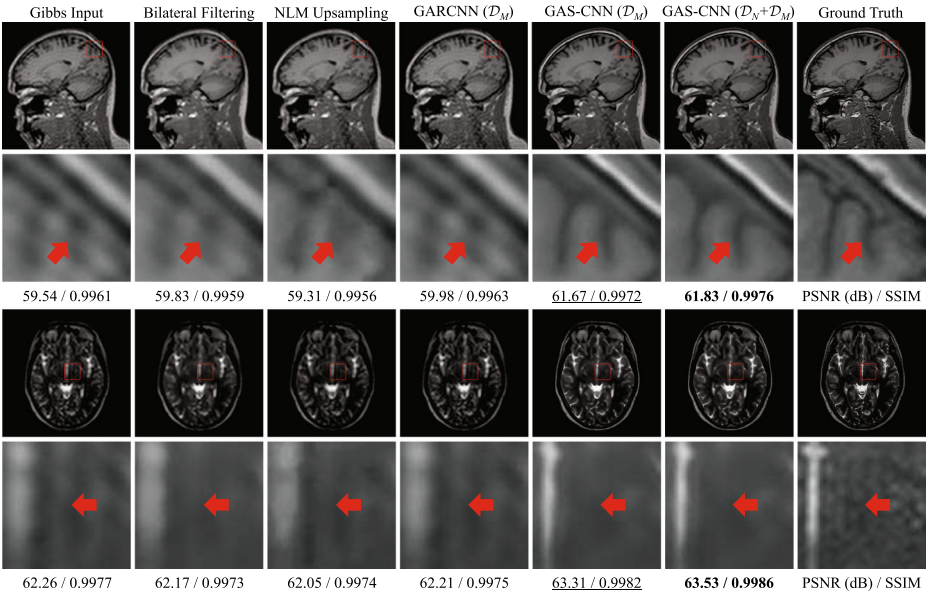
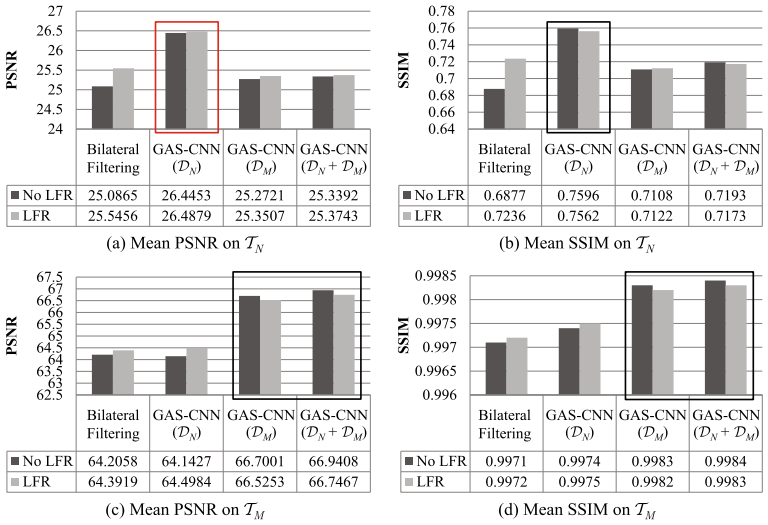


Fig. 13 The visual effects of the compared methods (top: PD image, bottom: T1 image). The proposed methods can eliminate Gibbs-ringing artifacts (bottom row) and improve the spatial resolution of the image (top row)



	Bilateral Filtering	GAS-CNN ($\mathcal{D}_N$)	GAS-CNN ($\mathcal{D}_M$)	GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)
No LFR	25.0865	26.4453	25.2721	25.3392
LFR	25.5456	26.4879	25.3507	25.3743

(a) Mean PSNR on $\mathcal{T}_N$

	Bilateral Filtering	GAS-CNN ($\mathcal{D}_N$)	GAS-CNN ($\mathcal{D}_M$)	GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)
No LFR	0.6877	0.7596	0.7108	0.7193
LFR	0.7236	0.7562	0.7122	0.7173

(b) Mean SSIM on $\mathcal{T}_N$

	Bilateral Filtering	GAS-CNN ($\mathcal{D}_N$)	GAS-CNN ($\mathcal{D}_M$)	GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)
No LFR	64.2058	64.1427	66.7001	66.9408
LFR	64.3919	64.4984	66.5253	66.7467

(c) Mean PSNR on $\mathcal{T}_M$

	Bilateral Filtering	GAS-CNN ($\mathcal{D}_N$)	GAS-CNN ($\mathcal{D}_M$)	GAS-CNN ($\mathcal{D}_N + \mathcal{D}_M$)
No LFR	0.9971	0.9974	0.9983	0.9984
LFR	0.9972	0.9975	0.9982	0.9983

