MRI Single Image In-Plane Super Resolution Using Mixed-Scale Sense CNN

Kornelius Podranski
MPI-CBS, Leipzig, Germany *
kpodranski@cbs.mpg.de

Konstantin Thierbach
MPI-CBS, Leipzig, Germany *
kthierbach@cbs.mpg.de

Nikolaus Weiskopf
MPI-CBS, Leipzig, Germany *
weiskopf@cbs.mpg.de

Nico Scherf
MPI-CBS, Leipzig, Germany *
nscherf@cbs.mpg.de

Abstract

In this work we combine a mixed-scale dense convolutional network and a structure preserving loss function, to increase the in-plane resolution of MRI images with sub-millimeter resolution. Despite having 20 times fewer parameters than SRCNN, this architecture can reconstruct high-resolution images in comparable quality and learns to better preserve high-frequency details than networks trained with L^2-loss. Preliminary experiments show that the architecture in general is working well and comparable to the well known SRCNN without any optimization of hyper-parameters.

1 Introduction

Single image super resolution (SR) aims at reconstructing a higher resolution (HR) image from a single low-resolution (LR) image. It has different of applications in medical imaging and beyond. In the field of high-resolution brain MRI it can help to address some of the limitations related to long acquisition times and low signal-to-noise ratio [1].

Despite its usefulness the SR-problem is highly ill-posed, as there are many possible solutions for a given LR image. As shown by [2] deep learning is a promising way of selecting likely HR images from the solution space. Yet most implementations suggested so far lack good reconstruction of high frequency details, which is mostly due to using mean squared error as loss during training [3, 4].

In sub-millimeter brain-MRI these fine structures are particularly interesting and important for downstream processing and analysis (e.g. extracting cortical layers and vasculature). Therefore we need a SR method which reconstructs those structures best.

2 Methods

To achieve this goal we combined a mixed-scale dense convolutional neural network architecture [5] with a structure preserving loss function [3, 6] in a proof-of-principle study.

2.1 Network Architecture and Objective Function

The mixed-scale dense (MS-D) convolutional network uses dilated convolutions and densely connects all feature maps. Thus features at different scales are captured at the same time using fewer parameters

*Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

than fully convolutional networks architectures [5]. This way we will be able to create a robust network trainable with fewer labeled datasets and applicable to whole slice or even whole volume processing in the future.

To better preserve and restore high-frequency structures in the image, the network is trained using the weighted sum of multi-scale structural similarity metrics (MS-SSIM) and L^1-norm (MS-SSIM+L^1) as proposed by [3]. The MS-SSIM utilized is an approximation of structural similarity metrics (SSIM)[6] at five different scales. The mean absolute error is added to better preserve brightness and luminance [3]. Even though the metric was originally designed to mimic human visual perception, it specifically accounted for the structures of different scale contained in an MRI image.

2.2 Experiments

First experiments were performed training the network with downsampled image patches from the publicly available 'atlasing of the basal ganglia' (ATAG) consortium dataset [7] to recover image slices in the original resolution of 0.7 mm^2 per pixel. Low resolution samples were generated by blurring each slice with a Gaussian kernel ($\sigma = \sqrt{\lg_2(n)}$, with upscaling factor $n = \{2, 3\}$) and subsampling it by the upscaling factor [2]. In preparation of the SR-process each slice was upscaled again by the same factor using bilinear interpolation [3]. Training was performed on overlapping image patches of 33x33 pixels from 10 subjects. For evaluation image volumes were fed to the trained network slice by slice.

We compared performance of the network against the super resolution convolutional network (SR-CNN) [2], trained on the same set of patches and with the same objective function. We also compared the results to linear and bicubic upsampling without further processing. In a different experiment we trained the MS-D network using L^1- and L^2-norm as an objective function, to better understand the influence of the loss function.

3 Results

Preliminary results for two and three times super resolution show that most of the structures in the image were refined compared to bicubic interpolation. Comparisons of signal-to-noise ratio and structural similarity for the different reconstructions are given in Table 1. In our experiments peak signal-to-noise ratio (PSNR), SSIM [6] and MS-SSIM [8] were increased with regards to the input image and bicubic interpolation. With the tested settings SRCNN outperformed the MS-D network at 2x SR in terms of PSNR, but both networks achieved comparable levels of structural similarity. At 3x SR MS-D network predictions were slightly better than the other SR methods. SRCNN results were comparable to bicubic interpolation with PSNR closer to bilinear interpolation.

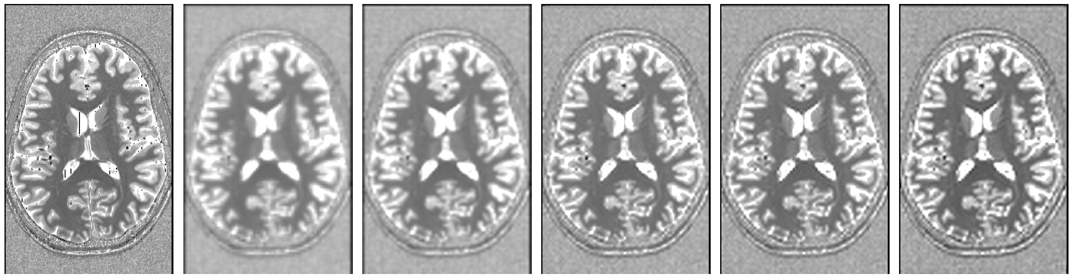
Table 1: Image metrics for different SR methods (networks trained with MS-SSIM+L^1-objective)

	2x SR			3x SR		
Method	PSNR [dB]	SSIM	MS-SSIM	PSNR [dB]	SSIM	MS-SSIM
bilinear	22.58	0.716	0.941	21.55	0.635	0.906
bicubic	23.19	0.757	0.956	22.19	0.684	0.926
SRCNN	24.89	0.853	0.980	21.71	0.681	0.926
MS-D-Net	23.77	0.822	0.968	22.54	0.739	0.952

In regards to MS-D network trained with different loss functions, visual inspection showed a notable increase in image structure for MS-SSIM+L^1, as can be seen in Figure 1. The calculated image metrics showed best performance for MAE, though.

4 Discussion and Conclusion

The first preliminary results are encouraging although the MS-D network is not performing better than SRCNN at 2x SR. At 3x SR instead SRCNN seems to fail to learn with structure preserving loss and the given hyperparameters (performance is comparable to bicubic upsampling, see Table 1). We



(a) (b) (c) (d) (e) (f)

Figure 1: Exemplary slice from validation dataset T1 map for 3x super resolution. Images from left to right show: (a) original HR image, (b) input data for prediction: bilinear upsampled LR image, (c) bicubic upsampled LR image and predictions from MS-D network trained with loss (d) L^1, (e) L^2 and (f) MS-SSIM+L^1.

think that the MS-D architecture is a promising candidate for robust MRI image processing, providing specifically the features needed for SR and other tasks related to multi-scale image structure.

The metrics reported here are biased towards the trained loss. To get a more meaningful and relevant criterion of reconstruction quality, we will compare the performance improvements in common postprocessing pipelines on the different SR results.

As a next step, we will optimize the MS-D network size and parameters to increase performance. Exploiting the considerably lower number of parameters of the MS-D architecture we will evaluate the benefits of processing larger chunks of 3D volumes directly.

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 616905.

References

[1] E. Plenge, D. H. J. Poot, M. Bernsen, G. Kotek, G. Houston, P. Wielopolski, L. van der Weerd, W. J. Niessen, and E. Meijering, "Super-resolution methods in MRI: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time?," *Magnetic Resonance in Medicine*, vol. 68, pp. 1983–1993, Nov. 2012.

[2] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in *Computer Vision – ECCV 2014*, Lecture Notes in Computer Science, (Zurich, Switzerland), pp. 184–199, Springer, Cham, Sept. 2014.

[3] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss Functions for Image Restoration With Neural Networks," *IEEE Transactions on Computational Imaging*, vol. 3, pp. 47–57, Mar. 2017.

[4] K. Hayat, "Super-resolution via deep learning," June 2017, arXiv:1706.09077.

[5] D. M. Pelt and J. A. Sethian, "A mixed-scale dense convolutional neural network for image analysis," *PNAS*, vol. 115, pp. 254–259, Jan. 2018.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, Apr. 2004.

[7] B. U. Forstmann, M. C. Keuken, A. Schafer, P.-L. Bazin, A. Alkemade, and R. Turner, "Multi-modal ultra-high resolution structural 7-Tesla MRI data repository," *Scientific Data*, vol. 1, p. 140050, Dec. 2014.

[8] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, vol. 2, (Pacific Grove, CA, USA), pp. 1398–1402, IEEE, Nov. 2003.