Medical Image Denoising in MRI Reconstruction Procedure

Dong Han(✉) and Ronny Velastegui

Norwegian University of Science and Technology, Gjøvik, Norway
dongha@stud.ntnu.no

Abstract. With the rapid development of computer technology, deep learning can be used in nearly every field, and it always has the potential to achieve a high efficiency performance. Specifically, in the field of medical images, it makes doctors possible to distinguish and diagnose diseases in a more accurate way. Medical images like any other form of imaging techniques are affected by noise and artifacts. There are many types of noise, such as quantum, random, electric, and gaussian noise, etc. The presence of noise affects image clarity and may obstruct the recognition and analysis of diseases. The traditional image denoising method has much more limitation when came to medical images, and the results cannot meet some specific medical image standards. Hence, denoising of medical images in deep learning can be an important technique for further medical image processing. In this work, we conducted a deep learning method, which is a sparse dilated convolution neural network based on compressed sensing, for medical image denoising.

Keywords: Medical image denoising · Medical resonance imaging · Compressed sensing · Dilated convolution · Deep learning

1 Introduction

The digital images are essential in the medical field in which it has been employed to anatomize human body. These medical images can be utilized to recognize different diseases. Unfortunately, medical images require high resolution, and the presence of noise is not beneficial to disease diagnosis and sometimes even leads to the possibility of misdiagnosis [2]. Hence, the need for effective image noise reduction technology is urgent. Researchers have noticed this problem and provided many examples and principles used in the process of noise reduction in medical images [26].

Computed tomography (CT) is to use X-rays to scan the human body layer by layer, and then according to the difference in absorption of X-rays by various tissues in the human body. The attenuation coefficient of X-rays in the human body is determined by using a certain mathematical method. After a specific computer processing, the two-dimensional distribution matrix of the attenuation coefficient values in the human body layer is obtained and transformed into the

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12950, pp. 115–130, 2021.
https://doi.org/10.1007/978-3-030-86960-1_9

gray-scale distribution of the image. This is a procedure to realize the modern medical imaging technology of building tomographic images [1]. There are mainly two types of noise in the CT image, quantum noise, and electronic noise [4].

Medical resonance imaging (MRI) is an imaging technology based on a nuclear physical phenomenon, which uses the signal generated by the resonance of the nucleus in a strong magnetic field to be reconstructed from the image [32]. Medical resonance (MR) diagnosis has been widely applied in clinics due to its advantages of high resolution, non-invasiveness, and non-radiation [8]. With the help of MRI, doctors more easily specify the difference between various types of tissues judged by these magnetic properties. Impulse and gaussian noise are the two main noise types in MRI. Most noise in MRI is introduced from the electronic circuits or coils and radio frequency according to the movement of the ions inside the human body [25]. Noise in MRI can cause random fluctuations and result in a low level of contrast because of the inclination in signals. This interferes with accurate subjective and numerical assessment and feature detection of MR images. In most medical application scenarios, MRI has a high signal-to-noise ratio, leading to gaussian noise [10].

Compressed sensing (CS), is a very efficient way to reconstruct the fast MRI and positively affects MRI denoising [18,19,28]. For the most traditional CS-MRI approaches, the CS-MRI model normally has two elements: the data fidelity part and the regularization part. The first one is always in the k-space form and the second is associated with sparsifying operation to avoid the overfitting. However, this conventional model can deal with the problem properly but there are some missing detailed information and noise-like artifact occurs during in the reconstruction procedure, especially when it comes up to the high-rate undersampled measurements [24]. Many related research works have been managed in order to alleviate this kind of issue and most studies can be classified as two parts. The one focus on introducing more accurate spasifying transforms or use the non-local means. For instance, the singular value decomposition (SVD) [12] and blind compressive sensing (BCS) [21], which are both based on basic discrete cosine transform (DCT), patch-based directional wavelets based on bandelet transform [29], L1 norm method in data fidelity constraints [23], sparsifying dictionary for removing the aliasing [31], and non-local total variation regularization for noise elimination [20]. Nonetheless, the approaches mentioned above have good performance in improving the accuracy of the reconstructed MR image but with high computational cost or requirement for stacking considerable non-local operators [30]. Another part is about the feature restoration and refinement which uses the augmented lagrangian (AL) based on SENSitivity Encoding (SENSE) [31]. However, although this strategy reconstructs the most detailed fine structure, it still brings some noise to the output MR image.

Deep neural networks evolve very fast and have achieved pleasant performance in many research areas because they have efficient learning abilities for different datasets. Analyzing unstructured data and the ability to convey high-quality results are the principal superiorities of deep neural networks. Nevertheless, the normal data-driven network requires huge datasets, and it is difficult to

acquire huge medical data because of the sensitivity of privacy. The deep learning method is one of the approaches to alleviate this issue by training the network with natural images to reconstruct the medical image [39]. Another method to mitigate this difficulty is the restoration of medical images through the model-based with a combination of deep learning networks [17]. As the deep learning method is developing in various areas, its utility in medical imaging denoising is confirmed, and their neural network has many varieties to choose from. Image denoising using deep learning techniques plays a vital part in many application areas of medical imaging such as CT and MRI images.

