FULL PAPER

Respiratory motion correction for free-breathing 3D abdominal MRI using CNN-based image registration: a feasibility study

[1]JUN LV, PhD, [2]MING YANG, PhD, [1,3]JUE ZHANG, PhD and [1,4]XIAOYING WANG, MD

[1]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China
[2]Vusion Tech Ltd. Co, Suzhou, China
[3]College of Engineering, Peking University, Beijing, China
[4]Department of Radiology, Peking University First Hospital, Beijing, China

Address correspondence to:
Jue Zhang
E-mail: *zhangjue@pku.edu.cn*
Xiaoying Wang
E-mail: *cjr.wangxiaoying@vip.163.com*

Objective: Free-breathing abdomen imaging requires non-rigid motion registration of unavoidable respiratory motion in three-dimensional undersampled data sets. In this work, we introduce an image registration method based on the convolutional neural network (CNN) to obtain motion-free abdominal images throughout the respiratory cycle.

Methods: Abdominal data were acquired from 10 volunteers using a 1.5 T MRI system. The respiratory signal was extracted from the central-space spokes, and the acquired data were reordered in three bins according to the corresponding breathing signal. Retrospective image reconstruction of the three near-motion free respiratory phases was performed using non-Cartesian iterative SENSE reconstruction. Then, we trained a CNN to analyse the spatial transform among the different bins. This network could generate the displacement vector field and be applied to perform registration on unseen image pairs. To demonstrate the feasibility of this registration method, we compared the performance of three different registration approaches for accurate image fusion of three bins: non-motion corrected (NMC), local affine registration method (LREG) and CNN.

Results: Visualization of coronal images indicated that LREG had caused broken blood vessels, while the vessels of the CNN were sharper and more consecutive. As shown in the sagittal view, compared to NMC and CNN, distorted and blurred liver contours were caused by LREG. At the same time, zoom-in axial images presented that the vessels were delineated more clearly by CNN than LREG. The statistical results of the signal-to-noise ratio, visual score, vessel sharpness and registration time over all volunteers were compared among the NMC, LREG and CNN approaches. The SNR indicated that the CNN acquired the best image quality (207.42 ± 96.73), which was better than NMC (116.67 ± 44.70) and LREG (187.93 ± 96.68). The image visual score agreed with SNR, marking CNN (3.85 ± 0.12) as the best, followed by LREG (3.43 ± 0.13) and NMC (2.55 ± 0.09). A vessel sharpness assessment yielded similar values between the CNN (0.81 ± 0.03) and LREG (0.80 ± 0.04), differentiating them from the NMC (0.78 ± 0.06). When compared with the LREG-based reconstruction, the CNN-based reconstruction reduces the registration time from 1h to 1min.

Conclusion: Our preliminary results demonstrate the feasibility of the CNN-based approach, and this scheme outperforms the NMC- and LREG-based methods.

Advances in knowledge: This method reduces the registration time from ~1h to ~1min, which has promising prospects for clinical use. To the best of our knowledge, this study shows the first convolutional neural network-based registration method to be applied in abdominal images.

INTRODUCTION

Respiratory motion is a significant source of error for the MRI of the upper abdomen. The scans are commonly performed during a breathholding period, and a healthy adult can hold his/her breath for approximately 20~30 s,[1] which could limit the image quality, resolution and coverage.[2] In addition, it is difficult for them to hold their breath if they are critical or paediatric patients. For free-breathing acquisition, respiratory gating is commonly used to

scan. However, it only accepts data at end-expiration, minimizing the motion artefacts at the expense of additional scan time. Additionally, respiratory gating does not always perform well if the respiratory rhythm is irregular.[3] To overcome this problem, navigator echoes are integrated into the imaging sequence.[4–8] However, the drawbacks of the navigator include not only the low efficiency of the scan[9] but also the ignorance of the time difference between the motion and navigator acquisition. Recently, self-navigation techniques[10,11] have been proposed to extract the respiratory motion signal[9,12,13] from the acquired data itself.

According to the respiratory signal, each spoke is related to the breathing position at which it was acquired. The respiratory signal is discretized into a set of bins, which allows the reconstruction of near motion-free images in each bin. One high-quality datum is acquired at an end-expiratory acceptance window which is called B_{HQ}. Then, several lower quality images which are obtained at reaming respiratory positions will be registered to B_{HQ}. Thus, an accurate registration method has great meaning for facilitating an accurate diagnosis.

Former researches[14,15] has usually used the hierarchical adaptive local affine registration method (LREG).[16] LREG is an optimization-based method, in which the transformation parameters are iteratively updated to optimize an objective function that reflects the accuracy of the registration. Thus, LREG requires a high computational cost and could not acquire registration data in real-time, which significantly limits its clinical application. Recently, the effectiveness of CNN has been shown in a wide range of medical image processing tasks, such as left ventricle[17]/brain[18–21]/prostate[22] segmentation, disease classification[23–25] and registration[26,27] tasks. However, to the best of our knowledge, it has not been applied to the abdominal image registration. To address these problems, we adopted a convolutional neural network (CNN) regression approach for real-time registration. The whole structure of our registration network consists of a CNN (ConvNet) regressor, a spatial transformer, and a resampler, as in.[27] The network takes pairs of fixed (B_{HQ}) and moving (B_{LQ}) images as inputs, and it outputs moving images that are warped to the fixed images.

We believe that the strong non-linear modelling capability of the CNN will directly estimate the transformation parameters among the images of different bins. The registration process will be accomplished within submilliseconds. To demonstrate the feasibility of our introduced approach, we compared the CNN against a non-motion corrected (NMC) method and the LREG-based method using the measures of signal-to-noise ratio (SNR), assessment of image sharpness, visual image quality rated by two experienced radiologists and registration time.

