Motion correction in MRI with large movements using deep learning and a novel hybrid loss function

Lei Zhang[1], Xiaoke Wang[1], Michael Rawson[2], Radu Balan[3], Edward H. Herskovits[1], Linda Chang[1], Ze Wang[1], and Thomas Ernst[1]

[1]Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

[2]Department of Mathematics, University of Maryland, College Park, MD, USA

[3]Department of Mathematics and Center for Scientific Computation and Mathematical Modeling, University of Maryland, College Park, MD, USA,

Introduction:

Patient motion continues to be a major problem in MRI, despite tremendous advances using post-processing or prospective correction strategies[1,2], and more recently deep learning (DL).[3-5] While encouraging, these studies represent a direct application of existing network architectures to relatively minor motion artifacts, leaving much room for further optimizations. We propose a two-stage and multiple-loss based MRI motion correction network (MC-Net). In stage 1, MC-Net was trained with an L1-loss function to drive the network output to the reference (motion-free images). Stage 2 was then trained to minimize the L1 and Total Variation (TV) of the residual images (output-reference). The goal was to suppress overall motion artifacts which often affect tissue boundaries the most.

Methods

Simulation of Motion Corrupted Images

78 motion trajectories of 128 points each were synthesized from in-vivo head movements measured with PACE fMRI[1] (all <2mm/2°, each with 6DOF), and multiplied by eight. Trajectories were offset such that the k-space center remained in the origin point.

The source images were 3D MP-RAGE scans of 52 subjects (1mm isotropic; 256×256×160) without motion artifacts. For each of the 78 trajectories, the motion-corrupted k-space trajectory was calculated[2] (resolution 128×128×128) and sampled with non-uniform FFT (NuFFT)[6]. Ten sagittal slices were extracted from the resulting motion-corrupted 3D datasets (5mm increments;

total 780 slices per subject). Two simulations were performed: in-plane rotations only, and all in-plane parameters (translations, z-rotation). After removing boundary slices, the final dataset contained 41,750 slices (128×128). This was divided into a training set (35 subjects/28,800 slices), validation set (5 subjects/4,320 slices), and test set (12 subjects/8,630 slices).

MC-Net

Fig. 1 shows the architecture of MC-Net, which was derived from UNet[7]. The major novelty is the hybrid loss function L, which combines a L1-loss and Total Variation loss (TV)[8]:

$$\mathrm{L} = \mathrm{alpha} * \mathrm{L1} + \mathrm{beta} * \mathrm{TV} \quad (1)$$

$$TV = \sum_{i,j}((I(i+1,j) - I(i,j))^2 + (I(i,j+1) - I(i,j))^2)^{1.25} \quad (2)$$

where I is a corrupted image; i, j are row/column indices.

The second novelty is the two-stage training strategy. Stage 1 uses L1-loss only [set (alpha,beta)=(1,0)] to suppress the overall motion-induced artifact. The pre-trained stage 1 model is then fine-tuned in stage 2 by turning on the TV-loss component [(alpha, beta)=(1, 1)]; this penalizes boundary artifacts in addition to overall artifacts.

For comparison with the two-stage MC-Net (abbreviated "*L1+TV-ft"*), single-stage models trained with the L1-loss[5] and L1 + TV loss only were also implemented ("*L1"* and "*L1 + TV"* models). All models were trained in 100 epochs, using the Adam optimization method[9]. The structural similarity index measure (SSIM) with the "clean" image as reference[10] was used to quantify the performance of each method. The quality of images was also assessed visually by 2 experienced readers.

Results

Representative results of L1, L1+TV, and L1+TV-ft are shown in Fig.2 (rotation only artifact) and Fig.3 (rotation plus translation). All DL methods improved image quality, but the MC-Net was particularly effective for severe motion artifacts (Fig.3c).

Fig. 4 compares the SSIM (mean±std) for various models. The proposed 2-stage L1+TV-ft method improved the image quality slightly better than conventional approaches, both for the rotation only and rotation+translation models.

