3T MRI super-resolution using 3D cycle-consistent generative adversarial network

Huy Quang Do, David Helbert, Pascal Bourdon, Mathieu Naudin, Carole Guillevin, Rémy Guillevin

3T MRI super-resolution using 3D cycle-consistent generative adversarial network

Huy Do[1,2], David Helbert[1,2], Pascal Bourdon[1,2], Mathieu Naudin[2,3], Carole Guillevin[2,3,4], Remy Guillevin[2,3,4]

[1] Xlim-ASALI, CNRS U-7252, University of Poitiers, France
[2] I3M Common Laboratory CNRS-Siemens, University and Hospital of Poitiers, France
[3] Poitiers University Hospital, CHU; Poitiers, France
[4] DACTIM-MIS/LMA Laboratory University of Poitiers, UMR CNRS 7348; Poitiers, France

Abstract

High-resolution magnetic resonance imaging (MRI) provides detailed anatomical information critical for clinical application diagnosis. However, current MRIs are acquired at clinical resolutions due to the limit of physical, technological, and economic considerations. On the other hand, existing approaches require paired MRI images as training data, which are difficult to obtain on existing datasets when the alignment between high and low-resolution images has to be implemented manually.

Within the scope of project, we aim to provide an end-to-end system to solve the super-resolution method on 3D MRI. Our proposed method derives from recent neural network developments and does not require paired data for efficient training. By integrating different models with separated functions, our 3D super-resolution CycleGAN (SRCycleGAN) achieved compelling results on MRI volumes. The output is close with ground-truth, showing a low distortion on different scaling factors. Besides, we also compare our method against different GAN-based methods in this field to highlight the performance.

Deep learning, neural network, generative model, MRI, medical image analysis, super-resolution

1 Introduction

Magnetic resonance imaging (MRI) is widely used in medical imaging because of its non-invasive assessment of the body's anatomy and physiology in health and disease while providing the best contrast resolution on soft tissues. MRI with high quality in the clinical and research domain is preferred

because it can provide critical structural details with a smaller voxel size, enabling accurate image analysis. However, MRI images are acquired with a finite resolution and less spatial coverage limited by the signal-to-noise ratio (SNR) or long-time scanning [18]. For example, a 3T MRI scanner may take from 2 to 48 hours to produce a high-resolution (HR) output, depending on the clinical-pathological question and the size of the area being scanned. Thus, improving medical image quality is becoming a potential direction with strong values for research and practical worlds.

Super-resolution (SR) is a process that produces high-resolution images from single- or multi-frame low-resolution images. In medical analysis, it becomes a very potential solution as a post-processing technique to increase the spatial resolution of medical scans after the acquisition process [15].

Before learning-based methods became widely used, resolution enhancement on medical images mostly relied on statistical methods such as interpolation [9], dictionary mapping [12, 13], regularization [14]. However, non-learning methods are limited by concepts of data representation. They are not robust methods, and the performance is not stable on images with abundant structural details.

In recent years, many learning-based methods have presented excellent performances to improve the quality of images. Convolutional neural networks (CNN) can model various complex structures while keeping cost time low for the sub-sequence process to produce output images after the training phase. Since Dong et al. [10] firstly introduced SRCNN, it has inspired many follow-ups approaches applied on different types of medical image [4]. In terms of MRI, there are several studies based on these methods have been proposed. For example, [20, 24] proposed a dense network with full use of hierarchical features to reconstruct high-resolution brain MRI; [16] used a residual network to reconstruct 3D HR cardiac volume from multiple 2D LR slices, or [21] introduced a 3D CNN for musculoskeletal MRI super-resolution.

However, most of the super-resolution methods for medical images in the real world require low and high-resolution MRI arranged in pairs to secure the efficient learning of the model. It is not easy to obtain a paired medical dataset. As mentioned above, the scanning phase to produce high-resolution MRI might take a very long time, while data size is usually limited due to the tedious and time-consuming task of paired data retrieval.

To address the problem, our work aims to provide a method that performs super-resolution tasks on MRI through unpaired training. The proposed method, entitled 3D super-resolution Cycle-consistent GAN (SRCycleGAN), developed based on recent researches in neural networks, CycleGAN [2] and Residual Dense Network (RDN) [3]. By integrating two networks into a new architecture, the proposed network can be executed on any publicly available MRI dataset through an unpaired training process, thus overcoming the limitations explained earlier. For evaluation, we compare

the performance of proposed methods to tricubic-interpolation method and GAN-based methods for super-resolution in 2D space: the ESRGAN [8] and our previous work on super-resolution CycleGAN [22].