METHODS AND MATERIALS
Data acquisition
27 healthy volunteers (21 males, 6 female, 36 years) took part in this experiment and informed consent was obtained from each participant. The participants were told to keep still and breathe normally. The study was approved by the local ethical review committee. All of the scans were performed on a clinical 1.5 T MRI system (Ingenia, Philips Healthcare, Best, Netherlands) equipped with a 16-channel anterior coil and a 16-channel posterior coil. MR data acquisition was performed using a three-dimensional (3D) golden angle-radial stack-of-stars (SOSs)[28] sequence (Figure 1a). The relevant imaging parameters were as follows: slice thickness = 3 mm with over contiguous sampling; flip angle = 10 degrees; field of view = $450 \times 450 \times 249$ mm²; sense factor along z = 1.41; number of read-out points in each spoke = 400 with two times oversampling; spatial resolution = $1.00 \times 1.00 \times 3.00$ mm³; and repetition time (TR)/echo time (TE) = 4.88/2.06 ms. A total of 751 spokes were acquired for each partition, with a total scan time of 216.22 s. Another group of acquisitions with the higher F–H resolution doubled was also applied to the same volunteers. All of the other parameters of these scans were identical.

Data binning
In our work, an adaptive method was used to estimate the respiratory signal. The one-dimensional fast Fourier transform along the feet–head (FH) direction to the centre k-space profiles was applied to compute the projection profiles of the 3D volume. The respiratory motion detection was performed by first aligning the projection profiles into a two-dimensional (2D) image (Figure 1b), for which no *a priori* respiratory training phase was required. Following by envelope extraction of the image, the respiratory motion signal would be acquired (Figure 1c). Then, the continuously acquired golden-angle radial data sets were divided into three respiratory bins (Figure 1d), in which the spokes were in the same motion position, similar to a previous work.[13]

Reconstruction of the respiratory bins
Reconstruction was developed and performed using the Recon2.0 platform by Philips. 3D undersampled data (Figure 1d) for each bin were reconstructed by gridding[29] followed by SENSE to unfold the warped image along the F–H direction. It took ~80 s to reconstruct a 3D volume using a workstation equipped with 16 GB DDR RAM and 2 intel XEON E5-1620 CPUs.

Motion modelling
Unlike Buerger et al[9] where the LREG tool[16] was directly used to register reconstructed bin images to a common respiratory position, here, we proposed to introduce a CNN-based registration algorithm (Figure 1e).

Figure 2 illustrates the whole structure of the convolutional neural network. First, the network takes concatenated pairs of moving (Bin2/Bin3) and fixed (Bin1) images as its input, which is then followed by five hidden layers: (1) convolutional layer with 64 3×3 filters; (2) max-pooling layer of 2×2; (3) convolutional layer with 128 3×3 filters per map; (4) convolutional layer of 128 maps of filter size 3×3 ; and (5) max-pooling layer of 2×2. Finally, the output layer includes two kernels that indicate the 2D displacement of the input image pairs.

Next, as shown in Figure 1, the spatial transformer generates a dense displacement vector field that will be used to warp the moving image to the fixed image. Since abdominal motion