(d) Mean SSIM on $\mathcal{T}_M$

Fig. 14 The influence of low frequency replacement (LFR) on the compared methods. The red rectangle shows the improvement of model performance, while the black rectangles indicate the degradation of model performance

4.7 Implementation efficiency

The average execution time of bilateral filtering, GARCNN ($\mathcal{D}_M$) [53], REF-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_M$) on $\mathcal{T}_M$ is evaluated in this section. We do not include NLM [37] here because it is an iterative optimization-based method that takes a lot of time ($>$ 1000 ms) to process a 128 $\times$ 128 image. Since the image size in $\mathcal{T}_M$ is 256 $\times$ 256 and we only focus on running time here, we record the results of 128 $\times$ 128 and 512 $\times$ 512 by simply resizing the images in $\mathcal{T}_M$ before feeding them into the model. The results are calculated on a Omnisky supercomputer workstation equipped with 2 Intel (R) Xeon (R) CPUs (E5-2630 v4, 2.2GHz) and 64GB memory. Table 4 presents the comparison between these methods (ms / image). As can be seen, bilateral filtering runs faster than deep CNN-based methods. This is because deep models have a large number of model parameters, requiring a lot of computation resources. In addition, Table 4 also indicates that the efficiency of REF-CNN ($\mathcal{D}_M$) and GAS-CNN ($\mathcal{D}_M$) is approximately equivalent.

5 Discussion and future work

5.1 Difference with other methods

The network used in this paper is a typical image generation model which is derived from the EDSR model [35]. Therefore, some of the previous models for image generation are similar to it, e.g., SRCNN [12], FSRCNN [13], DRCN [26], DRRN [49], VDSR [25], ESPCNN [44], SRDenseNet [50], EDSR/MDSR [35] and RDN [59] etc. Nevertheless, we also summarize the differences between the GAS-CNN and the EDSR model:

Table 4 Execution time comparison between bilateral filtering and the proposed method (ms / image)

Image Size	128 × 128	256 × 256	512 × 512
Bilateral Filtering	17.9	22.8	44.1
GARCNN ($\mathcal{D}_M$)	49.2	78.5	97.8
REF-CNN ($\mathcal{D}_M$)	71.8	110.8	272.9
GAS-CNN ($\mathcal{D}_M$)	72.3	109.3	273.8

- The input and output images of the models have different sizes. Image super-resolution usually requires to change the size of the original image, but the GAS-CNN model aims at recover a artifact-free image from an input image with the same size (zero-padding).
- The ESC between the outermost layers of the model enables it to learn the residual between **x** and **y**, which does not exist in the original EDSR model.
- The GAS-CNN is mainly used to study the impact of knowledge transfer from natural images to MR images on the problem of Gibbs-ringing artifact suppression. However, the EDSR is especially designed to solve the task of single image super-resolution (originally for natural images)

5.2 Calculation of evaluation indexes

We utilized the built-in functions of MATLAB (2016b) to calculate the values of PSNR and SSIM. Due to different data types,[5] the PSNR peaks and the SSIM dynamic ranges are set to the maximum of the corresponding data type. In doing this, we expect that they keep consistent across different examples. But this leads to the relatively high PSNR and SSIM values for MR images. However, the relative relationship among the results and the conclusion will not change with different dynamic ranges.

5.3 Image degradation

The present work only studies the case of 4-fold downsampling in both directions. However, there may be cases where image degradation is less severe in practice. For example, 2-fold downsampling in both directions or just in phase coding direction. In these cases, Gibbs-ringing artifacts may not be as obvious as in this work. However, the main purpose of this paper is to verify the effectiveness of knowledge transfer from natural images to MR images in the regression problem of dense predication. Besides, the GAS-CNN is also believed to be effective and capable of improving image quality in these cases. Investigating the different downsampling scales, directions and degradations of Gibbs-ringing artifact suppression and further improving the performance of the model will be a part of our future work.

[5]Usually, natural images are stored as uint8 (e.g., png, bmp, jpg etc.), while MR images are stored as int16 (e.g., dicom, nifti, etc.) with a much larger dynamic range.

6 Conclusion

In this work, a novel method based on CNN and transfer learning is presented to suppress Gibbs-ringing artifacts in MR images. The network used in this work is derived from the EDSR model [35] with several adjustments. Several techniques have been applied to stabilize model training, including constant scaling, L_1 loss, and multilevel residual mechanism. Knowledge transfer from natural images to MR image is adopted to further improve the model performance. Experimental results show that the proposed method outperforms other methods by a large margin both quantificationally and qualitatively. Moreover, we have experimentally verified that transfer learning can indeed boost the performance of deep regression models.