In Section 2, we introduce details of the network architecture used in our experiments. For evaluation, we compare model performance with other SR methods: interpolation, super-resolution GAN-based methods on 2D data, and different states of SRCycleGAN on 3D data in Section 4. Then, we discuss experimental results in Section 5

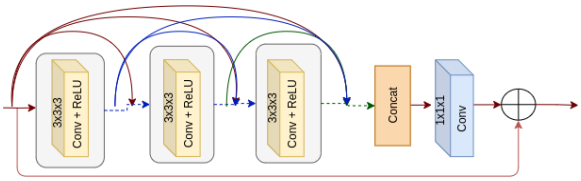

Figure 1: Details of residual dense block in 3D. RDB contains several convolutional layers followed by a ReLU activation function in continuous connection. The number of features through convolutional layers is increased with a certain growth rate, which aims at synthesizing information from input. The higher growth rate can further improve the performance of the network.

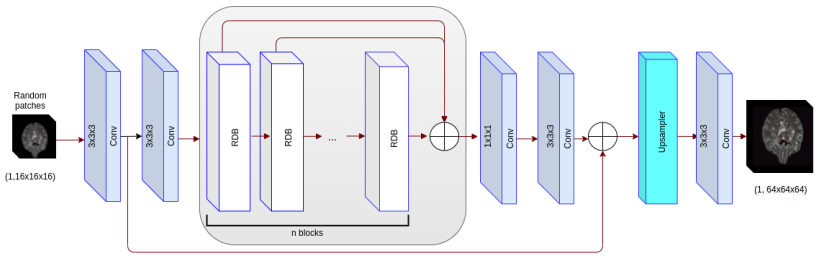

Figure 2: The architecture of the generator. It contains several RDBs for feature extraction. In the beginning, shallow features are extracted from raw input through two single convolutional layers. Then, features from RDBs are extracted and concatenated to form global features. A convolutional layer is used to adaptively fuse a range of features with different levels, followed by 3×3 convolutional layer to further extract features for global residual learning. Finally, the shallow and fused features are concatenated to obtain the dense feature before conducting up-scaling.

2 Method

2.1 Network architecture

2.1.1 Generative network architecture

Our generator is a 3D implementation of residual dense blocks (RDB) - a combination of residual blocks [11] and dense blocks [17] with the removing of batch normalization layers as building units in Figure 1. RDB was proved to reduce computational time, memory usage and speed up the training process while using fewer parameters.

The details of the generators are shown in Figure 2. Based on RDN architecture, the generators use both local and global features to synthesize information. By concatenating components of blocks, local features are synthesized from the states of preceding RDBs and whole layers in the current RDB. Then, global features are stacked to make full use of features from all preceding layers. Before upsampling, feature output has information of all subsequent layers, which preserves the feed-forward nature and extracts local densely features.

2.1.2 Adversarial network architecture

Following the fundamental CycleGAN [2], our final model aims at translating LR MRI volume into an HR MRI volume without the requirement of paired images for training. The network contains two generators to produce images between LR and HR MRI and two discriminators to predict output be real or generated. The final objective is to train the generator G to fool a differentiable discriminator D that is trained to distinguish generated SR images from LR images. Once well trained, the generator can create highly similar solutions to real images and thus difficult to classify by D. This encourages perceptually superior solutions residing in the subspace, or manifold, of real images.

The main difference between the two generators is the up/down-sampling operators. We use the deconvolutional as the upsampler for 3D data in this generator. Based on [6], it has been proved to work similarly to an upsampling operator on SR tasks. We use a convolutional layer as a downscaling operator to reconstruct LR from HR MRI in the downsampling generator in terms of downsampling.

In terms of discriminators, it is similar to a CNN contains several convolutional layers mixed with an instance normalization layer followed by Leaky ReLU (LReLU) activation to extract information from 3D volume and label whether it is a real or generated image. The depth of the network is customizable depending on the number of mixed blocks in the network. In the end, a convolutional layer with a single output channel is used to produce values ranging from 0 (reconstructed MRI) to 1 (real MRI).

3 Experiments

3.1 Dataset

We focus on implementing a method that can perform super-resolution through unpaired training. For a dataset with high quality enough and a large amount of MRIs, through the degradation process which used a down-sampling operator with fixed scaling factors [6], we expect to obtain the LR MRI volume with the minimum loss value on pixels and reduce the appearance of new artifacts.