Fig. 5 shows that the SSIM decreases with the motion magnitude (standard deviation across 128 time points), from >0.8 for small motions (sub-mm/degree) to <0.5 for severe movements. The proposed 2-step L1+TV-ft approach improved image quality for motion magnitudes >1, but was especially effective for large rotation plus translation movements (Fig.4, center).

Discussion:

The MC-Net can correct motion artifacts in brain scans, in agreement with previous work[3-5]. We based our experiments on a customized UNet architecture[7] because it is widely used in medical image analysis. Also, the UNet architecture is complex enough to perform the motion correction task, but also simple enough to demonstrate the merits of the proposed hybrid loss. Implementation using the open-source framework Keras[11] with GPU acceleration might allow real-time motion correction. The proposed novel hybrid loss and two-stage training was slightly better in improving the final image quality than existing approaches, and was especially beneficial for correcting large rotational and translational motion artifacts.

Conclusion:

We propose a simulation framework and DL method for motion correction in MRI. Since the method is data-driven and independent of data acquisition or reconstruction, it may be suitable for routine clinical practice.

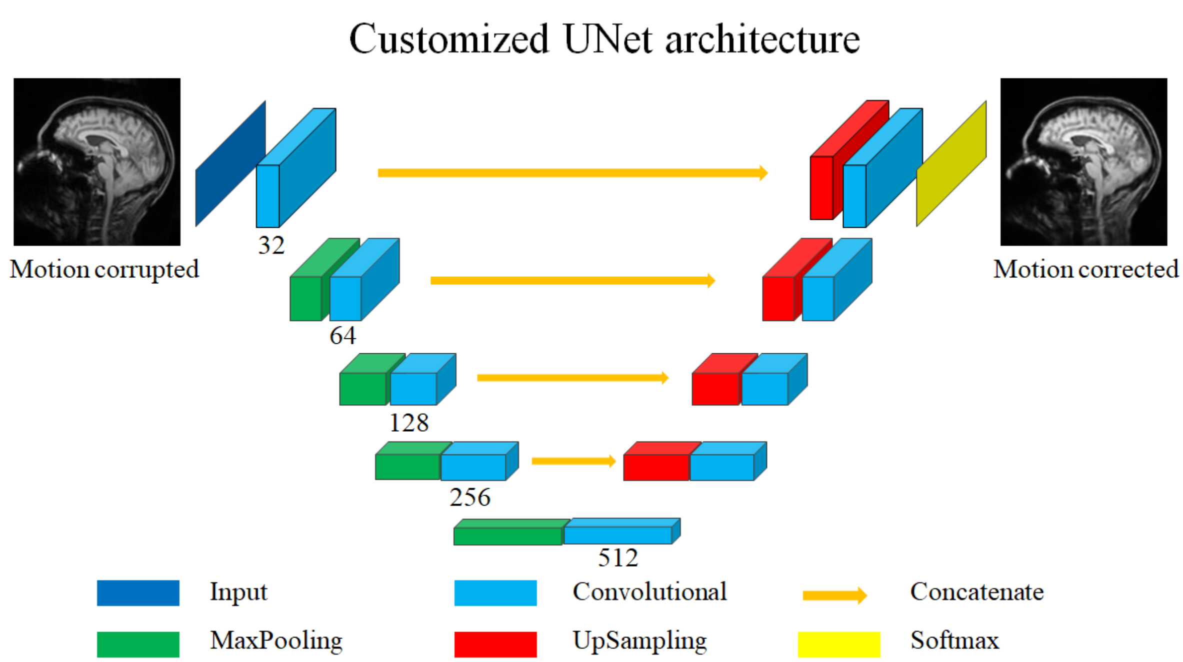

Figure 1. Customized UNet architecture.