An important feature of the proposed method is that it formalizes Gibbs-ringing artifact suppression as a problem of image generation akin to image super-resolution and/or inpainting, thus it is capable of reducing the Gibbs-ringing artifacts and increasing the spatial resolution of the images simultaneously, e.g., Figs. 12 and 13. This feature also helps to improve image quality and visualization.

Acknowledgements The work is supported in part by the National Key Research and Development Program of China (No. 2016YFC0100 800 and 2016YFC0100802).

References

1. Agustsson E, Timofte R (2017) NTIRE 2017 challenge on single image super-resolution: dataset and study. In: 2017 IEEE conference on computer vision and pattern recognition workshops, pp 1122–1131
2. Archibald R, Gelb A (2002) A method to reduce the gibbs ringing artifact in mri scans while keeping tissue boundary integrity. IEEE Trans Med Imaging 21(4):305–319
3. Bae W, Yoo J, Ye JC (2017) Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification. arXiv:1611.06345 [cs.CV]
4. Bertalmio M, Vese L, Sapiro G, Osher S (2003) Simultaneous structure and texture image inpainting. IEEE Trans Image Process 12(8):882–889
5. Bertalmio M, Sapiro G, Caselles V, Ballester C (2005) Image inpainting. Siggraph 4(9):417–424
6. Block KT, Uecker M, Frahm J (2008) Suppression of mri truncation artifacts using total variation constrained data extrapolation. Int J Biomed Imaging 2008(2):184123
7. Boyd JP (2005) Trouble with gegenbauer reconstruction for defeating gibbs' phenomenon: Runge phenomenon in the diagonal limit of gegenbauer polynomial approximations. J Comput Phys 204(1):253–264
8. Carroll J, Carlson N, Kenyon GT (2017) Phase transitions in image denoising via sparsely coding convolutional neural networks. arXiv:1710.09875 [cs.NE]
9. Constable RT, Henkelman RM (1991) Data extrapolation for truncation artifact removal. Magn Reson Med 17(1):108–118
10. Ding Z, Fu Y (2017) Robust transfer metric learning for image classification. IEEE Trans Image Process 26(2):660–670
11. Ding Z, Shao M, Fu Y (2018) Incomplete multisource transfer learning. IEEE Trans Neural Netw Learn Syst 29(2):310–323
12. Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell 38(2):295–307
13. Dong C, Loy CC, Tang X (2016) Accelerating the super-resolution convolutional neural network. In: Computer vision – ECCV 2016, pp 391–407. Springer International Publishing
14. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HMT (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118
15. Glasner D, Bagon S, Irani M (2009) Super-resolution from a single image. In: IEEE international conference on computer vision, pp 349–356
16. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the 13th international conference on artificial intelligence and statistics AISTATS, pp 249–256
17. Gottlieb D, Shu CW (1997) On the gibbs phenomenon and its resolution. Siam Rev 39(4):644–668

