Improving the feasibility of deep learning based *S*uper-*R*esolution MRI using *N*oisy high-resolution *R*eference data (SRNR)

Jiaxin Xiao[1], Zihan Li[2], Ziyu Li[3], Berkin Bilgic[4, 5], Jonathan R. Polimeni[4, 5], Susie Y. Huang[4, 5], Qiyuan Tian[4, 5]
[1]Department of Electronic Engineering, Tsinghua University, Beijing, P.R. China; [2]Department of Biomedical Engineering, Tsinghua University, Beijing, P.R. China; [3]FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom; [4]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA; [5]Harvard Medical School, Boston, MA, USA.

Introduction: Neural networks (NNs) have been shown to outperform conventional super-resolution methods and are widely adopted in biomedical imaging[1-4]. Nonetheless, the training of NNs requires high-quality high-resolution reference data acquired in many subjects, as a result reducing the practical feasibility and accessibility of NN-based super-resolution. Higher resolution necessitates not only a larger number of data samples and longer acquisition times to encode a whole-brain volume, but also multiple signal averages to improve the inherently low signal-to-noise ratio (SNR).

In this work, we leverage the mechanism of an unsupervised NN-based denoising method Noise2Noise[5], and train NNs to map low-resolution images to their residuals compared to noisy high-resolution reference images. We systematically demonstrate that results from NNs trained using noisy and high-SNR references are similar for both simulated and empirical data. SRNR suggests a smaller number of repetitions of high-resolution reference data can be used to simplify the training data preparation for super-resolution MRI, thereby benefiting a broader range of its clinical and neuroscientific applications.

Methods: *Simulated data.* High-SNR T1w images acquired at $0.7\times0.7\times0.7$ mm^3 resolution of 30 healthy subjects from the Human Connectome Project (HCP)[6, 7] WU-Minn-Ox Consortium were used as the ground-truth high-resolution data. Low-resolution thick-slice images were generated by down-sampling the ground-truth data along the superior-inferior direction to $0.7\times0.7\times3.5$ mm^3 resolution. Noisy high-resolution data were simulated by adding Gaussian noise (μ=0, σ=0.2, 0.4, 0.6, …, 1.6$\times\sigma$ of brain voxel intensities) to the native data as the noisy reference high-resolution data (Fig. 1a).

Empirical data. Highly accelerated (R=3×3) high-resolution ($0.8\times0.8\times0.8$ mm^3, 10 repetitions) and lower-resolution ($1\times1\times1$ mm^3, 1 repetition) T1w data were acquired on 9 healthy subjects using a Wave-MPRAGE sequence (five on a MAGNETOM Skyra scanner and four on a MAGNETOM Prisma scanner, Siemens Healthineers). For each subject, the 10 repetitions of high-resolution T1w data were co-registered to the first repetition and averaged using FreeSurfer's "mri_robust_template" function[8-11]. The low-resolution T1w data were co-registered to the 10-repetition averaged volume using NiftyReg's "reg_aladin" function.

Network and training. A modified U-Net (MU-Net) with all pooling and up-sampling layers removed was adopted and implemented using Tensorflow. The input of MU-Net was the low-resolution image volume up-sampled to the target high-resolution using cubic spline interpolation. The output of MU-Net was the residual volume between the input and target volume. The training and validation were performed on the simulation data of 15 subjects with the noisy and high-SNR high-resolution data as the target as well as the empirical data of 4 subjects with the single-average and 10-average data as the target respectively, with Adam optimizers and L2 loss.

Evaluation. The mean absolute error (MAE), peak SNR (PSNR), and structural similarity index (SSIM) were computed to quantify the similarity between up-sampled, super-resolved images and ground-truth high-resolution images, as well as between super-resolved images from different models.

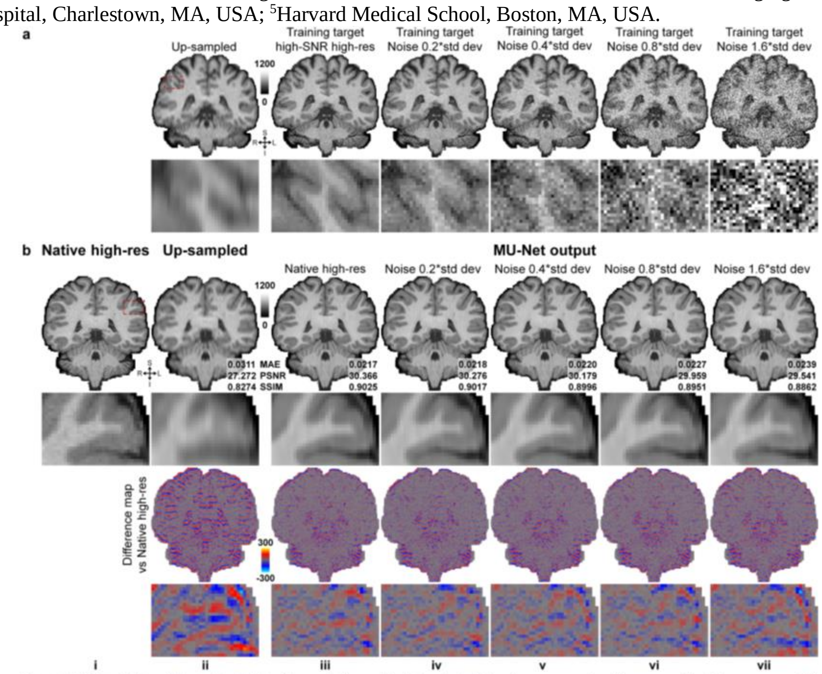

Figure 1. Simulation data and results. Coronal slices of training data (a) and super-resolved images with difference map (b).

a — Group mean (±std dev) of MAE, PSNR, SSIM compared to native high-resolution image

	Up-sampled	Clean target	Target noise 0.2*std dev	Target noise 0.4*std dev	Target noise 0.6*std dev	Target noise 0.8*std dev	Target noise 1.0*std dev	Target noise 1.2*std dev	Target noise 1.4*std dev	Target noise 1.6*std dev
MAE	0.0313 ± 0.00162	0.0214 ± 0.00209	0.0215 ± 0.00205	0.0218 ± 0.00200	0.0222 ± 0.00197	0.0225 ± 0.00196	0.0229 ± 0.00188	0.0231 ± 0.00196	0.0235 ± 0.00191	0.0238± 0.00187
PSNR (dB)	27.17 ± 0.444	30.26 ± 0.686	30.22 ± 0.665	30.10 ± 0.649	29.99 ± 0.646	29.86 ± 0.627	29.75 ± 0.602	29.64 ± 0.621	29.52 ± 0.602	29.43 ± 0.571
SSIM	0.830 ± 0.01381	0.906 ± 0.01691	0.906 ± 0.01660	0.904 ± 0.01647	0.900 ± 0.01619	0.899 ± 0.01647	0.896 ± 0.01588	0.896 ± 0.01633	0.893 ± 0.01620	0.890 ± 0.01635

b — Group mean (±std dev) of MAE, PSNR, SSIM compared to output of model trained with clean target

	Target noise 0.2*std dev	Target noise 0.4*std dev	Target noise 0.6*std dev	Target noise 0.8*std dev	Target noise 1.0*std dev	Target noise 1.2*std dev	Target noise 1.4*std dev	Target noise 1.6*std dev
MAE	0.0095 ± 0.00033	0.0096 ± 0.00034	0.0100 ± 0.00031	0.0104 ± 0.00037	0.0109 ± 0.00038	0.0114 ± 0.00035	0.0118 ± 0.00040	0.0122± 0.00037
PSNR (dB)	36.23 ± 0.329	36.32 ± 0.343	36.19 ± 0.283	35.75 ± 0.327	35.52 ± 0.304	35.11 ± 0.319	34.85 ± 0.311	34.61 ± 0.257
SSIM	0.982 ± 0.00114	0.982 ± 0.00120	0.981 ± 0.00105	0.979 ± 0.00127	0.977 ± 0.00115	0.976 ± 0.00138	0.974 ± 0.00147	0.972 ± 0.00134

Table 2. Simulation image similarity metrics. Group mean (± std) of the whole-brain averaged MAE, PSNR, and SSIM of up-sampled low-resolution and super-resolved images across 15 evaluation subjects compared to native high-resolution (a) and high-SNR training results (b).

Results: Super-resolved images from NNs trained with noisy high-resolution references were visually (Figs. 1, 3) and quantitatively (Tabs. 2, 4) similar to those from NNs trained with high-SNR reference data. For the simulated data, the similarity gradually decreased as the noise level in the reference data increased, as expected (Tab. 2). Even for the noisiest case (Fig. 1, vii), the similarity compared to the ground-truth data only slightly decreased (Tab. 2a, MAE: 0.0214 vs. 0.0238, 11.2%; PSNR: 30.26 dB vs. 29.43 dB, -2.7%; SSIM: 0.906 vs. 0.890, -1.77%) and the inter-result similarity was high (Tab. 2b, MAE: 0.0122; PSNR: 34.61 dB; SSIM: 0.972).

For the empirical data, the similarity of results using single-average and 10-average reference data compared to the ground-truth data was highly similar (Tab. 4, MAE: 0.0203 vs. 0.0215, 5.9%; PSNR: 30.90 dB vs. 30.44 dB, -1.6%; SSIM: 0.930 vs. 0.925, -0.54%). The inter-result similarity was also high (MAE: 0.0097; PSNR: 36.75 dB; SSIM: 0.988).

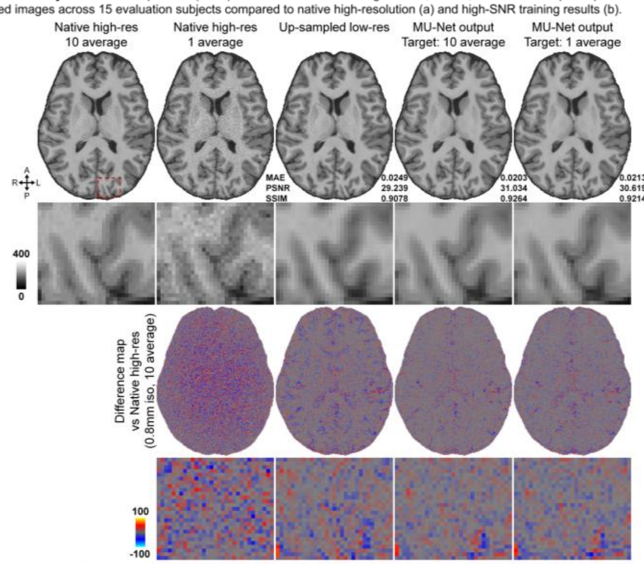

Figure 3. Empirical image results. Exemplary axial image slices from 10 average & single-average high-resolution, up-sampled, and super-resolved images with difference maps.

	(i) Up-sampled	(ii) 10 avg	(iii) 1 avg	(iv) 1 rep vs 10 avg
MAE	0.0247 ± 0.00246	0.0203 ± 0.00222	0.0215 ± 0.00225	0.0097 ± 0.00086
PSNR (dB)	29.25 ± 1.097	30.90 ± 1.105	30.44 ± 1.071	36.75 ± 1.036
SSIM	0.913 ± 0.01210	0.930 ± 0.01140	0.925 ± 0.01139	0.988 ± 0.00153

Table 4. Empirical image similarity metrics. Group mean (± std) of average MAE, PSNR, and SSIM compared to native high-resolution (i-iii) or 10 average target training results (iv) across 5 evaluation subjects.

Discussion: SRNR shows comparable efficacy of training using noisy versus high-SNR high-resolution reference data for super-resolution MRI. The success of the SRNR approach suggests that a smaller number of repetitions of high-resolution reference data for averaging can be adopted to achieve slightly compromised super-resolution performance and improve feasibility and accessibility. Using a single repetition also avoids additional data re-sampling during image co-registration, thus avoiding potential blurring of the high-resolution reference data.

References: [1] Pham. IEEE. 2017:197-200. [2] Chaudhari. MRM. 2018;80(5):2139-2154. [3]Chen. Springer. 2018:91-99. [4] Tian. Cerebral Cortex. 2021;31(1):463-482. [5] Lehtinen. arXiv preprint arXiv:1803.04189. 2018. [6] Glasser. Neuroimage. 2013;80:105-124. [7] Glasser. Nature neuroscience. 2016;19(9):1175-1187. [8] Reuter. Neuroimage. 2010;53(4):1181-1196. [9] Fischl. NeuroImage. 1999;9(2):195-207. [10] Dale. NeuroImage. 1999;9(2):179-194. [11] Fischl. NeuroImage. 2012;62(2):774-781.