Figure 1. The process used to reconstruct the free-breathing 3D MRI data set. (a) Golden-angle radial SOS trajectory. (b) Projection profiles of the 3D volume derived from the K-space centre [central line in (a)]. (c) Respiratory signal is binned into three bins. (d) According to the respiratory position, the corresponding K-spaces are obtained for each bin. Each of them is reconstructed separately. (e) The CNN Net estimates the spatially corresponding image patches from the moving (Bin2/Bin3) and fixed images (Bin1). (f) The spatial transformer generates the DVF which will be used to warp the moving image to the fixed image. (g) The network is trained by back propagating a similarity metric as a cost function.3D, three-dimensional; CNN, convolutional neural network; DVF, displacement vector field; SOS, stack-of-star.
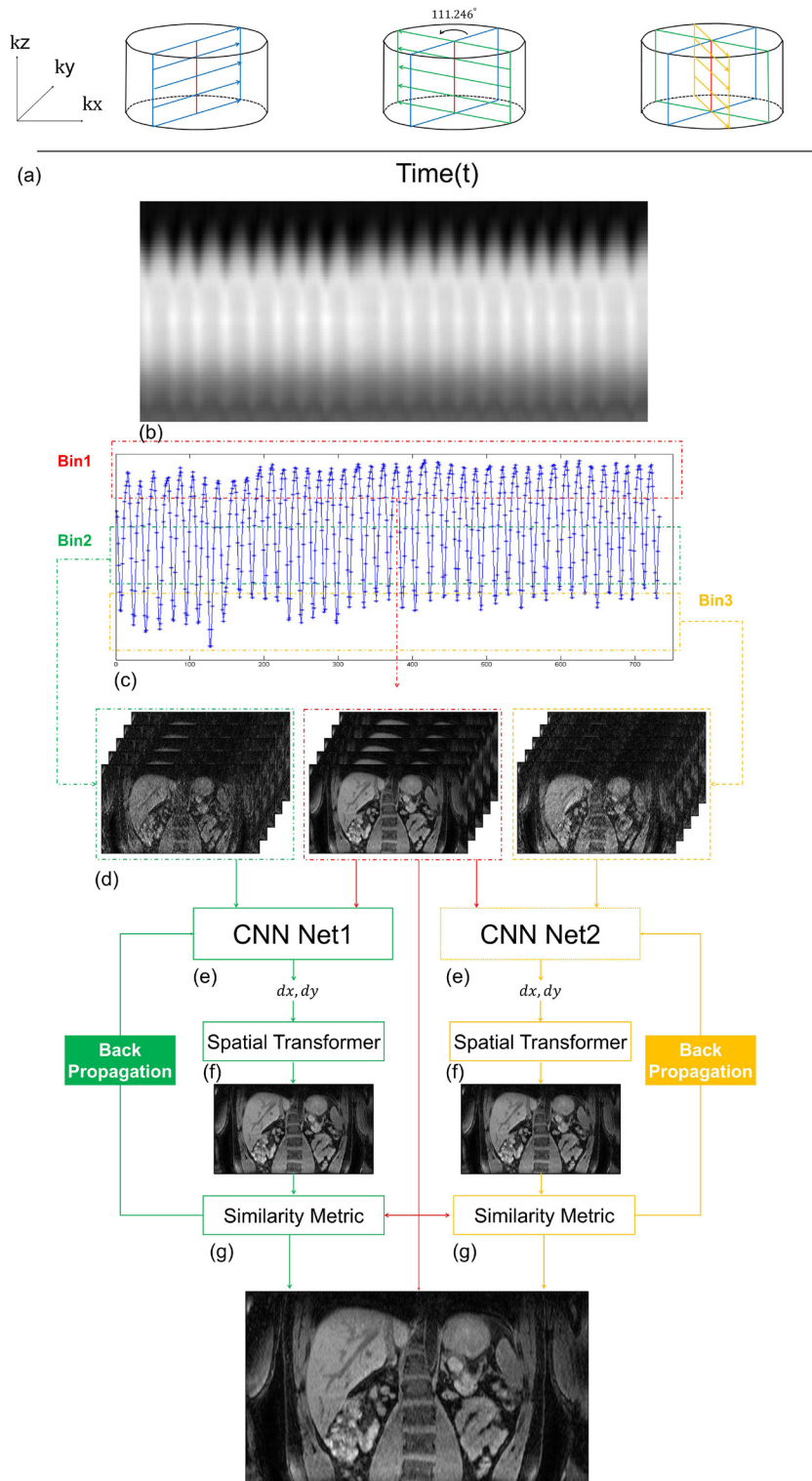
Figure 2. 2D convolutional neural network structure. The network takes two 2D patches from moving (Bin2/Bin3) and fixed (Bin1) images at the same location and generates two 2D momentum predictions of the patches in the *x* and *y* spatial directions. C: 2D convolution layer. S: pooling layer. Parameters for the C and S layers: number and size of the filter kernel. 2D, two-dimensional.
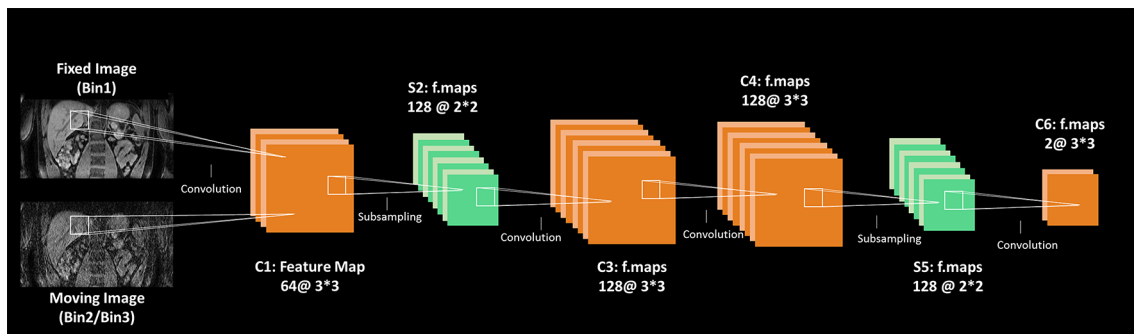
leads to a large amount of local variation, we adopted a cubic B-spline[30] transformer.

In the training stage, the normalized cross-correlation (NCC) between pairs of moving and fixed images using minibatch stochastic gradient descent (Adam)[31] was implemented as a cost function that must be optimized.

Therefore, NCC was adopted as our cost function while performing the actual registration in this study. The NCC of the fixed image and the moving image is defined by

$$\text{NCC}\left(I_1, I_2\right) = \sum_x \sum_y \frac{\left(I_1\left(x,y\right) - \overline{I_1}\right)\left(I_2\left(x,y\right) - \overline{I_2}\right)}{\sigma_{I_1}\sigma_{I_2}}$$

where I_1 is the fixed image, I_2 is the moving image, and σ_{I_1} and σ_{I_2} are the standard deviations of I_1 and I_2, respectively. After training, the network can be applied to the registration of the test images.

Network training and testing

To evaluate the effects of various sizes of image patches, several additional experiments were conducted. Experiment 1 was designed with overlapping patches of size 28×28; Network 2 used image patches of 64×64; and Network 3 analysed the full image as input. We performed the registration in the coronal view for each experiment. Thus, we trained two networks in all. When referring to the motion in the axial and sagittal view, we used only the 2D affine registration.

We trained each of the networks using a data set of 2490 images with size of 448×448. We randomly split the data set into the training set (80% of the original data set) and test set (20% of the original data set). During the training stage, the moving and the reference image patch were randomly selected, which corresponded to slices of the same subject but belonging to different bins. Each CNN was trained until convergence with batches of 256 image pairs in 2000 iterations. The training took approximately 18, 6, and 2 h on NVIDIA GPUs (GTX 1080), separately. The filter weights of each layer were initialized by drawing randomly from a Gaussian distribution with zero mean and standard deviation 0.001 (and 0 for biases). In this study, Tensorflow[32] was used for implementing this model with a learning rate of 10–4 on a Linux machine (64-bit Ubuntu 14.04 LTS; Cuda 7.5).

Statistical analysis

The introduced method was compared with the NMC approach and reconstruction framework of Buerger's [9] in terms of the SNR, image sharpness, and visual image scoring by two experienced radiologists with 12 and 10 years of experience in clinical MRI interpretation and registration time.

As seen in Figure 3, the SNR was calculated based on two separate regions of interest from a single image: one (the rectangle at the top left) in the tissue to determine the signal intensity and the other (the rectangle at the bottom left) in the image background to measure the noise intensity.[33,34] The image sharpness was estimated using the CoroEval software.[35] The vessel sharpness was measured based on 25 manually defined 1D intensity profiles, similar to Buergers' works[13,36] (Figure 3). The overall vessel sharpness was determined as the mean sharpness over each of the 25 selected profiles. The two radiologists were asked to "score the sharpness of the main boundaries and features of the images" on a scale of 0 (extreme blurring) to 4 (no blurring). The scores by the two observers were averaged for evaluation. The SNR values, vessel sharpness and registration time were compared using repeated measures analysis of variance (*post hoc* Bonferroni correction) with Greenhouse-Geisser correction, which was used to test for differences between the reconstruction methods.

Figure 3. Coronal reconstruction image from Subject 1. The rectangles indicate the signal (top left) and noise (bottom left) intensity, respectively. The 1D intensity profiles (three slashed lines) were manually defined to estimate the vessel sharpness. 1D, one-dimensional.
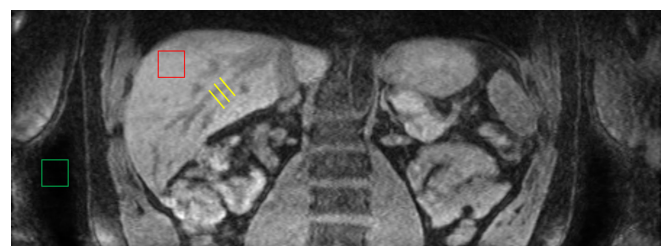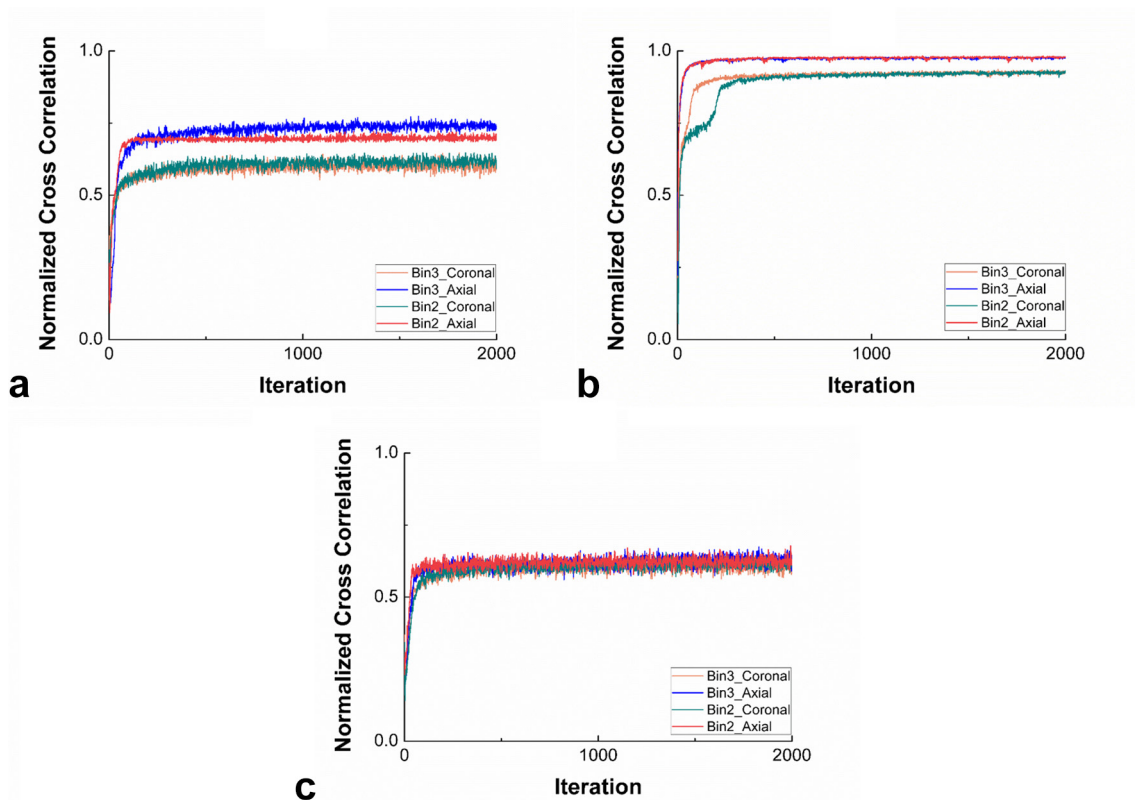
Figure 4. Normalized cross-correlation value over 2000 iterations for the three experiments. (a) Image patch of size 28. (b) Image patch of size 64. (c) Whole image as input.

The visual score measurements were evaluated using a Wilcoxon signed-rank test, with a significance threshold of $p < 0.05$.

RESULTS

Figure 4 shows the normalized cross-correlation value over 2000 iterations during training for the three experiments. All the networks converged quickly at first and then stabilized, which shows that the NCC loss of all of the models converges as the number of iterations increases. However, each network has a different NCC value. The network with the 64×64 patch size (Figure 4b) has the highest performance, followed by the 28×28 patch size (Figure 4a) and the whole image network (Figure 4c).

As seen in Figure 5, the longer dashed line delineates the upper edge of the liver in Bin1, and the shorter dashed line represents the upper edge of the liver in each bin. Before registration, the locations of the shorter dashed lines have large variations among the different bins. However, after registration, Bin2 and Bin3 have nearly been corrected to almost the same place.

All the bins were combined to form a high-quality composite image. Multiple slice orientations for the NMC, LREG-based reconstruction and CNN-based reconstruction for volunteers 1, 6, 8 are shown in Figure 6. In Figure 6, it can be seen that clear noise and the obscure edge of the liver are still remaining in the

Figure 5. The corresponding undersampled image reconstructions of pre- and post-registration are shown in a coronal view.
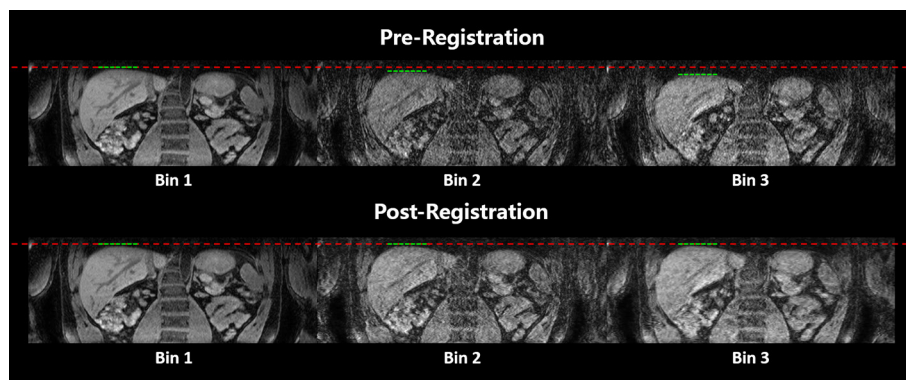
Figure 6. Coronal (top), sagittal (middle) and axial (bottom) slices for volunteers 1, 6 and 8 (including zoom-in images; arrows note some main differences). (a, d, g, j, m, p) NMC (non-motion corrected): Several structures in the image are corrupted by motion. (b, e, h, k, n, q) LREG: Some structures appear sharper than NMC, but presents vessel broken and distorted. (c, f, i, l, o and r) CNN: a sharper reconstruction is obtained. CNN, convolutional neural network; LREG, local affine registration method; NMC, non-motion corrected.
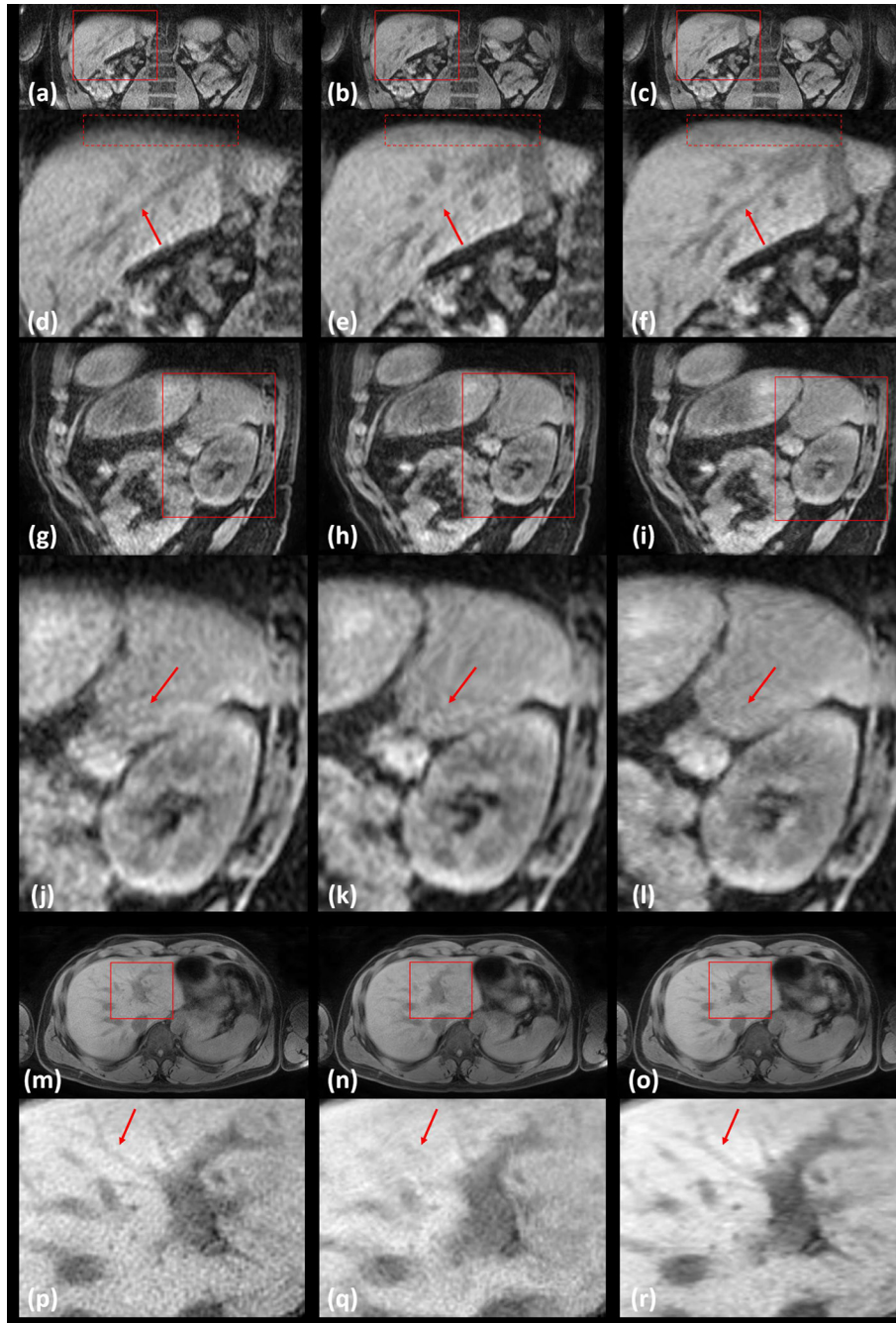
NMC reconstruction. While the LREG and CNN yield images with similar quality, they reduce most of the noise presented in the NMC reconstruction. However, the coronal images indicate that LREG has caused broken blood vessels, while the vessels of CNN are sharper and more consecutive, as highlighted by the red arrows in Figure 6b,c. As shown in the sagittal view, LREG has distorted and blurred the liver contours compared with NMC and LREG (indicated with red arrows). At the same time,

zoom-in axial images present that vessels (indicated with the red arrow) were delineated more clearly by CNN than LREG.

The vessel sharpness and registration time both differed significantly among the methods when tested with repeated measures analysis of variance with the Greenhouse-Geisser correction ($p = 0.013 < 0.05$ and $p = 2.6e-4 < 0.001$, respectively). *Post hoc* testing with the Bonferroni correction was significant for the

Table 1. Comparison between the NMC reconstruction, the LREG tool-based reconstruction and our introduced method (CNN)

	NMC	LREG	CNN
SNR	116.67 ± 44.70	187.93 ± 96.68	207.42 ± 96.73
Visual score	2.55 ± 0.09	3.43 ± 0.13	3.85 ± 0.12
Vessel sharpness	0.78 ± 0.06	0.80 ± 0.04	0.81 ± 0.03
Registration time	0.23 ± 0.02 s	67.14 ± 23.03 min	1.09 ± 0.21 min

CNN, convolutional neural network; LREG, local affine registration method; NMC, non-motion corrected.

comparison of vessel sharpness as measured by LREG compared with CNN ($p = 0.003 < 0.05$) and all of the registration time comparisons ($p = 1.98e-6$ and $2.36e-6 < 0.001$, respectively), with the exception of the measurement that compared NMC with CNN ($p = 1$). Higher visual scores were obtained using CNN compared with LREG, and comparisons yielded statistically significant differences ($p = 0.03 < 0.05$). Although there was a general trend towards improved SNR values using the CNN compared with the LREG, no statistically significant differences were obtained. When compared with the LREG-based reconstruction, the CNN-based reconstruction remarkably reduces the registration time, from ~1 h to ~1 min ($p = 2.36e-6 < 0.001$).

DISCUSSION

In this study, we introduced a novel registration method based on the CNN to improve the alignment of the images derived from different bins. First, three highly undersampled images that cover the whole respiratory signal were reconstructed. Second, the CNN-based registration approach was employed to combine all three bin images. This technique was validated on 10 healthy subjects and was compared with the NMC method as well as the LREG method.

The training loss converged well and quickly in each experiment, and the network with image patches of 64×64 shows the highest NCC value. Kervrann et al[37] demonstrated that image patch-size selection is a complex problem, since it depends on the image contents. Zhao et al indicated that a larger patch ($65 \times 65 \times 3$) provided more information and helped improve the network's performance. During the forward pass of a pooling layer, although an image patch-size of 28 can represent fine details,[38] it might neglect the relationships between the local textures. Additionally, a network that uses the whole image as input will not catch up the textures of the image pairs, which leads to the poorest result. However, in the case of patches with the larger size of 64, it can take care of the local geometries and maintain global consistency.[38] Thus, this type of size of patch image is fully adaptive for our image registration. Remarkably, as can be observed in Figure 5, with the proposed method, the Bin2 and Bin3 images at different positions of the breathing cycle have been registered well to the high-quality Bin1 image. To train the network in the coronal position is reasonable, since the principal respiratory motion predominately occurs in the FH direction.[39] Considering the image quality, obviously, the CNN registration method gave significantly better results than the NMC and LREG methods. The high score of the SNR in the CNN and LREG indicates that non-rigid registration can remove misalignments induced by noise and blurring of the liver in both

the centre position and edge location. The vessel sharpness values of the CNN and LREG are similar, and they are distinguished as superior to the NMC. The inaccuracy of the local registration is responsible for the broken and distorted vessels in the LREG group. The visual quality of the images of different groups was evaluated by two radiologists. Both radiologists scored the CNN group higher than the NMC and LREG groups. These results make sense because the purpose of using the CNN network is to learn discriminative features that represent complex morphological features in image patches accurately and concisely. This deep network is working to acquire the non-linear relationship between the moving and fixed image patch: first, it will learn the hierarchical image patterns by seeking simple features. Then, from the previous layers, it could acquire more complex representations.[26] In addition, Wu et al[26] indicated that the trained deep learning network selected features that more accurately capture the complex morphological patterns in the image patches. Thus, this approach resulted in better anatomical correspondences, which ultimately improved the image registration performance. In regard to the registration time, the proposed method acquired remarkably higher computational efficiency than the LREG method in that the execution time was reduced from ~1 h to ~1 min. This finding can be explained by the fact that in the testing state, our method does not require any training images. The CNN-based image registration is executed with the well-trained model in one pass.[40–42] However, LREG must solve these complex optimization problems iteratively.

Our study has several limitations. One limitation is that the introduced method is based on 2D images, which transforming the 3D volume one by one to a reference image using an NCC metric. Thus, to extend the applicability of our method, registration based on the 3D network will be investigated. If we can adopt a 3D network architecture, we will train the network only once. The network takes two 3D patches from the moving and target image as the input, and outputs three 3D initial momentum patches (one for each of the x,y,z dimensions). Another limitation is the lack of real data sets from patients. Different results might occur when performing the same testing and training steps in a group of patient subjects. Future studies will include the use of real patients with different lesions.

CONCLUSIONS

In conclusion, a CNN-based approach for real-time registration has been introduced to improve the alignments of the images. Our preliminary results demonstrate that the feasibility of the CNN-based method and that this scheme outperforms the NMC- and LREG-based methods. In addition, this method

reduces the registration time from ~1 h to ~1 min, which has promising prospects for clinical use. To the best of our knowledge, this study is the first CNN-based registration method to be applied in abdominal images.

As shown in Table 1, the vessel sharpness and registration time both differed significantly among the methods when tested with repeated measures ANOVA with the Greenhouse-Geisser correction ($p = 0.013 < 0.05$ and $p = 2.6\text{e}{-}4 < 0.001$, respectively).

REFERENCES

1. Bernstein MA, King KF, Zhou XJ. *Handbook of MRI pulse sequences.* Burlington MA: Elsevier Google Scholar; 2004.

2. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. *J Magn Reson Imaging* 2015; **42**: 887–901. doi: https://doi.org/10.1002/jmri.24850

3. Chavhan GB, Babyn PS, Vasanawala SS. Abdominal MR imaging in children: motion compensation, sequence optimization, and protocol organization. *Radiographics* 2013; **33**: 703–19. doi: https://doi.org/10.1148/rg.333125027

4. Lauzon ML, Mahallati H, Frayne R. Time-efficient breath-hold abdominal MRI at 3.0 T. *AJR Am J Roentgenol* 2006; **187**: 649–57. doi: https://doi.org/10.2214/AJR.05.1157

5. Lin W, Huang F, Börnert P, Li Y, Reykowski A. Motion correction using an enhanced floating navigator and GRAPPA operations. *Magn Reson Med* 2010; **63**: 339–48. doi: https://doi.org/10.1002/mrm.22200

6. Nielles-Vallespin S, Mekkaoui C, Gatehouse P, Reese TG, Keegan J, Ferreira PF, et al. In vivo diffusion tensor MRI of the human heart: reproducibility of breath-hold and navigator-based approaches. *Magn Reson Med* 2013; **70**: 454–65. doi: https://doi.org/10.1002/mrm.24488

7. van der Kouwe AJ, Benner T, Dale AM. Real-time rigid body motion correction and shimming using cloverleaf navigators. *Magn Reson Med* 2006; **56**: 1019–32. doi: https://doi.org/10.1002/mrm.21038

8. Welch EB, Manduca A, Grimm RC, Ward HA, Jack CR. Spherical navigator echoes for full 3D rigid body motion measurement in MRI. *Magn Reson Med* 2002; **47**: 32–41. doi: https://doi.org/10.1002/mrm.10012

9. Buerger C, Clough RE, King AP, Schaeffter T, Prieto C. Nonrigid motion modeling of the liver from 3-D undersampled self-gated golden-radial phase encoded MRI. *IEEE Trans Med Imaging* 2012; **31**: 805: 805: 15. doi: https://doi.org/10.1109/TMI.2011.2181997

10. Stehning C, Börnert P, Nehrke K, Eggers H, Stuber M. Free-breathing whole-heart coronary MRA with 3D radial SSFP and self-navigated image reconstruction. *Magn Reson Med* 2005; **54**: 476–80. doi: https://doi.org/10.1002/mrm.20557

11. Welch EB, Rossman PJ, Felmlee JP, Manduca A. Self-navigated motion correction using moments of spatial projections in radial MRI. *Magn Reson Med* 2004; **52**: 337–45. doi: https://doi.org/10.1002/mrm.20151

12. Cruz G, Atkinson D, Buerger C, Schaeffter T, Prieto C. Accelerated motion corrected three-dimensional abdominal MRI using total variation regularized SENSE reconstruction. *Magn Reson Med* 2016; **75**: 1484–98. doi: https://doi.org/10.1002/mrm.25708

13. Buerger C, Prieto C, Schaeffter T. Highly efficient 3D motion-compensated abdomen MRI from undersampled golden-RPE acquisitions. *MAGMA* 2013; **26**: 419–29. doi: https://doi.org/10.1007/s10334-013-0370-y

14. Buerger C, Clough RE, King AP, Schaeffter T, Prieto C. Nonrigid motion modeling of the liver from 3-D undersampled self-gated golden-radial phase encoded MRI. *IEEE Trans Med Imaging* 2012; **31**: 805: 805: 15. doi: https://doi.org/10.1109/TMI.2011.2181997

15. Cruz G, Atkinson D, Buerger C, Schaeffter T, Prieto C. Accelerated motion corrected three-dimensional abdominal MRI using total variation regularized SENSE reconstruction. *Magn Reson Med* 2016; **75**: 1484–98. doi: https://doi.org/10.1002/mrm.25708

16. Buerger C, Schaeffter T, King AP. Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation. *Med Image Anal* 2011; **15**: 551–64. doi: https://doi.org/10.1016/j.media.2011.02.009

17. Zhou T, Icke I, Dogdas B, Parimal S, Sampath S, Forbes J. Automatic segmentation of left ventricle in cardiac cine MRI images based on deep learning. In: Styner M. A, Angelini E. D, eds. *Proceedings of SPIE.* **vol. 101331W**. Orlando, FL, USA: International Society for Optics and Photonics; 2017.

18. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 2016; **30**: 108–19. doi: https://doi.org/10.1016/j.media.2016.01.005

19. Zhang W, Li R, Deng H, Wang L, Lin W, Ji S, et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *Neuroimage* 2015; **108**: 214–24. doi: https://doi.org/10.1016/j.neuroimage.2014.12.061

20. Guo Y, Wu G, Commander LA, Szary S, Jewells V, Lin W, et al. Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features. *Med Image Comput Comput Assist Interv* 2014; **17**(Pt 2): 308-15.

21. Hu S, Wei L, Gao Y, Guo Y, Wu G, Shen D. Learning-based deformable image registration for infant MR images in the first year of life. *Med Phys* 2017; **44**: 158: 158: 70. doi: https://doi.org/10.1002/mp.12007

22. Cheng R, Roth HR, Lu L, Wang S, Turkbey B. Active appearance model and deep learning for more accurate prostate segmentation on MRI. In: Styner M. A, Angelini E. D, eds. *Proceedings of SPIE.* **vol. 9784**. Orlando, FL, USA: International Society for Optics and Photonics; 2016.

23. Li F, Tran L, Thung KH, Ji S, Shen D, Li J. A robust deep model for improved classification of AD/MCI patients. *IEEE J Biomed Health Inform* 2015; **19**: 1610–6. doi: https://doi.org/10.1109/JBHI.2015.2429556

24. Amit G, Ben-Ari R, Hadad O, Monovich E, Granot N, Hashoul S. Classification of breast MRI lesions using small-size training sets: comparison of deep learning approaches. In: Armato S. G, Petrick N. A, eds. *Proceedings of SPIE.* **vol. 10134**. Orlando, FL, USA: International Society for Optics and Photonics; 2017.

25. Mehrtash A, Sedghi A, Ghafoorian M, Taghipour M, Tempany CM, Wells WM, et al. Classification of clinical significance of MRI prostate findings using 3D

convolutional neural networks. *Proc SPIE Int Soc Opt Eng* 2017; **10134**. doi: https://doi.org/10.1117/12.2277123

26. Wu G, Kim M, Wang Q, Munsell BC, Shen D. Scalable high-performance image registration framework by unsupervised deep feature representations learning. *IEEE Trans Biomed Eng* 2016; **63**: 1505–16. doi: https://doi.org/10.1109/TBME.2015.2496253

27. de Vos BD, Berendsen FF, Viergever MA, Staring M, Išgum I. End-to-end unsupervised deformable image registration with a convolutional neural network. In: *Lecture notes in computer science*. **10553**. QC, Canada: Springer; 2017. pp. 170406065.

28. Feng L, Grimm R, Block KT, Chandarana H, Kim S, Xu J, et al. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn Reson Med* 2014; **72**: 707–17. doi: https://doi.org/10.1002/mrm.24980

29. Beatty PJ, Nishimura DG, Pauly JM. Rapid gridding reconstruction with a minimal oversampling ratio. *IEEE Trans Med Imaging* 2005; **24**: 799–808. doi: https://doi.org/10.1109/TMI.2005.848376

30. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* 1999; **18**: 712–21. doi: https://doi.org/10.1109/42.796284

31. Kingma D, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:14126980. 2014.

32. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:160304467. 2016.

33. Kaufman L, Kramer DM, Crooks LE, Ortendahl DA. Measuring signal-to-noise ratios in MR imaging. *Radiology* 1989; **173**: 265–7. doi: https://doi.org/10.1148/radiology.173.1.2781018

34. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. *J Magn Reson Imaging* 2007; **26**: 375–85. doi: https://doi.org/10.1002/jmri.20969

35. Schwemmer C, Forman C, Wetzl J, Maier A, Hornegger J. CoroEval: a multi-platform, multi-modality tool for the evaluation of 3D coronary vessel reconstructions. *Phys Med Biol* 2014; **59**: 5163–74. doi: https://doi.org/10.1088/0031-9155/59/17/5163

36. Buerger C, Prieto C, Schaeffter T. Highly efficient 3D motion-compensated abdomen MRI from undersampled golden-RPE acquisitions. *MAGMA* 2013; **26**: 419–29. doi: https://doi.org/10.1007/s10334-013-0370-y

37. Kervrann C, Boulanger J. Optimal spatial adaptation for patch-based image denoising. *IEEE Trans Image Process* 2006; **15**: 2866–78. doi: https://doi.org/10.1109/TIP.2006.877529

38. Wang Y, Qiu Y, Thai T, Moore K, Liu H, Zheng B. A two-step convolutional neural network based computer-aided detection scheme for automatically segmenting adipose tissue volume depicting on CT images. *Comput Methods Programs Biomed* 2017; **144**: 97–104. doi: https://doi.org/10.1016/j.cmpb.2017.03.017

39. Noorda YH, Bartels LW, Viergever MA, Pluim JP. Subject-specific four-dimensional liver motion modeling based on registration of dynamic MRI. *J Med Imaging* 2016; **3**: 015002–02. doi: https://doi.org/10.1117/1.JMI.3.1.015002

40. Mardani M, Gong E, Cheng JY, Vasanawala S, Zaharchuk G, Alley M. Deep generative adversarial networks for compressed sensing automates MRI. arXiv preprint arXiv:170600051. 2017.

41. Dong C, Loy CC, Loy CC, Loy CC, He K, Tang X. *Learning a deep convolutional network for image super-resolution. European conference on computer vision*. **vol. 8692**: Springer; 2014. link.springer.com/chapter/10.1007/978-3-319-10593-2_13.

42. Bahrami K, Shi F, Rekik I, Shen D. Convolutional neural network for reconstruction of 7T-like images from 3T MRI using appearance and anatomical features. In: *International workshop on large-scale annotation of biomedical data and expert label synthesis*: Springer; 2016. pp. 39–47.