A convolution neural network (CNN) is a special artificial neural network, which is designed to preserve the spatial relationship between data, with few connections between each layer. The traditional CNN includes four main parts: convolutional layer, activation layer, pooling layer, and fully connected layer. CNN can form an efficient representation of input data. The input data in CNN are stack in an array and then flowed through layers that retain these correlations. Different parameters in each layer can achieve optimization in the end. Like a standard artificial neural network, it uses backpropagation and gradient descent for training [22].

In specific, deep learning has good performance in solving the noise artifact that cannot be mitigated purely in the CS model as mentioned previously. In this paper inspired by classical CNN, we mainly introduce a dilated convolution network with batch normalization based on variable splitting for medical MR image denoising. Two metrics are used for evaluating image quality: the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM). We combine the model-based approach with the deep learning method, and the sparse dilated network (SD-Net) is introduced to improve denoising performance during the training procedure of the variable splitting-based CS model. The inconsistency of network between training and testing performance is analyzed, the potential reasons are given. Finally, the SD-Net-H network based on hybrid dilated convolution (HDC) is proposed to solve the grid problem in SD-Net and better denoising performance is achieved.

2 Preliminaries

2.1 Compressed Sensing

In CS, sensing describes that in order to express and restore a certain signal or object, a certain sensing method is adopted for it (the sensing here includes the imaging, the sampling of the continuous signal, etc.) to obtain this measurement under this sensing modality so that the subsequent signal can be reconstructed and analyzed or for other applications [7]. There are mainly three types of sensing, full-sampling, over-sampling, and under-sampling. Without signal priors, we record the minimum required sensing sample dimension as the critical sampling rate. Then you can also choose a measurement with less than the critical rate. In this case, we call it under-sampling. This will compress the full sensing, that is, compressed [5]. According to the classic signal processing theory, the target

image cannot be restored uniquely and perfectly. This type of problem is called the ill-posed problem, and CS is used to solve ill-posed inverse problems. CS is able to recover the target image perfectly in the case of under-sampling with an effective signal prior [23].

2.2 Batch Normalization

Batch normalization (BN) was suggested by Google researchers [13] and was also implemented in GoogLeNet [33]. The BN algorithm accelerates the training process to a large extent and relaxes the conditions for network initialization. The BN can be used without Dropout, and in the meantime, the recognition effect of the network can be improved to a certain extent. There are widespread applications in new networks ResNet [3,27].

In the neural network training process, input sample features are generally normalized so that the data becomes a distribution with a mean of 0 and a standard deviation of 1, or a distribution with a range of 0 to 1. The sample feature distribution is scattered, if data is not normalized, it may cause the neural network to learn slowly or even difficult to learn [9]. BN can be seen as a component in the neural network that normally functions with an activation function [15]. The algorithm of BN is as follows:

The range of input values x in a mini-batch is $B = \{x_i ... x_m\}$; the parameters to be learned are γ, β and final output value is y_i.

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{1}$$

$$\sigma^2{}_B = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2 \tag{2}$$

$$\hat{x_i} = \frac{x_i - \mu_B}{\sqrt{\sigma^2{}_B + \epsilon}} \tag{3}$$

$$y_i = \gamma \hat{x_i} + \beta \tag{4}$$

First, the mean μ_B and the variance $\sigma^2{}_B$ of the small size of training data are calculated, then used to normalize the batch of training data to obtain a 0–1 distribution. The ϵ is a tiny positive number in the normalization step $\hat{x_i}$ to exclude the situation when the number is divided by zero. Finally, do the scale transformation and offset by y_i: adjust x_i by multiply by γ, and add β to increase the offset to get y_i, where γ is the scale factor and β is the translation factor. This step is the essence of BN, the normalized x_i will be restricted to a normal distribution, which will reduce the learning ability of the network. To solve this problem, two new parameters y_i, β are introduced and learned by the network during training. The role of BN in deep neural networks is undeniable: if neural network training encounters slow convergence, or "gradient explosion" and other untrainable situations occur, BN is a good way to solve those problems. Simultaneously, BN can also be added to accelerate model training and even improve model accuracy [14].

2.3 Dilated Convolution

The receptive field indicates the input area which can be "sensed" by the neuron in the neural network. In CNN, the calculation of an element on the feature map is affected by a certain area on the input image. This area is the receptive field of the element. The neurons in deeper layers can see the larger input area. The larger receptive field can contain more detailed information of the input image which is detected. Normally, increasing the size of the receptive field can collect much more detailed information. There mainly two ways to get the wide receptive field. One is increasing depth, and another is increasing width of CNN. In the first method, it is time-consuming and difficult to learn for networks. The second approach might introduce a number of parameters so that the network structure becomes more complex.