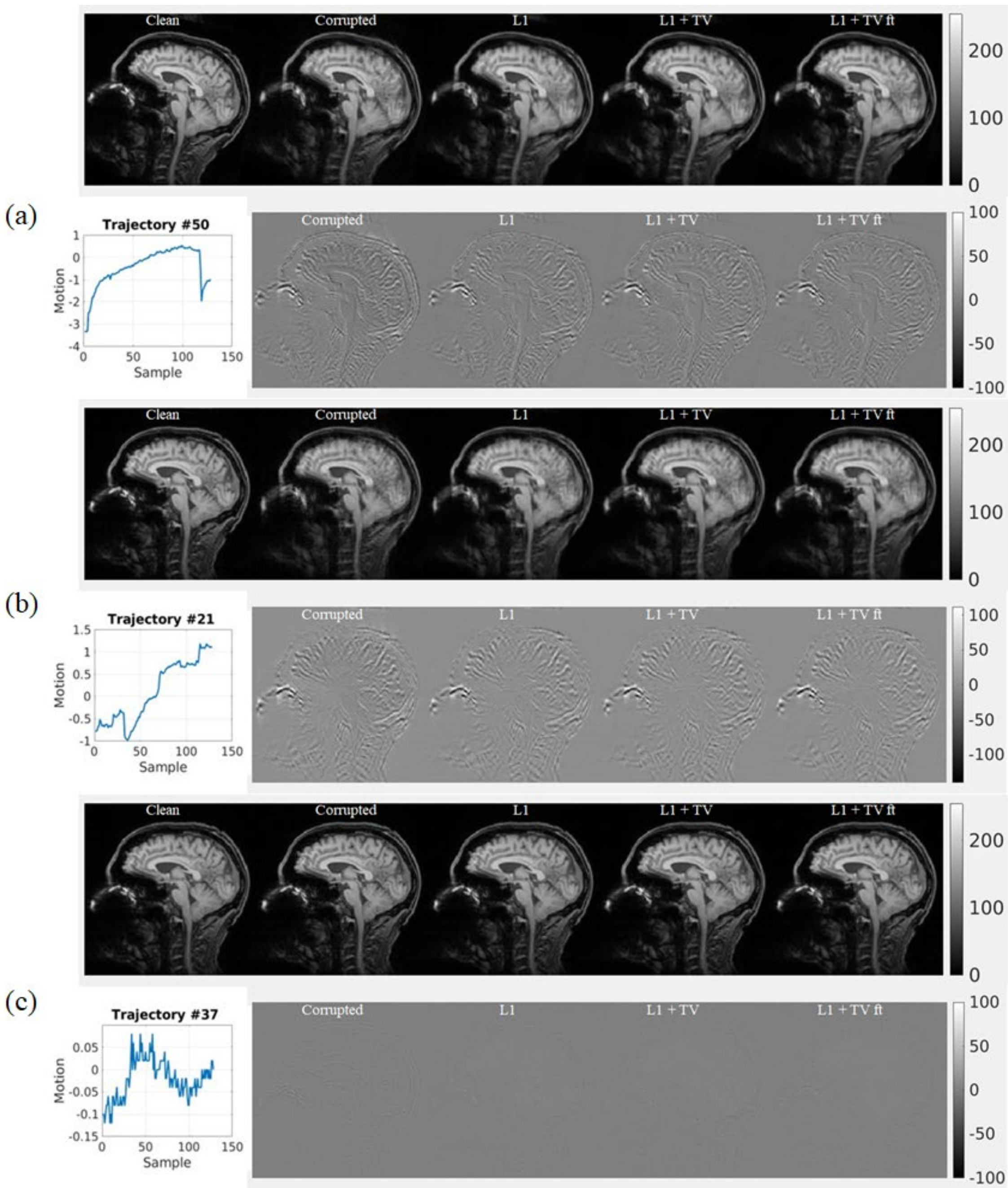

Figure 2: Rotation-only results. The first row of each subfigure shows the clean image, raw corrupted image, motion correction result of L1, motion correction result of L1+TV, and motion correction result of L1+TV-ft. The second row of each subfigure shows the motion trajectory, and the respective residual images for each condition relative to the clean image, respectively.