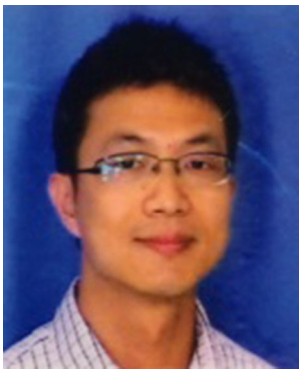
18. Gottlieb D, Shu CW, Solomonoff A, Vandeven H (1992) On the gibbs phenomenon i: recovering exponential accuracy from the fourier partial sum of a nonperiodic analytic function. J Comput Appl Math 43(1-2):81–98
19. Gulshan V, Peng L, Coram MT (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402–2410
20. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
21. He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. In: European conference on computer vision (ECCV), pp 630–645
22. Hu Y, Gao X, Li J, Huang Y, Wang H (2018) Single image super-resolution via cascaded multi-scale cross network. arXiv:1802.08808 [cs.CV]
23. Jain V, Seung HS (2008) Natural image denoising with convolutional networks. In: Proceedings of the 21st international conference on neural information processing systems, pp 769–776
24. Kellner E, Dhital B, Kiselev VG, Reisert M (2016) Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn Reson Med 76(5):1574–1581
25. Kim J, Lee JK, Lee KM (2016) Accurate image super-resolution using very deep convolutional networks. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 1646–1654
26. Kim J, Lee JK, Lee KM (2016) Deeply-recursive convolutional network for image super-resolution. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 1637–1645
27. Kingma DP, Ba JL (2014) Adam: A method for stochastic optimization. arXiv:1412.6980v9 [cs.LG]
28. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: International conference on neural information processing systems, pp 1097–1105
29. Latt JL, Morrison A, Radgowski A, Tobin J, Viswanathan A (2016) Technical report: improved fourier reconstruction using jump information with applications to mri. arXiv:1610.03764v1 [cs.NA]
30. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
31. Lecun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1(4):541–551
32. Li J, Lu K, Huang Z, Zhu L, Shen HT (2018) Heterogeneous domain adaptation through progressive alignment. IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2018.2868854
33. Li J, Lu K, Huang Z, Zhu L, Shen HT (2018) Transfer independently together: a generalized framework for domain adaptation. IEEE Transactions on Cybernetics. https://doi.org/10.1109/TCYB.2018.2820174
34. Liang ZP, Boada FE, Constable RT, Haacke EM (1992) Constrained reconstruction methods in mr imaging. Reviews of Magnetic Resonance in Medicine 4(2):67–185
35. Lim B, Son S, Kim H, Nah S, Lee KM (2017) Enhanced deep residual networks for single image super-resolution. In: 2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW), pp 1132–1140
36. Litjens G, Kooi T, Bejnordi BE, Aaa S, Ciompi F, Ghafoorian M (2017) A survey on deep learning in medical image analysis. Med Image Anal 42(9):60–88
37. Manjon JV, Coupe P, Buades A, Fonov V, Collins DL, Robles M (2010) Non-local mri upsampling. Med Image Anal 14(6):784–792
38. Ouyang W, Luo P, Zeng X et al (2014) Deepid-net: multi-stage and deformable deep convolutional neural networks for object detection. arXiv:1409.3505
39. Pan SJ, Yang Q (2010) A survey on transfer learning. IEEE Trans Knowl Data Eng 22(10):1345–1359
40. Perrone D, Aelterman J, Pižurica A, Jeurissen B, Philips W, Leemans A (2015) The effect of gibbs ringing artifacts on measures derived from diffusion MRI. Neuroimage 120:441–455
41. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252
42. Sarra SA (2006) Digital total variation filtering as postprocessing for chebyshev pseudospectral methods for conservation laws. Numerical Algorithms 41(1):17–33
43. Shen D, Wu G, Suk HI (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19(1):221–248
44. Shi W, Caballero J, Huszar F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z (2016) Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 1874–1883
45. Sun Y, Chen Y, Wang X, Tang X (2014) Deep learning face representation by joint identification-verification. In: Advances in neural information processing systems, pp 1988–1996
46. Sun Y, Liang D, Wang X, Tang X (2015) DeepID3: face recognition with very deep neural networks. arXiv:1502.00873v1
47. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR), pp 1–9

48. Szegedy C, Reed S, Erhan D, Anguelov D (2014) Scalable, highquality object detection. arXiv:1412.1441
49. Tai Y, Yang J, Liu X (2017) Image super-resolution via deep recursive residual network. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 2790–2798
50. Tong T, Li G, Liu X, Gao Q (2018) Image super-resolution using dense skip connections. In: 2017 IEEE international conference on computer vision (ICCV), pp 4809–4817
51. Veraart J, Fieremans E, Jelescu IO, Knoll F, Novikov DS (2016) Gibbs ringing in diffusion mri. Magn Reson Med 76(1):301–314
52. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
53. Wang Y, Song Y, Xie H, Li W, Hu B, Yang G (2017) Reduction of gibbs artifacts in magnetic resonance imaging based on convolutional neural network. In: International congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI), pp 1–5
54. Wang S, Ding Z, Fu Y (2018) Cross-generation kinship verification with sparse discriminative metric. IEEE Trans Pattern Analysis Mach Int PP(99):1–1
55. Wang T, Sun M, Hu K (2018) Dilated deep residual network for image denoising. arXiv:1708.05473 [cs.CV]
56. Yan H, Mao J (1993) Data truncation artifact reduction in mr imaging using a multilayer neural network. IEEE Trans Med Imaging 12(1):73–77
57. Zeiler MD, Fergus R (2013) Visualizing and understanding convolutional networks. In: European conference on computer vision, pp 818–833
58. Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y (2018) Image super-resolution using very deep residual channel attention networks. arXiv:1807.02758
59. Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y (2018) Residual dense network for image super-resolution. arXiv:1802.08797
60. Zhao X, Zhang Y, Zhang T, Zou X (2018) Channel splitting network for single mr image super-resolution. arXiv:1810.06453
61. Zhu L, Huang Z, Li Z, Xie L, Shen HT (2018) Exploring auxiliary context: discrete semantic transfer hashing for scalable image retrieval. IEEE Trans Neural Netw Learn Syst 29(11):5264–5276
62. Zhu L, Huang Z, Liu X, He X, Sun J, Zhou X (2017) Discrete multi-modal hashing with canonical views for robust mobile landmark search. IEEE Trans Multimed 19(9):2066–2079
63. Zhu L, Shen J, Xie L, Cheng Z (2017) Unsupervised visual hashing with semantic assistant for content-based image retrieval. IEEE Trans Knowl Data Eng 29(2):472–486

Xiaole Zhao received the B.S. and the M.S. degrees from the School of Computer Science and Technology, Southwest University of Science and Technology (SWUST), Mianyang, China, in 2013 and 2016 respectively. He is currently working toward the Ph.D. degree in the High Field Magnetic Resonance Brain Imaging Laboratory of Sichuan and the School of Life Science and Technology, the University of Electronic Science and Technology of China (UESTC), Chengdu, China. His research interests include computer vision and image processing, machine learning, medical image analysis and deep learning techniques.

Huali Zhang received the B.S. degree in engineering from the School of Biomedical Engineering, South-Central University for Nationalities (SCUEC), Wuhan, China, in 2016. She is currently a graduate student in the High Field Magnetic Resonance Brain Imaging Laboratory of Sichuan, School of Life Science and Technology, University of Electronic Science and Technology of China (UESTC), Chengdu, China. Her research interests include computer vision, image localization and deep learning in medical image applications.

Yuliang Zhou received the B.S. degrees from the School of Biomedical Engineering, the Kunming University of Science and Technology (KMUST), Kunming, China, in 2013. M.S. He is now working towards M.S. candidate with the Neural Information Group of the School of Life Science and Technology, the University of Electronic Science and Technology of China (UESTC), Chengdu, China. His research interests mainly include magnetic resonance imaging processing and reconstruction by machine learning and deep learning algorithms.



Wei Bian received his B.S. degree in computer engineering from Chengdu University of Technology and his Ph.D. degree in Bioengineering from University of California Berkeley and University of San Francisco. He is currently an adjunct associate professor of the University of Electronic Science and Technology of China and the chief scientist in the advanced clinical applications of the Alltech Medicine. His research interests include image acquisition, reconstruction, post-processing and clinical applications of magnetic resonance imaging.

Tao Zhang received the BS degree in 1996 and MS degree in 1999 from Shanghai Jiao Tong University, Shanghai, China, and the Ph. D in Mechanical Engineering from Florida State University in 2004. He is currently a professor in the School of Life Science and Technology of the University of Electronic Science and Technology of China. From 2005 to 2012, he worked as a senior scientist at the General Electric (GE) Global Research Center. From 2012, he joined and worked as the Chief Technology Officer and Vice President of Alltech Medical System in China. Dr. Zhang's research interests include the magnetic resonance imaging (MRI) system technologies and its clinical applications, MRI system hardware development, and the applications of MRI technique in neuroscience and brain research. He has published over 30 technical papers and has been awarded over 20 patents.

Xueming Zou received the MS degree in Physics from the Oregon State University (OSU), Corvallis, Oregon, USA in 1986 and the Ph.D. degree in Bio-Medical Physics from the Massachusetts Institute of Technology, Cambridge, in 1990. He is a professor in the School of Life Science and Technology of the University of Electronic Science and Technology of China (UESTC), and the group leader in the High-field Magnetic Resonance Brain Imaging Key Laboratory of Sichuan Province. He worked as the Founder, President and CEO of USA Instruments which scratch to number one in global market share in MRI Coils, from 1993 to 2002. He also served as the Global Vice President and General Manager for Greater China Region of MR Business Division of GE. In 2005, Dr. Zou founded the AllTech Medical Systems Co. LTD, a China Corporation operating in Chengdu and worked as the president. His research interests include MRI system physics, advanced medical diagnostic imaging applications, and big data in medical imaging.

Affiliations

Xiaole Zhao[1] · Huali Zhang[1] · Yuliang Zhou[1] · Wei Bian[1] · Tao Zhang[1,2,3] · Xueming Zou[1,2,3]

Huali Zhang
uestczhanghuali@163.com

Yuliang Zhou
zhouyuliang1225@163.com

Wei Bian
wei.bian@alltechmed.com

Tao Zhang
taozhangjin@gmail.com

Xueming Zou
mark.zou@alltechmed.com

1 School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, 611731, China

2 High Field Magnetic Resonance Brain Imaging Lab of Sichuan, Chengdu, 611731, China

3 China Key Lab for Neuro Information of Ministry of Education, Chengdu, 611731, China