Diffusion MRI Spatial Super-Resolution Using Generative Adversarial Networks

Enes Albay[(⊠)], Ugur Demir, and Gozde Unal

Istanbul Technical University, Istanbul 34469, Turkey
{albay,ugurdemir,gozde.unal}@itu.edu.tr

Abstract. Spatial resolution is one of the main constraints in diffusion Magnetic Resonance Imaging (dMRI). Increasing resolution leads to a decrease in SNR of the diffusion images. Acquiring high resolution images without reducing SNRs requires larger magnetic fields and long scan times which are typically not applicable in the clinical settings. Currently feasible voxel size is around $1 \, \text{mm}^3$ for a diffusion image. In this paper, we present a deep neural network based post-processing method to increase the spatial resolution in diffusion MRI. We utilize Generative Adversarial Networks (GANs) to obtain a higher resolution diffusion MR image in the spatial dimension from lower resolution diffusion images. The obtained real data results demonstrate a first time proof of concept that GANs can be useful in super-resolution problem of diffusion MRI for upscaling in the spatial dimension.

Keywords: Magnetic resonance imaging (MRI)
Diffusion MRI (dMRI) · Super resolution
Generative adversarial networks (GANs)

1 Introduction

Water molecules undergo random movement and diffuse in an environment due to second law of thermodynamics. Diffusion phenomenon enables us to map fibrous substances using principles of magnetic resonance imaging (MRI). Diffusion magnetic resonance imaging (dMRI) takes advantage of signal attenuation that takes place due to diffusion of water molecules in a tissue that is being imaged. Although the signal attenuates isotropically in a free water environment, the signal shows varying attenuations in a restricted environment. This gives an opportunity of in vivo imaging of the internal structure of the human brain white matter, which contains fibrous material that restricts water molecules movements in some directions while water molecules move freely in other directions [10].

Even though dMRI allows microscopic imaging of the white matter at very high magnetic fields, spatial resolution of dMRI is restricted clinically because with the current technology, very high magnetic fields cannot be used ante-mortem. Furthermore, long scan times are not clinically feasible for microscopic

---

E. Albay and U. Demir—Equal contribution.

resolution. At lower magnetic fields, signal to noise ratio (SNR) becomes problematic for small voxel sizes. Currently, for diffusion image volumes, clinically applicable voxel size is about 1 mm^3 [11], which is relatively coarse with respect to underlying microstructure of brain tissue. Diameter of neuronal axons in brain white matter is at most 30 μm [7], therefore, a typical voxel contains thousands of fiber populations, possibly lying along different directions with crossings, splaying, or kissing architectures. Hence, increasing both spatial resolution and angular resolution of dMRI using post-processing techniques is desirable and would aid in post-analysis of dMRI data.

In this paper, we present a post-processing method to generate higher spatial resolution dMRI volumes based on an end-to-end generative adversarial network (GAN) framework [5]. GANs learn a mapping from low resolution diffusion MRI data to synthesize a high resolution counterpart. Its main difference from conventional methods is that GANs learn a non-linear model from pairs of low resolution-high resolution data rather than performing a blind interpolation.

2   Related Works

Only a few spatial super resolution methods for diffusion MRI were presented in the literature. Conventional methods for super resolution are typically based on up-sampling with interpolation of low resolution data. An early super resolution approach to diffusion data is based on combination of two shifted images to create an up-sampled image [14], which led to blurry results. Alternatively, a track density approach was presented to obtain super resolution in white matter fiber tracts based on tractography information, however, this method does not up-sample the underlying spatial structure of the diffusion images [1]. A Markov chain Monte Carlo method, the Metropolis-Hastings algorithm is utilized by [19] to create a generative model of local information and sharpens images according to local structure while increasing spatial resolution. This is different from our approach as it does not actually directly learn the data distribution. A recently suggested method proposes using RGB image enhancement method with diffusion images, however, not leading to clear results [18]. It is observed that diffusion weighted images are blurry and ODFs are corrupted with respect to the ground truth data. Recent studies on texture synthesis have shown that convolutional neural networks and adversarial training can be successfully applied to super-resolve images at high upscale factors [3,12]. This was the motivation of our method, which is presented next.

3   Method

We introduce a deep GAN based single slice super-resolution model that takes a down-sampled low resolution dMRI axial slice I_{LR} and synthesizes its high resolution counterpart $\hat{I}_{HR}$. Each down-sampled axial slice from a brain volume is upscaled to the desired resolution with a certain scale factor through bi-cubic interpolation, which is called I_{LR}^{bc}. In this paper, we exemplify the spatial super

resolution model with an up-sampling factor of two. The generative model takes I_{LR}^{bc}, and tries to expose the high frequency details by exploring the context of low-resolution image. The trained network resolves the blurriness and generates sharp images filled with estimated missing details. Overall flow of our method is depicted in Fig. 1.
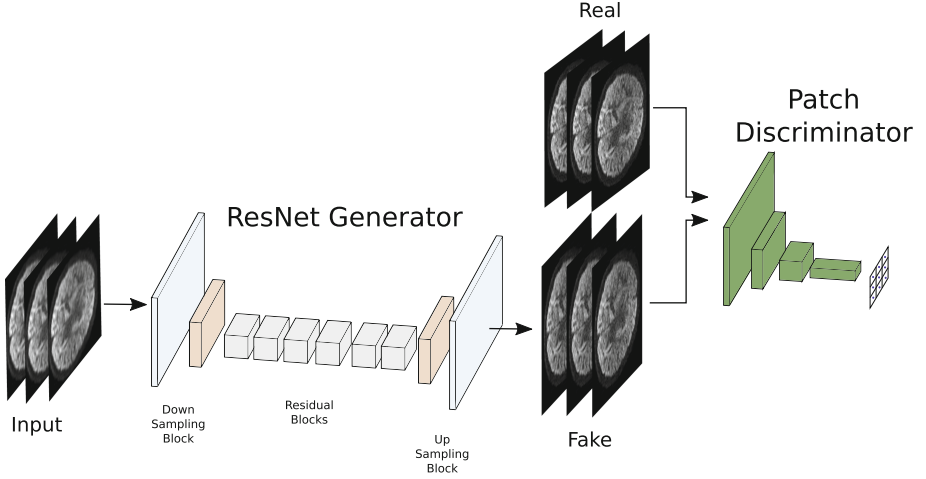

Fig. 1. General structure of architecture. Input slices are x2 upscaled by bi-cubic interpolation. Generator network takes input I_{LR}^{bc} images and synthesizes "fake" high-resolution slices I_{HR}. Discriminator network evaluates artificially generated slices and produces an adversarial loss.

During training of the generative model, an adversarial training [5] approach is used in order to produce more realistic looking outputs. Training procedure intends to minimize the combination of an adversarial loss produced by the discriminator network, and a pixel-wise reconstruction loss (an L2 Loss) to conditionally generate samples from the high resolution image distribution. Details of this procedure are described next.

3.1 Generative Adversarial Networks

GANs have been used to figure out distribution of the input data by learning a mapping from a noise variable to the data space [5]. Recent studies show that once the distribution is learned, the model can be used to generate realistic looking samples [6,12]. Apart from sample generation, GANs are also used to learn a mapping between contextually paired two images [6]. In our super-resolution problem, low resolution image I_{LR}^{bc} is given as a condition to the generator and it is expected that our model learns a mapping G that translates I_{LR}^{bc} to I_{HR}.

There are two different neural networks in the adversarial training phase. The generative network G corresponds to mapping function between the input and

the output. The purpose of the discriminator network D is distinguishing the real images from the artificially synthesized ones. While the network G aims to fool D, at the same time, D is trained to improve its accuracy. This optimization problem corresponds to a minimax game, which can be formulated as:

$$\min_{\theta_G} \max_{\theta_D} E[\log D(I_{HR})] + E[\log(1 - D(G(I_{LR}^{bc})))]. \tag{1}$$

As long as D successfully classifies its input, G benefits from the gradient provided by the D network via its adversarial loss.

Generator Network. The architecture of the generative network is ResNet, which is composed by following the guidelines described in [8]. It consists of down-sampling layers, residual blocks and up-sampling layers. There are two down-sampling layers and each one consists of a convolution layer with stride set to 2, batch normalization layer and Leaky ReLU (LReLU) activation. There are six residual blocks in the architecture. The up-sampling blocks recover spatial resolution of the activation maps in order to reach desired height and width for a slice. An up-sampling layer contains resized convolution [13], batch normalization and LReLU activation. Additionally, a 7×7 convolution layer with a Tanh activation is added to end of the network.

Discriminator Network. We utilize a patch based discriminator network PatchGAN [6] design which evaluates local patches of the generated image and gives an average score as a measure instead of considering the whole input. This gives more robust results than the vanilla GAN. Our patch based discriminator has 6 convolution layers followed by batch normalization except the first and the last layers. First 5 layers have LReLU activation and the last convolution layer pass its outputs to Sigmoid activation.

3.2   Training Objective

The main objective function is formed by combining the reconstruction and adversarial losses. The total loss function is optimized with back-propagation by using Adam optimizer [9]. L2 pixel-wise distances between the synthesized image and the ground truth are used as reconstruction loss. Even though it forces the network to produce a blurry output, it guides the network to roughly predict texture colors and low frequency details. Discriminator network computes a score according to quality of the generator outputs and, and is used as an adversarial loss as described in Eq. 1.

Total loss function defines the objective used in the training phase. Each component of the total loss function is governed by a coefficient λ:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{rec} + \lambda_2 \mathcal{L}_{adv}. \tag{2}$$

where $\mathcal{L}_{rec}$ is reconstruction loss and $\mathcal{L}_{adv}$ is adversarial loss.

4   Experiments

4.1   Dataset

A diffusion dataset obtained from the Human Connectome Project (HCP) is used [17], where 29 diffusion subjects are randomly selected, and 25 are used to train the network and four of them are used in testing. HCP diffusion images are multi-shell, from which a single shell is extracted for each subject, which resulted in 108 diffusion volumes per subject. It was shown that the best b-value with an SNR of 30 for a non-diffusion weighted volume is between 3000 and 4000 s/mm^2 [16]. As the SNR of the HCP data is greater than 30 for the non-diffusion weighted volume, a single shell that has b-value 2000 s/mm^2 is selected. DIPY [4] library is used in all the analysis. All 108 diffusion-weighted volumes including non-diffusion weighted volumes are used in training the network, and the super resolution model is applied for up-sampling of all diffusion volumes in the test stage. In Fig. 2, sample visual results for diffusion-weighted images with I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} from a selected subject are shown. It can be seen that I_{LR}^{bc} is blurry and our network produces $\hat{I}_{HR}$ image with a success.
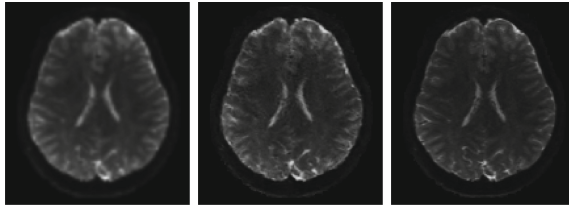

Fig. 2. Diffusion-weighted images for I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} of selected one subject, respectively.

4.2   SNR Comparison and FA (Fractional Anisotropy) Maps

SNR values of I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} are compared to measure how image generation introduced noise to diffusion data. The same ROI is used to compare each of the images. SNR values are computed according to most signal attenuation direction approach [2].

Corpus Callosum (CC) is segmented automatically using fractional anisotropy (FA) values. SNR values in the CC region in x, y ad z-directions are compared for four different subjects in Table 1. It can be observed that $\hat{I}_{HR}$ shows closer SNR values to I_{HR} than those of the I_{LR}^{bc}.

As a second quantitative evaluation, FA histograms are calculated for each subject. Figure 3 depicts the histograms for two of the subjects. The histograms show that I_{HR} and $\hat{I}_{HR}$ exhibit very similar distributions for FA values that are greater than 0.4. Other two subjects displayed similar distributions.

The generated FA maps and color FA maps are shown in Fig. 4 for one of the subjects. It can be observed that I_{HR} and $\hat{I}_{HR}$ have similar FA and color FA

Table 1. SNR values comparison in the various directions. Generated output has similar SNR values with the ground truth.

	Subject 1			Subject 2			Subject 3			Subject 4		
Direction	I_{LR}^{bc}	$\hat{I}_{HR}$	I_{HR}	I_{LR}^{bc}	$\hat{I}_{HR}$	I_{HR}	I_{LR}^{bc}	$\hat{I}_{HR}$	I_{HR}	I_{LR}^{bc}	$\hat{I}_{HR}$	I_{HR}
b0	76.05	**62.64**	50.19	72.52	**63.04**	51.02	40.8	**37.26**	29.1	113.47	**92.87**	83.51
x-dir	12.15	**10.92**	9.2	12.39	**11.59**	10.3	5.99	**4.62**	5.24	16.08	**12.82**	14.29
y-dir	39.27	**32.14**	27.44	36.05	**34.21**	28.7	18.91	**16.35**	15.25	52.0	**43.43**	45.52
z-dir	33.75	**25.79**	25.99	31.44	**26.42**	27.95	18.78	**14.12**	15.89	56.57	**43.44**	49.33

(a) Subject 1   (b) Subject 2
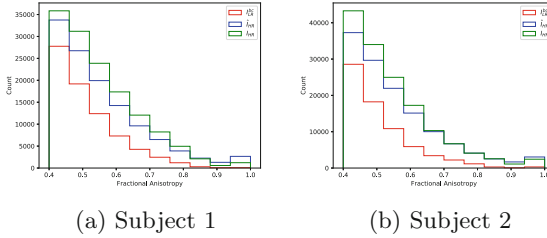
Fig. 3. FA distributions for two subjects. Green shows I_{HR} distribution, blue shows $\hat{I}_{HR}$ and red shows I_{LR}^{bc} (Color figure online)
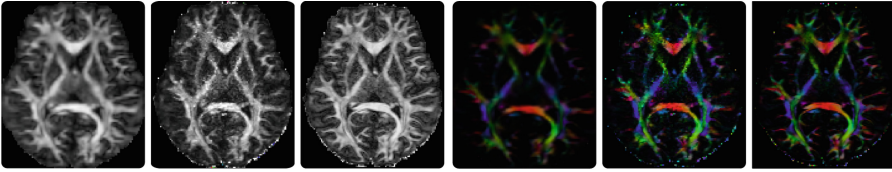
Fig. 4. Color FA maps are shown for I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} images respectively. Red indicates right-left axis, green for anterior-posterior axis and blue for inferior-superior axis diffusion. (Color figure online)

maps while the baseline bi-cubic interpolation introduces attenuation and blur in the FA maps.

4.3   Tensor and ODF Analysis

For further evaluation of the quality of the reconstructed high resolution diffusion volumes, the diffusion tensor models are constructed for $\hat{I}_{HR}$, I_{LR}^{bc} and I_{HR}. Figure 5 shows the results for one of the subjects. Similar tensor orientations and strengths at the crossing points of CC and corticospinal tracts (CST) are observed for the $\hat{I}_{HR}$ and I_{HR} of test subjects.

The orientation distribution functions (ODFs) are generated using constrained spherical deconvolution (CSD) [15] over the high resolution diffusion volumes. In Fig. 6, reconstructed ODFs are shown for one subject. It was observed that I_{LR}^{bc} has bigger artifactual side lobes.
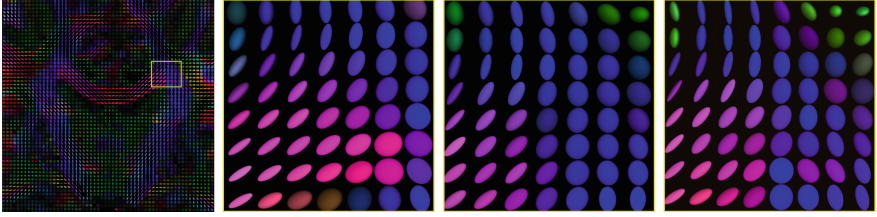
Fig. 5. Reconstructed tensors of diffusion image for one selected subject. Leftmost image is a general coronal view. Other views show I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} images at the area of crossing between CC and CST, respectively. Red indicates right-left axis, green for anterior-posterior axis and blue for inferior-superior axis diffusion. (Color figure online)
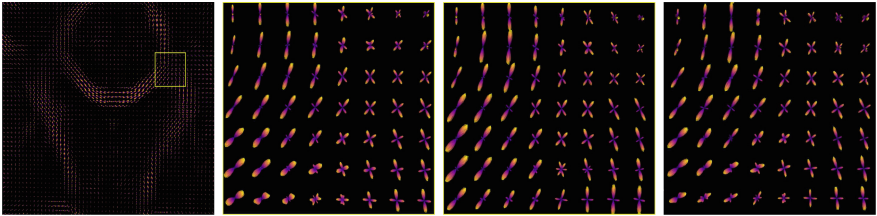

Fig. 6. Reconstructed ODFs over high resolution diffusion images. Leftmost image is a general coronal view. Crossing between CC and CST is shown in the yellow rectangle. Other views show ODFs of I_{LR}^{bc}, $\hat{I}_{HR}$ and I_{HR} images at the area of crossing between CC and CST, respectively. (Color figure online)

5   Conclusions

In this paper, for the first time, an end-to-end super-resolution method based on GANs is presented for dMRI data. This approach does not assume any model, does not simply interpolate existing data but learns a data-driven generative mapping. The experimental quantitative results such as distribution of FA values and SNR as well as qualitative results such as FA maps, color FA maps, reconstructed tensors and ODFs demonstrate that GANs produce promising results to create higher resolution data using low resolution dMRI input. Although our work shows a preliminary proof of concept with GANs to increase the spatial resolution of dMRI twofold, our future work investigates further tuning of networks with larger training sets, increasing the resolution to triple, quadruple, or higher scale factors, and extending our work to angular up-sampling.

References

1. Calamante, F., Tournier, J.D., Heidemann, R.M., Anwander, A., Jackson, G.D., Connelly, A.: Track density imaging (TDI): validation of super resolution property. Neuroimage **56**(3), 1259–1266 (2011)
2. Descoteaux, M., Deriche, R., Le Bihan, D., Mangin, J.F., Poupon, C.: Multiple q-shell diffusion propagator imaging. Med. Image Anal. Front. Neuroinformatics **15**(4), 603–621 (2011)
3. Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8692, pp. 184–199. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10593-2_13
4. Garyfallidis, E.: Dipy, a library for the analysis of diffusion MRI data. Front. Neuroinformatics **8**, 8 (2014)
5. Goodfellow, I., et al.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 27, pp. 2672–2680. Curran Associates, Inc. (2014)
6. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. arxiv (2016)
7. Johansen-Berg, H., Behrens, T.E.: Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy. Academic Press (2013)
8. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
9. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. CoRR abs/1412.6980 (2014). http://arxiv.org/abs/1412.6980
10. Le Bihan, D., Basser, P.J.: Molecular diffusion and nuclear magnetic resonance. In: Diffusion and Perfusion Magnetic Resonance Imaging, pp. 5–17 (1995)
11. Le Bihan, D., Iima, M.: Diffusion magnetic resonance imaging: what water tells us about biological tissues. PLoS Biol. **13**(7), e1002203 (2015)
12. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. CoRR abs/1609.04802 (2016)
13. Odena, A., Dumoulin, V., Olah, C.: Deconvolution and checkerboard artifacts. Distill (2016). https://doi.org/10.23915/distill.00003, http://distill.pub/2016/deconv-checkerboard
14. Peled, S., Yeshurun, Y.: Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging. Magn. Reson. Med. **45**(1), 29–35 (2001)
15. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. Neuroimage **35**(4), 1459–1472 (2007)
16. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage **23**(3), 1176–1185 (2004)
17. Van Essen, D.C., et al.: The human connectome project: a data acquisition perspective. Neuroimage **62**(4), 2222–2231 (2012)

18. Yang, Z., He, P., Zhou, J., Wu, X.: Non-local diffusion-weighted image super-resolution using collaborative joint information. Exp. Ther. Med. **15**(1), 217–225 (2018)
19. Yap, P.-T., An, H., Chen, Y., Shen, D.: A generative model for resolution enhancement of diffusion MRI data. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8151, pp. 527–534. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40760-4_66