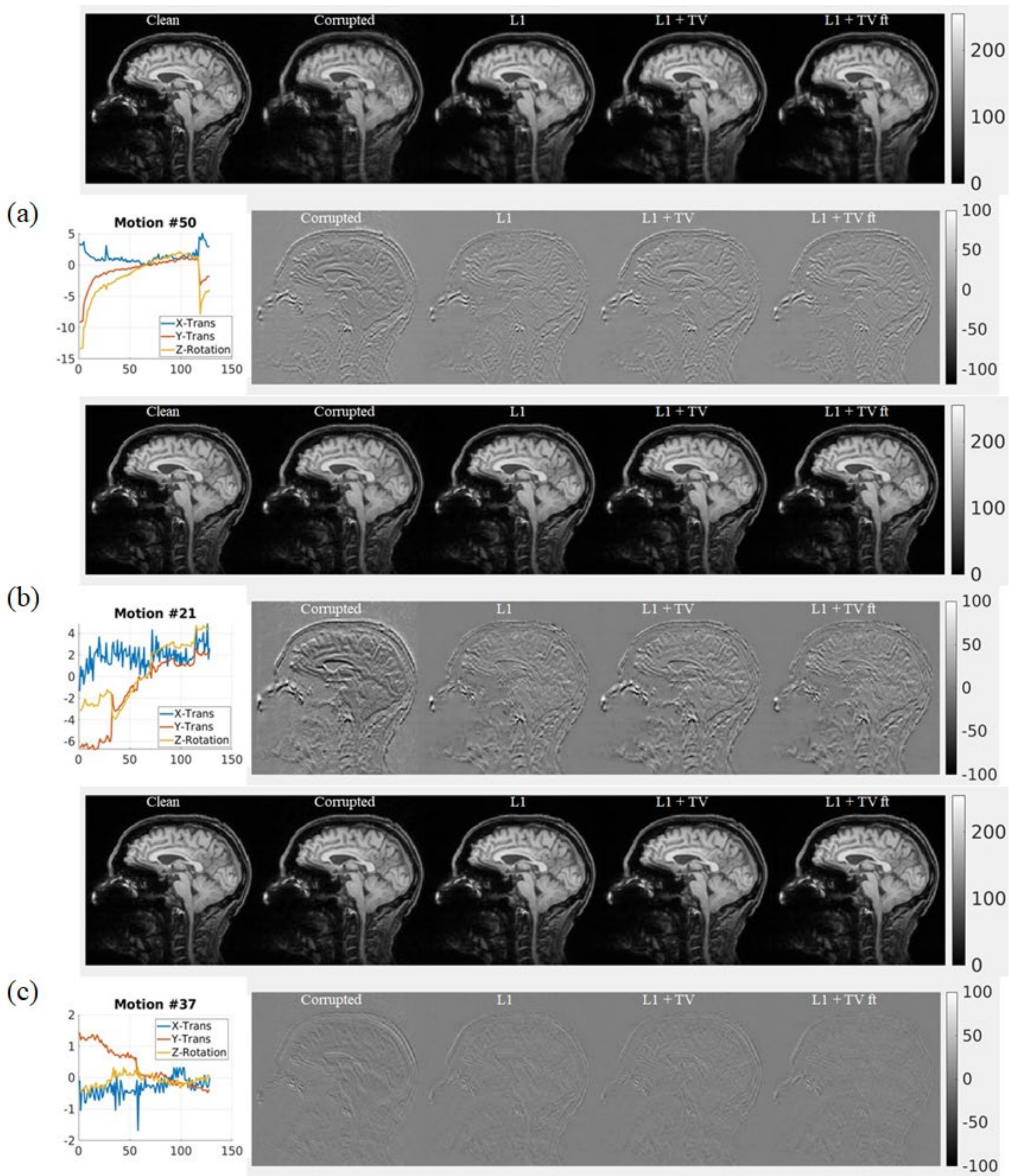

Figure 3: Rotation plus translation results. The first row of each subfigure shows the clean image, raw corrupted image, motion correction result of L1, motion correction result of L1+TV, and motion correction result of L1+TV-ft. The second row of each subfigure shows the motion trajectory, and the respective residual images for each condition relative to the clean image, respectively.

Motion	Corrupted image	L1	L1+TV	L1+TV -t
Rotation only	0.715 ± 0.093	0.738 ± 0.084	0.738 ± 0.085	0.743 ± 0.083
Rotation plus translation	0.574 ± 0.126	0.640 ± 0.077	0.642 ± 0.076	0.645 ± 0.076

Figure 4. SSIM (mean ± standard deviation) of the motion-corrupted images, and the output of the L1, L1 + TV, and L1+TV-ft methods, using "clean" source images as a reference. The proposed L1+TV-ft method performs slightly better than conventional approaches.