In this work, we used the BraTS2018 dataset [7] - a dataset containing 3T MRI images with different types of sequences as training data. It is a very popular dataset in the field of multimodal brain MRI segmentation [**?**,23]. BraTS contains 285 volumes in the training set and 59 volumes in the validation set, including T1, T2, T1-ce 3T MRI. The variety of the dataset is confirmed when samples were acquired with different clinical protocols and various scanners from multiple institutions. The field of view is $155 \times 240 \times 240$, wherein slice thickness is 1mm. Training set in the BraTS is downgraded into LR volumes with fixed scale factors. All image resolution changes are executed during the training phase.

3.2 Training setting

		Tricubic	ESRGAN	2D SRCycleGAN	3D CycleGAN w/o DC	3D SRCycleGAN
BraTS	2x	30.24	40.12	42.57	44.75	**51.22**
	4x	27.26	35.85	36.76	35.78	**37.82**

Table 1: Average value of PSNR (dB) for scale factor $\times 2$ and $\times 4$ on BraTS dataset

Due to the size of the network and GPU memory capacity, the model is trained on patches. For each training batch, high-resolution volumes are randomly extracted from $155 \times 240 \times 240$ into different patcches with size $64 \times 64 \times 64$; corresponding to a size of $64/s \times 64/s \times 64/s$ on LR volumes where s is the scaling factor. Based on [13] and [20], the low-resolution samples can be obtained through a down-sampling by isotropic scaling factors and Gaussian filter. Each generator contains 3 RDBs, where each block includes three dense blocks in residual connection. The batch size is set as 2. The learning rate is initialized to 0.0001, and decay starts after every 200 epochs. The ADAM optimizer is used to update network weights based on training data. The training CycleGAN model takes an average of 45 hours with an NVIDIA Quadro P4000 8GB for 250 epochs.

4 Result

For testing, the MRIs from the BraTS test set (with size mentioned in Section 3.1) are degraded at scaling factor x2, or x4 to simulate low-resolution MRI. Then, the proposed model performs the reconstruction process back to HR MRI. Figure 3 shows the visualization of model performance on T1-weight 3T. The results of models on generated high-resolution MRIs are very exploitable. SRCycleGAN can maintain fine features of the main object compared to ground truth in qualitative visual quality.

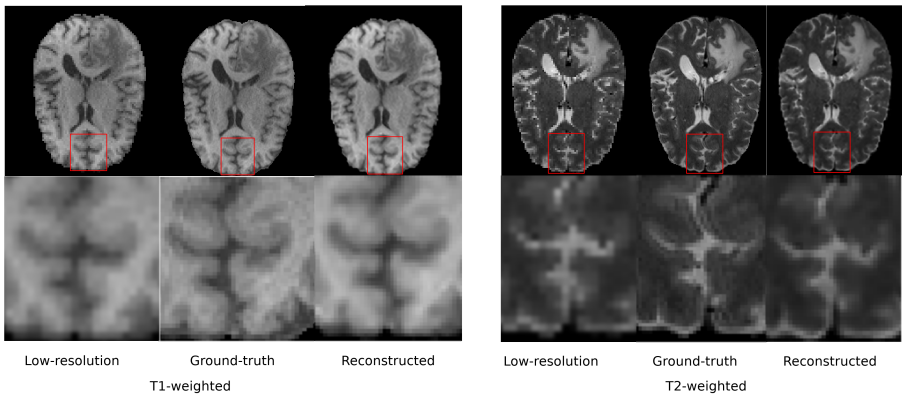

Figure 3: Visualization of 3D model performance on BraTS dataset. On randomly selected sample, zoom-ins are shown in the red box

Since there is no current benchmark dataset or state-of-the-art methods in the super-resolution for medical images, we want to evaluate the performance of different GAN-based models. We compare the performance of two states of 3D SRCycleGAN to interpolation (tricubic interpolation) and GAN-based method (2D ESRGAN [8] and 2D SRCycleGAN [22]) for measurement purposes. The 2D SRCyleGAN is our previous study on before working on 3D model. The primary difference between 2D and 3D versions is the difference of components in network architecture and also the model complexity. For example, in upscaling part, the 2D CycleGAN uses a sub-pixel function from ESPCNN [6] while the proposed 3D model uses transpose convolutional layers. A model in 3D space might provide an outstanding result compared to methods in 2D space due to the data structure and uniformity of 3D volumes. Quantitative results on table 1 show the average distortion and fidelity values obtained on the test set by all tested methods. Image quality measurements in terms of PSNR show that the 3D SRCycleGAN method can achieve relatively low distortion for $4\times$ and $2\times$ scale factor. In general, both implementations of 3D CycleGAN provides super-resolution MRI images with detailed textures, compared to 2D models and tricubic interpolation.

5 Discussion

The complexity of GAN-based models is very considerable, with millions of parameters. It increases along with the increase of model depth or the size of the input. Therefore, to reduce the computational cost of the model, training is performed on patches. Doing so can also ensure the diversity of data.

MRI volumes contain different or more complex spatial variations, correlations, and statistical properties than natural images, limiting the SR imaging performance of most traditional methods. Moreover, because the sampling and degradation operations are coupled and ill-posed, SR tasks cannot be performed beyond a marginal degree using traditional methods, which cannot effectively restore some fine features and suffer from the risk of producing a blurry appearance and new artifacts. In an unpaired training, the generated output cannot be compared to the target output to improve performance due to the difference in voxel values.

Although the reconstructed output at high-resolution levels is quite close to the original input in all aspects, at both the training and testing phase, we also observe the limitation of model performance in terms of blur on small objects, although the reconstructed images are fairly detailed. This issue might come from the degradation process. In some cases, when the resolution of input is not satisfied enough, the over degradation might lead to the lack of information of objects, and later is the not-fully reconstructed.

6 Conclusion

In this paper, we proposed SRCycleGAN to solve the super-resolution on medical images. The advantage of self-learning between two different classes can be used to perform the unpaired training. By integrating network architectures, we use CycleGAN to match the problem of SR. Evaluation of reconstructed images on both T1 and T2-weighted MRI shows exploitable results with low distortion and detailed texture. We also compared SRCycleGAN with different methods in the same domains to have an objective perspective about model performance. The quantitative evaluation shows that the SRCycleGAN is better and more measured than other methods on different scaling factors.

References

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

[2] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.

[3] Zhang, Yulun, et al. "Residual dense network for image super-resolution." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[4] Zeng, Wei, et al. "A comparative study of CNN-based super-resolution methods in MRI reconstruction and its beyond." Signal Processing: Image Communication 81 (2020): 115701.

[5] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[6] Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[7] Menze, Bjoern H., et al. "The multimodal brain tumor image segmentation benchmark (BRATS)."IEEE transactions on medical imaging 34.10 (2014): 1993-2024.

[8] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European Conference on Computer Vision (ECCV) Workshops. 2018.

[9] Siu, Wan-Chi, and Kwok-Wai Hung. "Review of image interpolation and super-resolution." Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference. IEEE, 2012.

[10] Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." European conference on computer vision. Springer, Cham, 2014.

[11] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[12] Li, Yinghua, et al. "Super-resolution of brain MRI images using over complete dictionaries and non local similarity." IEEE Access 7 (2019): 25897-25907.

[13] Rueda, Andrea, Norberto Malpica, and Eduardo Romero. "Single-image super-resolution of brain MR images using overcomplete dictionaries." Medical image analysis 17.1 (2013): 113-132.

[14] Tourbier, Sébastien, et al. "An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization." NeuroImage 118 (2015): 584-597.

[15] Van Reeth, Eric, et al. "Super-resolution in magnetic resonance imaging: a review." Concepts in Magnetic Resonance Part A 40.6 (2012): 306-325.

[16] Oktay, Ozan, et al. "Multi-input cardiac image super-resolution using convolutional neural networks." International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2016.

[17] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[18] Plenge, Esben, et al. "Super-resolution methods in MRI: can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time?." Magnetic resonance in medicine 68.6 (2012): 1983-1993.

[19] Zhang, Yongbing, et al. "Multiple cycle-in-cycle generative adversarial networks for unsupervised image super-resolution." IEEE transactions on Image Processing 29 (2019): 1101-1112.

[20] Pham, Chi-Hieu, et al. "Brain MRI super-resolution using deep 3D convolutional networks." 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE, 2017.

[21] Chaudhari, Akshay S., et al. "Super-resolution musculoskeletal MRI using deep learning." Magnetic resonance in medicine 80.5 (2018): 2139 2154.

[22] Do, H., Bourdon, P., Helbert, D., Naudin, M., Guillevin, R. (2021). "7T MRI super-resolution with Generative Adversarial Network". In IST Electronic Imaging 2021 Symposium.

[23] Bakas, Spyridon, et al. "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features." Scientific data 4.1 (2017): 1 13.

[24] Chen, Yuhua, et al. "Brain MRI super resolution using 3D deep densely connected neural networks." 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018