Adopting larger filters or increasing the number of layers will greatly increase the number of parameters to be calculated, which means that more memory resources will be required. Some researchers came up with dilated convolution for CNN in 2015 [38], so that only a small part of the weight can increase the receiving field. Dilated Convolution is to inject holes into the classical CNN to enlarge the receptive field of the network. Compared with the original normal convolution operation, dilation convolution has one more parameter: dilation rate, which refers to the total intervals between the convolution kernel points. For instance, the dilated rate of the conventional convolution operation is 1.

Receptive field size can be calculated by $2(dilated\ rate - 1) * (kernel\ size - 1) + kernel\ size$. According to this calculation, when the kernel size is 3×3, the receptive field size of standard convolution is 3, and the dilation convolution with dilated rate 2 has a larger receptive field size of 7. In other words, it is a 3×3 size kernel with 9 weights, and other elements are all set to zero. Therefore, it still has the same parameter as standard convolution but with the increased receptive field size 7.

3 Proposed Network

In MRI, CS is defined to realize authentic restoration based on a small set of k-space data instead of using the fully k-space data. The traditional CS-MRI model has been introduced and there are mainly two parts of this model. One is the data fidelity term and another is the regularization term. The first term guarantees the solution accords with the degradation process. The second term enforces desired property of the output.

$$\min_{x} \frac{\gamma}{2} \sum_{i=1}^{n_c} ||DFS_i m - y_i||_2^2 + R(m) \tag{5}$$

This is one of the general CS-MRI equations proposed by Jinming Duan et al. [6]. As stated by CS concepts, the image m which is reconstructed can be estimated by solving the unconstrained optimization problem shown above. The first term is the data fidelity term. m is MR image needed to be reconstructed,

y_i is the undersampled k-space data, F is the undersampled Fourier encoding matrix. D is the matrix that some positions are filled with zeros which can be seen as a sampling function. S_i is the ith coil sensitivity which is a diagonal matrix, and λ is a weight to make a trade-off between the two terms. The second term is a common sparse regularization term and it has many variants.

3.1 Variable Splitting

With intention of solving this optimisation issue, the variable splitting method can be apply to break down the relation between variable and other functions. In specific, the auxiliary variables u is introduced and it is equal to m and also enable $S_i m$ equal to x_i. According to the penalty function concept, these constraints are fed back into the model and the problem is converted to multi-optimisation problem:

$$\min_{m,u,x_i} \frac{\gamma}{2}\sum_{i=1}^{n_c}||DFx_i - y_i||_2^2 + R(u) + \frac{\alpha}{2}\sum_{i=1}^{n_c}||x_i - S_i m||_2^2 + \frac{\beta}{2}||u - m||_2^2 \tag{6}$$

here α and β are added into formula as penalty weights. In order to optimise (6), we can split it into three independent optimisation problems:

$$\begin{aligned} u^{k+1} &= \arg\min_u \frac{\beta}{2}||u - m^k||_2^2 + R(u) \\ x_i^{k+1} &= \arg\min_{x_i} \lambda\sum_{i=1}^{n_c}||DFx_i - y_i||_2^2 + \frac{\alpha}{2}\sum_{i=1}^{n_c}||x_i - S_i m^k||_2^2 \\ m^{k+1} &= \arg\min_m \frac{\alpha}{2}\sum_{i=1}^{n_c}||x^{k+1} - S_i m||_2^2 + \frac{\beta}{2}||u^{k+1} - m||_2^2 \end{aligned} \tag{7}$$

where k is the index number of kth iteration. Employing gradient descent and least-square (LS) to the middle and last equations of (7), the (8) could be computed out.

$$\begin{aligned} u^{k+1} &= denoiser(m^k) \\ x_i^{k+1} &= F^{-1}((\lambda D^T D + \alpha I)^{-1}(\alpha F S_i m^k + \lambda D^T y_i)) \\ m^{k+1} &= (\beta I + \alpha\sum_{i=1}^{n_c} S_i^H S_i)^{-1}(\beta u^{k+1} + \alpha\sum_{i=1}^{n_c} S_i^H x_i^{k+1}) \end{aligned} \tag{8}$$

Now, the conventional CS problem (5) is converted into a denoising operation and two mathematical optimization questions. The middle equation is called data consistency because it intuitively shows the consistency between frequency data and spatial data. The bottom equation combines the weight acquired from the previous two equations.

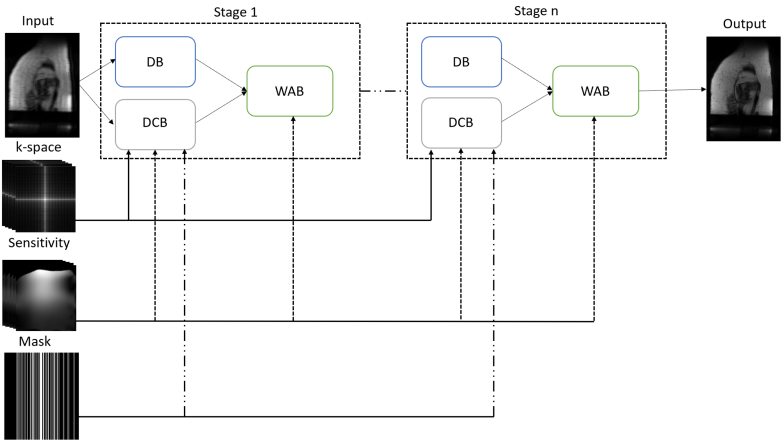


Fig. 1. The framework of VS-Net

In the end, the proposed framework based on Eq. (8) is constructed. The architecture of the network is shown in Fig. 1.

This network consists of three parts: denoiser block (DB), data consistency block (DCB), and weighted average block (WAB), they are coincided with Eq. (8) from upper to bottom respectively.

The basic CNN in the denoiser block is used [6]. Due to the simple architecture, it has limited performance in the denoising procedures. For the denoising block, there many the state of art denoising CNN can be implemented. In this research, inspired by sparse representation (SR) for image denoising in [34], we proposed a sparse dilated CNN denoising architecture to make an improvement on denoising outcomes. In specific, it could increase the depth of the denoising network by enlarging the receptive field without greatly increasing the running time and memory usage.

The presented architecture of sparse dilated network (SD-Net) is shown in Fig. 2.

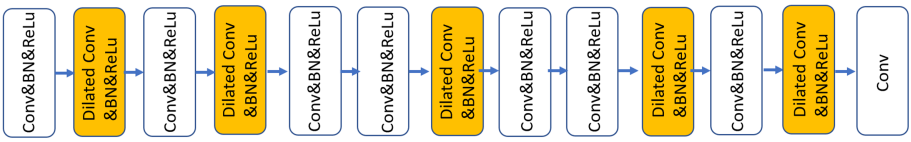


Fig. 2. The structure of the sparse dilated network (SD-Net)

In specific, the defined SD-Net based on dilated convolution, batch normalization, and standard convolution is unlike the normal denoising block as it incorporates a sparse mechanism. The 13 layers SD-Net is made of two main parts: the Dilated block and Conv block.

The Dilated block stands for the dilation convolution with the dilated rate of 2, BN, and the rectified linear function, ReLU [16] are linked. The Conv block is made of classical convolution, BN, and ReLU. The Dilated block is set at the second, fourth, seventh, tenth, and twelve layers in SD-Net.

The dilation convolutions are able to contain more detailed information [37]. According to this principle, the Dilated block can be addressed as high-powered points. The Conv block is set at the first, third, fifth, sixth, eighth, ninth, and eleventh layers in SD-Net, which can be regarded as low-powered points.

In particular, the kernel size applied in convolution layers is 3×3 which is a typical size used broadly. The input of the first layer is m (sensitivity-weighted under-sampled image) and it is actually a complex-valued image, in order to deal with complex values, we convert the complex value of the under-sampled image input into a common two-channel image. Therefore, in this case, the number of channels in the input is 2. In layers from 2 to 12, the input channels and output channels of filters are all 64. The dilated convolution layers are not placed in a contiguous sequence, which introduces the sparsity mechanism into the network by stacking some high-powered points and low-powered points. Furthermore, the SD-Net implements fewer high-powered points rather than many high-powered points to obtain a more comprehensive context. With this setting, SD-Net can not merely boost the denoising performance and ensure effective learning outcomes but also avoid the redundancy structure in the network.

3.2 Loss Function

The loss function plays an important role in the parameterization procedure and it could purify all parameters of the model as a single number that the improvements in this number indicate the quality of the network model. For MR reconstruction and denoising, the loss function is usually associated with the sameness relationship between the image which is reconstructed from the under-sampled image and an original image. For instance, the mean squared error (MSE) is a popular loss function that is also implemented in our network.

$$L(\Theta) = \min_{\Theta} \frac{1}{2n_i} \sum_{i=1}^{n_i} ||m_i^{n_{it}}(\Theta) - g_i||_2^2 \tag{9}$$

The n_i is the index number of training images, and g is the ground truth image, which is computed by fully sampling through $\sum_j^{n_c} S_j^H F^{-1} f_j$. The n_{it} is the stage number that indicates the iteration times going through the three network blocks DB, DCB and WAB. The f_j denotes the fully sampled raw data of the corresponded coil. Θ can be learned as the dilated parameters during the network training procedure. In specific, the Θ includes weight parameter λ in denoiser network and penalty weights α, β in DCB and WAB. They are all learnable during the training.

4 Experiments

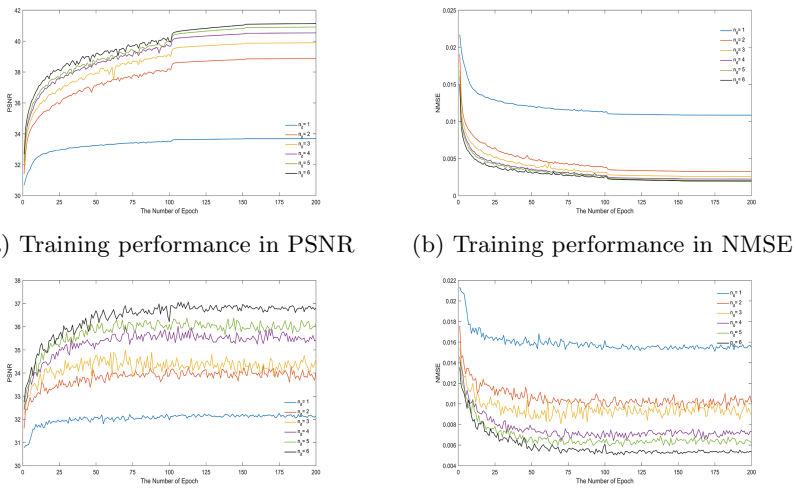
4.1 Dataset

Based on the framework of VS-Net, we incorporated the SD-Net into the denoiser block instead of the old one. For the dataset, the publicly clinical knee dataset is available in [11].

There are mainly 5 types of image acquisition protocols, we picked up coronal proton-density (PD) as the data of our experiments. For PD, the same 20 contents were scanned and the scan of each content contain around 40 slices with each slice has 15 channels. Coil sensitivity maps are also presented in the dataset were computed by BART [35]. In our case, we set the acceleration factor (AF) as 4-fold and use half of the subjects for training, the other half for testing. The trained SD-Net is to denoise the image during the reconstruction procedure of each 2D slice. Same as [6], the epoch number was set as 200 where the parameters of the network were normally optimized, with the learning rate equal to 10^{-3} and batch size 1.

4.2 Performance

With the purpose of investigating the influence of the stage number n_{it}, the n_{it} is set from 1–6 and visualized the training and testing numerical curves with epoch numbers in the first and last row of Fig. 3. The network has better performance when the number of stages is increased. This situation can be attributed to two possible reasons: the first reason is that the learning ability of the network is improved due to the increased weight; the second reason is that the optimization

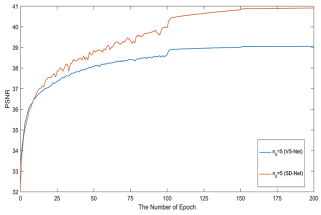


(a) Training performance in PSNR (b) Training performance in NMSE

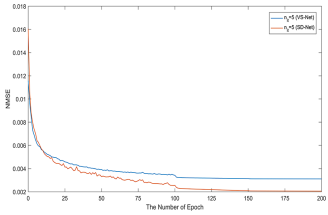
(c) Testing performance in PSNR (d) Testing performance in NMSE

Fig. 3. Performance of SD-Net from stage 1–6

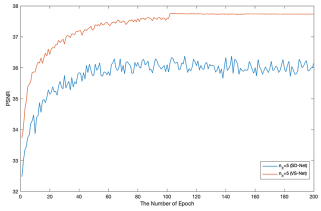
can be computed more precisely when the iterations are increased with larger stage numbers. Due to the limitation of GPU in the lab, we only increase the stage number to 6. Judging by the tendency of PSNR and NMSE in training and testing, it is possible to get better results by increasing the stage number to a larger number. However, it will introduce a high computational burden and is time-consuming.



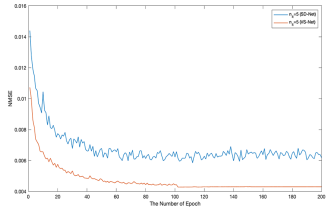
(a) Training performance in PSNR of stage 5



(b) Training performance in NMSE of stage 5



(c) Testing performance in PSNR of stage 5



(d) Testing performance in NMSE of stage 5

Fig. 4. Comparison of VS-Net and SD-Net

For the sake of exploring the effectiveness of denoising performance, the performance in stage number 5 is chosen to make a comparison between VS-Net and SD-Net. According to Fig. 4, in the same stage, it is obvious that our SD-Net has high performance either in PSNR or reconstruction error (NMSE) in training performance. However, we got slightly decreased performance for testing results, and the fluctuations in PSNR and NMSE lines of SD-Net are observed. Couples of reasons can lead to this situation, such as high learning rate, unsuitable batch size, and overfitting.

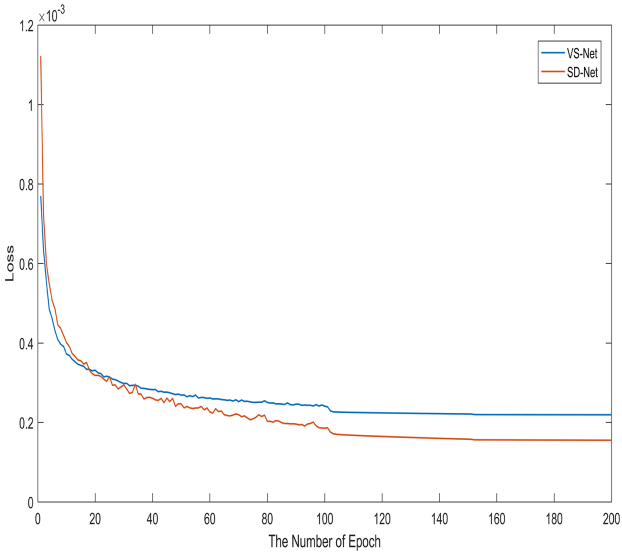
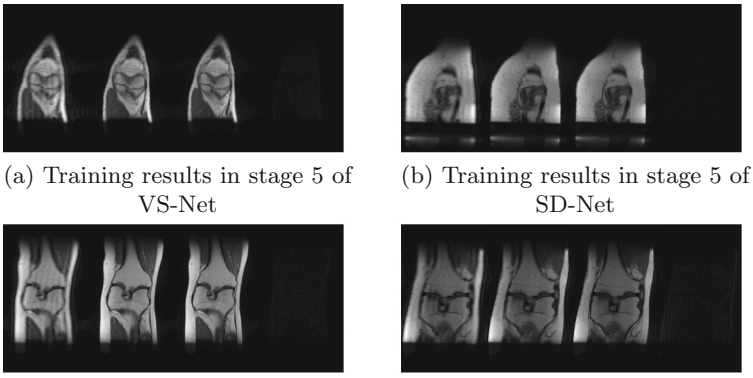


Fig. 5. Training loss in stage 5

According to the training loss performance in Fig. 5, we can notice that our SD-Net has a lower training loss than VS-Net. This effect might be brought with batch normalization operation in SD-Net. The batch normalization prevents the network from getting stuck in local minima and it can give a better error surface, the loss is low and converges to low training error. Additionally, batch normalization accelerates the training procedure and enables the network to converge in very few epochs.



(a) Training results in stage 5 of VS-Net

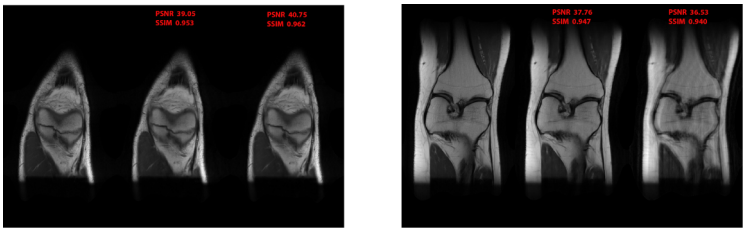
(b) Training results in stage 5 of SD-Net

(c) Testing results in stage 5 of VS-Net

(d) Testing results in stage 5 of SD-Net

Fig. 6. Visual Comparison of VS-Net and SD-Net, From left to right: under-sampled image, reconstructed image, ground truth and difference between reconstructed and ground truth image.

In Fig. 6, the different images are collected to show the accuracy consistency of various images. We can see the images reconstructed from corresponded undersampled images are very similar to their ground truth images, respectively.



(a) Training visual comparison (b) Testing visual comparison

Fig. 7. Visual comparison of VS-Net and SD-Net, From left to right: ground truth, reconstructed image from VS-Net, reconstructed image from SD-Net

In this comparison, both VS-Net and SD-Net were trained and tested in stage 5 in Fig. 7. For training visual comparison, the SD-Net has improvements in PSNR and SSIM. For testing visual comparison, the SD-Net has lightly unexpected performance both in PSNR and SSIM. We can see some local information is lost from the testing result of SD-Net. It can be caused by a gridding problem [36] in dilated convolution network. For the output of a certain layer in dilated convolution, the adjacent pixels are obtained from the convolution of independent subsets, and there is no correlation between each other. The gridding problem has two main features. One is the loss of local information which means that since the calculation method of the dilated convolution is similar to the checkerboard format, the convolution results obtained in a certain layer come from the independent set of the previous layer. Hence, there is no correlation among the convolution outputs of this layer. Another main feature is irrelevant information gathered from a distance. Due to the sparsely sampled input signal of the dilated convolution, there is no correlation between the information obtained by the long-distance convolution, which affects the result.

In order to mitigate the gridding effect in testing procedure in SD-Net. According to the paper [36], the hybrid dilated convolution (HDC) method is used to solve this problem. We combine the SD-Net with HDC to generate a new network SD-Net-H showing in Fig. 8 to make an improvement of SD-Net.

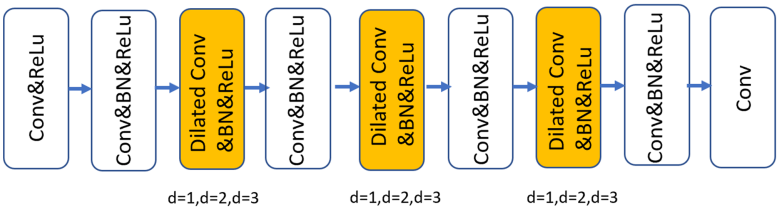


Fig. 8. SD-Net-H network structure

This solution forms a certain number of layers into a group, and then each group uses a continuously increasing dilated rate with repeating the same structure. Same as SD-Net, the low-powered points are introduced in the SD-Net-H to reduce the computational cost. In SD-Net-H, we use a combination of three different dilated rates (1, 2, 3). Moreover, the SD-Net-H network can obtain information from a wider range of pixels, avoiding the grid problem. At the same time, we can also adjust the receptive field arbitrarily by modifying the rate.

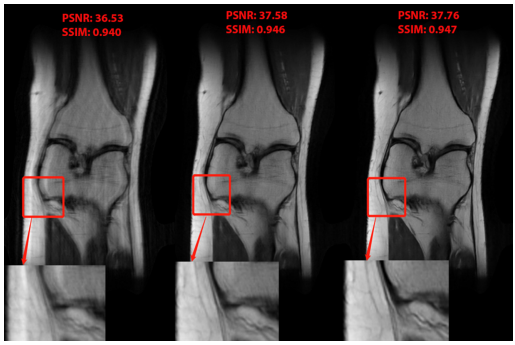


Fig. 9. Visual comparison of the reconstructed image during testing at stage 5, From left to right: SD-Net, SD-Net-H, VS-Net

As the testing results shown in Fig. 9, we can notice that there are some black and white stripes existing in reconstructed image from SD-Net while the SD-Net-H has alleviated the gridding effect and reduced noise artifacts. Besides, the local detailed information can be reconstructed more accurately than SD-Net by checking the same block areas of those three reconstructed images in terms of microstructure distribution. The global sparsity mechanism and local contiguous dilation are integrated together in the SD-Net-H to achieve a comparable performance with VS-Net. The PSNR and SSIM values computed from SD-Net-H and VS-Net are quite close to each other, which gives possibility to outperform the VS-Net by adjusting batch size and epoch number because SD-Net-H has

more complex structure in denoiser block. Moreover, since we introduced the local contiguous dilation, the SD-Net-H is more flexible in receptive field than VS-Net.

5 Conclusions

In this paper, we proposed the sparse dilated network (SD-Net) based on variable splitting (VS) for MRI denoising during an accelerated reconstruction procedure. We have introduced the SD-Net architecture and its excellent advantages for medical image denoising in a deep learning framework and every training stage greatly associated with the epoch number during the reconstruction process. For experimentation, we have demonstrated that the accuracy of SD-Net moderately improves with increasing stage number of network, and the parameters in each stage can be effectively learned due to the VS structure. Finally, we made comparisons and analyzed SD-Net with the empirical network VS-Net on an under-sampled Coronal PD dataset for 4-fold factors, and indicated its robustness. For training performance, we got better results but the testing performance is a little bit unexpected with some fluctuation. To mitigate this problem, we introduced HDC features into SD-Net to generated a new network. For future work, we will try to find specific reasons and propose an efficient solution for inconsistency between training and testing, and also achieve more higher performances in reconstruction with pleasant denoising results. The other non-medical state-of-the-art noisy images database can also be tested to see the benefit of the approach.

References

1. Buzug T.M.: Computed tomography. In: Kramme, R., Hoffmann, K.P., Pozos, R.S. (eds.) Springer Handbook of Medical Technology. Springer Handbooks, pp. 311–342. Springer, Berlin (2011). https://doi.org/10.1007/978-3-540-74658-4_16
2. Chow, L.S., Paramesran, R.: Review of medical image quality assessment. Biomed. Signal Process. Control **27**, 145–154 (2016)
3. Chung, Y.A., Weng, W.H.: Learning deep representations of medical images using siamese cnns with application to content-based image retrieval. arXiv preprint arXiv:1711.08490 (2017)
4. Diwakar, M., Kumar, M.: A review on CT image noise and its denoising. Biomed. Signal Process. Control **42**, 73–88 (2018)
5. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theor. **52**(4), 1289–1306 (2006)
6. Duan, J., et al.: VS-net: variable splitting network for accelerated parallel MRI reconstruction. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11767, pp. 713–722. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_78
7. Eldar, Y.C., Kutyniok, G.: Compressed Sensing: Theory and Applications. Cambridge University Press, Cambridge (2012)
8. Fatahi, M., Speck, O., et al.: Magnetic resonance imaging (MRI): a review of genetic damage investigations. Mutat. Res./Rev. Mutat. Res. **764**, 51–63 (2015)

9. Garbin, C., Zhu, X., Marques, O.: Dropout vs. batch normalization: an empirical study of their impact to deep learning. Multimedia Tools Appl. **79**, 1–39 (2020). ISSN: 1380-7501
10. Goyal, B., Agrawal, S., Sohi, B.: Noise issues prevailing in various types of medical images. Biomed. Pharmacol. J. **11**(3), 1227 (2018)
11. Hammernik, K., et al.: Learning a variational network for reconstruction of accelerated MRI data. Magn. Reson. Med. **79**(6), 3055–3071 (2018)
12. Hong, M., Yu, Y., Wang, H., Liu, F., Crozier, S.: Compressed sensing MRI with singular value decomposition-based sparsity basis. Phys. Med. Biol. **56**(19), 6311 (2011)
13. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
14. Ismail, A., Ahmad, S.A., Soh, A.C., Hassan, K., Harith, H.H.: Improving convolutional neural network (CNN) architecture (miniVGGNet) with batch normalization and learning rate decay factor for image classification. Int. J. Integr. Eng. **11**(4) (2019). ISSN: 2229-838X
15. Jung, W., Jung, D., Lee, S., Rhee, W., Ahn, J.H., et al.: Restructuring batch normalization to accelerate CNN training. arXiv preprint arXiv:1807.01702 (2018)
16. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (2017)
17. Lee, D., Yoo, J., Ye, J.C.: Deep artifact learning for compressed sensing and parallel MRI. arXiv preprint arXiv:1703.01120 (2017)
18. Liang, D., DiBella, E.V., Chen, R.R., Ying, L.: k-t ISD: dynamic cardiac MR imaging using compressed sensing with iterative support detection. Magn. Reson. Med. **68**(1), 41–53 (2012)
19. Liang, D., Liu, B., Wang, J., Ying, L.: Accelerating sense using compressed sensing. Magn. Reson. Med. Official J. Int. Soc. Magn. Reson. Med. **62**(6), 1574–1584 (2009)
20. Liang, D., Wang, H., Chang, Y., Ying, L.: Sensitivity encoding reconstruction with nonlocal total variation regularization. Magn. Reson. Med. **65**(5), 1384–1392 (2011)
21. Lingala, S.G., Jacob, M.: Blind compressive sensing dynamic MRI. IEEE Trans. Med. Imaging **32**(6), 1132–1145 (2013)
22. Lundervold, A.S., Lundervold, A.: An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik **29**(2), 102–127 (2019)
23. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn. Reson. Medi. Official J. Int. Soc. Magn. Reson. Med **58**(6), 1182–1195 (2007)
24. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing MRI. IEEE Signal Process. Mag. **25**(2), 72–82 (2008)
25. Macovski, A.: Noise in MRI. Magn. Reson. Med. **36**(3), 494–497 (1996)
26. Mredhula, L., Dorairangasamy, M.: An extensive review of significant researches on medical image denoising techniques. Int. J. Comput. Appl. **64**(14) (2013). ISSN: 0975-8887
27. Näppi, J.J., Hironaka, T., Yoshida, H.: Detection of colorectal masses in CT colonography: application of deep residual networks for differentiating masses from normal colon anatomy. In: Medical Imaging 2018: Computer-Aided Diagnosis. vol. 10575, p. 1057518. International Society for Optics and Photonics (2018)
28. Otazo, R., Kim, D., Axel, L., Sodickson, D.K.: Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. Magn. Reson. Med. **64**(3), 767–776 (2010)

29. Qu, X., et al.: Undersampled MRI reconstruction with patch-based directional wavelets. Top. Magn. Reson. Imaging **30**(7), 964–977 (2012)
30. Qu, X., Hou, Y., Lam, F., Guo, D., Zhong, J., Chen, Z.: Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator. Med. Image Anal. **18**(6), 843–856 (2014)
31. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging **30**(5), 1028–1041 (2010)
32. Reimer, P., Parizel, P.M., Meaney, J.F., Stichnoth, F.A.: Clinical MR imaging. Springer, Berlin (2010) https://doi.org/10.1007/978-3-540-74504-4
33. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
34. Tian, C., Zhang, Q., Sun, G., Song, Z., Li, S.: FFT consolidated sparse and collaborative representation for image classification. Arab. J. Sci. Eng. **43**(2), 741–758 (2018)
35. Uecker, M., et al.: Software toolbox and programming library for compressed sensing and parallel imaging. In: ISMRM Workshop on Data Sampling and Image Reconstruction, p. 41. Citeseer (2013)
36. Wang, P., et al.: Understanding convolution for semantic segmentation. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1451–1460. IEEE (2018)
37. Wang, Y., Wang, G., Chen, C., Pan, Z.: Multi-scale dilated convolution of convolutional neural network for image denoising. Multimedia Tools Appl. **78**(14), 19945–19960 (2019). https://doi.org/10.1007/s11042-019-7377-y
38. Yu, F., Koltun, V.: Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122 (2015)
39. Zhu, B., Liu, J.Z., Cauley, S.F., Rosen, B.R., Rosen, M.S.: Image reconstruction by domain-transform manifold learning. Nature **555**(7697), 487–492 (2018)