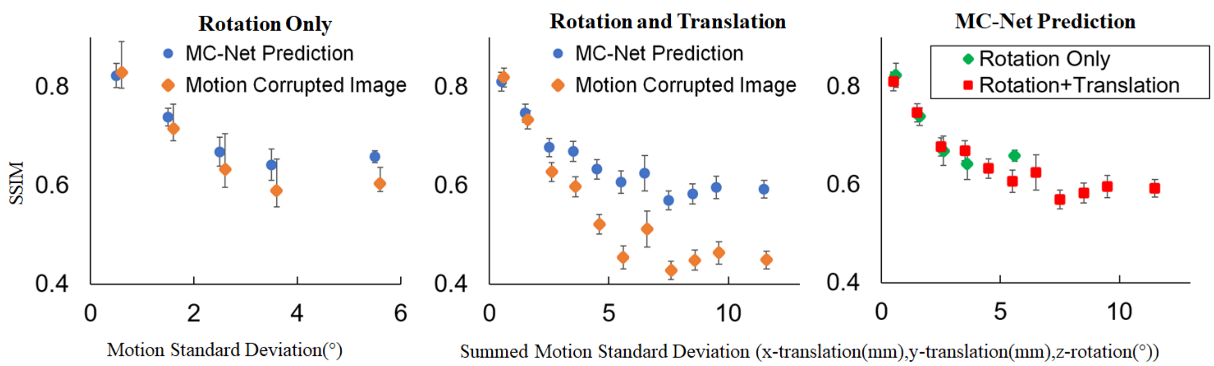

Figure 5: Dependence of the SSIM relative to "clean" reference images on the motion magnitude (standard deviation of motion across 128 time points) for the proposed MC-Net method. The SSIM provides a measure of image quality (0 = lowest, 1 = highest). The MC-Net successfully reduced the effect of motion for all conditions (left and center). The quality of MC-Net predication is similar for input images with only rotational motion and those with rotational plus translational motion (right).

Synopsis

Patient motion continues to be a major problem in MRI. We propose and validate a novel deep learning approach for the correction of large movements in brain MRI. Training pairs were generated using MRI data of high quality and corresponding synthesized images with artifacts based on real head movements. The images predicted by the proposed DL method from motion-corrupted data have improved image quality compared with the original corrupted images in terms of a quantitative metric and visual assessment by experienced readers.

Summary of Main Findings/Short Synopsis

We developed a novel deep learning approach for correction of large movements in brain MRI. The proposed method improved image quality compared with the motion corrupted images in terms of a quantitative metrics and visual assessment by experienced readers.

Acknowledgements

This work was supported by NIH grants 1R01 DA021146 (BRP) and NIH/NIA R01AG060054.

References:

1. Thesen S, Heid O, Mueller E, Schad LR. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magnetic Resonance in Medicine*. 2000;44(3):457-465.
2. Zahneisen B, Ernst T. Homogeneous coordinates in motion correction. *Magnetic Resonance in Medicine*. 2016;75(1):274-279.
3. Sommer, K., Saalbach, A., Brosch, T., Hall, C., Cross, N. M., & Andre, J. B. (2020). Correction of motion artifacts using a multiscale fully convolutional neural network. *American Journal of Neuroradiology*, 41(3), 416-423.
4. Usman, M., Latif, S., Asim, M., Lee, B. D., & Qadir, J. (2020). Retrospective motion correction in multishot MRI using generative adversarial network. *Scientific Reports*, 10(1), 1-11.
5. Pawar, K., Chen, Z., Shah, N. J., & Egan, G. F. (2019). Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation. *NMR in Biomedicine*, e4225.
6. Greengard, L., & Lee, J. Y. (2004). Accelerating the nonuniform fast Fourier transform. *SIAM review*, 46(3), 443-454.
7. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
8. Chollet, F. (2017). Deep Learning with Python.
9. Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. international conference on learning representations (2015).

10. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612.
11. Chollet, F., & others. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras