Received: 8 July 2021 | Revised: 28 October 2021 | Accepted: 2 November 2021

DOI: 10.1002/mp.15354

RESEARCH ARTICLE

MEDICAL PHYSICS

A deep learning method for eliminating head motion artifacts in computed tomography

Bin Su[1,¶] | **Yuting Wen**[2,¶] | **Yanyan Liu**[1,¶] | **Shu Liao**[3] | **Jianwei Fu**[1] | **Guotao Quan**[1] | **Zhenlin Li**[2]

[1] Shanghai United Imaging Healthcare Co., Ltd, Shanghai, China

[2] West China Hospital, Sichuan University, Chengdu, China

[3] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China

Correspondence
Guotao Quan, Shanghai United Imaging Healthcare Co., Ltd., Shanghai 201807, China.
Email: guotao.quan@united-imaging.com
Zhenlin Li, Department of Radiology, West China Hospital, Sichuan University, Chengdu 610041, China.
Email: HX_lizhenlin@126.com

¶B. Su, Y. Wen, and Y. Liu contributed equally to this work

Funding information
Research of Digital Theranostic Equipment, Grant/Award Number: 2016ZY04002047; National Science and Technology Planning Project of China, Grant/Award Number: 2016ZY04002047

Abstract

Purpose: Involuntary patient movement results in data discontinuities during computed tomography (CT) scans which lead to a serious degradation in the image quality. In this paper, we specifically address artifacts induced by patient motion during a head scan.

Method: Instead of trying to solve an inverse problem, we developed a motion simulation algorithm to synthesize images with motion-induced artifacts. The artifacts induced by rotation, translation, oscillation and any possible combination are considered. Taking advantage of the powerful learning ability of neural networks, we designed a novel 3D network structure with both a large reception field and a high image resolution to map the artifact-free images from artifact-contaminated images. Quantitative results of the proposed method were evaluated against the results of U-Net and proposed networks without dilation structure. Thirty sets of motion contaminated images from two hospitals were selected to do a clinical evaluation.

Result: Facilitating the training dataset with artifacts induced by variable motion patterns and the neural network, the artifact can be removed with good performance. Validation dataset with simulated random motion pattern showed outperformed image correction, and quantitative results showed the proposed network had the lowest normalized root-mean-square error, highest peak signal-to-noise ratio and structure similarity, indicating our network gave the best approximation of gold standard. Clinical image processing results further confirmed the effectiveness of our method.

Conclusion: We proposed a novel deep learning-based algorithm to eliminate motion artifacts. The convolutional neural networks trained with synthesized image pairs achieved promising results in artifacts reduction. The corrected images increased the diagnostic confidence compared with artifacts contaminated images. We believe that the correction method can restore the ability to successfully diagnose and avoid repeated CT scans in certain clinical circumstances.

KEYWORDS
CT, image reconstruction, machine learning

1 | INTRODUCTION

Patients are always coached to remain still during CT scans. However, motion can be difficult to avoid in certain clinical circumstances, such as nerve injury or cerebral hemorrhage. Furthermore, the problem can be exacerbated by unintended motions of the patient support such as a vibrating patient support.[1,2] Unexpected motion during CT scans causes data discontinuities which in turn introduce streaks or star artifacts

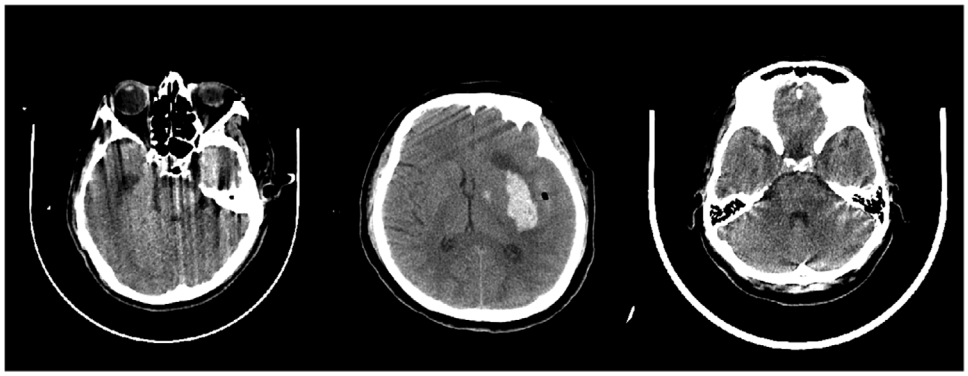

FIGURE 1 Clinical images with motion artifacts

originating from anatomical structures with a high gradient. As shown in Figure 1, motion-induced artifacts in head images mainly originate from the skull. These artifacts degrade the quality of reconstructed images and may also confuse diagnoses.[3,4]

Many ideas have been proposed for the correction of motion artifacts. Shortening scan time can freeze motion[5]; however, this method reduces the signal-to-noise ratio (SNR) compared with routine scanning. View weighting[6] is a commonly used technique for assigning lower weighting factors to views with discontinuities. This method can correct artifacts to a certain extent. However, as the motion pattern is often variable, a general set of weighting factors is not available. A software approach uses motion artifacts metrics (MAM) such as total variation for motion elimination.[7] In this approach, by using the estimated motion, the images are updated via motion correction reconstruction. The motion is then estimated by using the updated reconstructed images. The processes are repeated alternatively until the MAM value of the image is minimized. One last method worth mentioning is a projection-based motion estimation and compensation scheme.[8] This approach is based on image reconstruction using the ordered subset expectation maximization algorithm. The iterative methods have high computational complexity, and their penalty function is hard to design to achieve good performance.

Convolutional neural networks (CNNs) have yielded substantial benefits in medical image processing.[9–13] CNNs can efficiently learn high-level features from regions of pixels through a hierarchical framework. There are already numerous papers on CNNs-based image analysis, such as lesion classification, image enhancement and organ segmentation. CNN-based approaches have also been applied to image reconstruction for low-dose CT including low tube-current CT imaging and sparse-view CT imaging.

Here, we present a novel approach to correct motion artifacts. Instead of solving an inverse problem, we began with motion-free images, simulated motion artifacts, then used these image pairs to train a modified multi-scale neural network. Since the motion-induced artifacts mainly originate from the skull (an anatomical structure with a high gradient) and extend far beyond their origin, the multi-scale network took into account both a large reception field and the high image resolution. The overall process took into account several types of motion-induced artifact and took full advantage of the powerful learning capability of neural networks.

In the following sections, we introduce the simulation method and the modified neural network. Validation was performed on each component and on the overall method. After that, clinical images are shown demonstrating the feasibility of the proposed method. We end the paper with a discussion of the current results and some future related work.

2 | MATERIALS AND METHODS

2.1 | Motion simulation

The motion simulation algorithm processes data geometrically similar to the acquisition process during a CT scan. The 3D motion pattern of the object is predefined before the virtual data acquisition. For each view, the object moves with a small angle or distance or a combination of both. Then, 3D forward projection is performed on the object to get projection data.[14] A full circle of rotation is simulated and an appropriate reconstruction algorithm is used to reconstruct images with simulated motion artifacts.

A stack of motion free, volumetric head image is used as the scanned objects. As shown in Figure 2, 3D forward projection is performed on the volumetric images view by view. Rigid motions, including rotations, translations and oscillations, are applied on the object changing view by view. After acquiring the projection data in a full rotation, we back-projected the data to get the associated motion simulated images through a filtered back-projection (FBP) operation.[14]

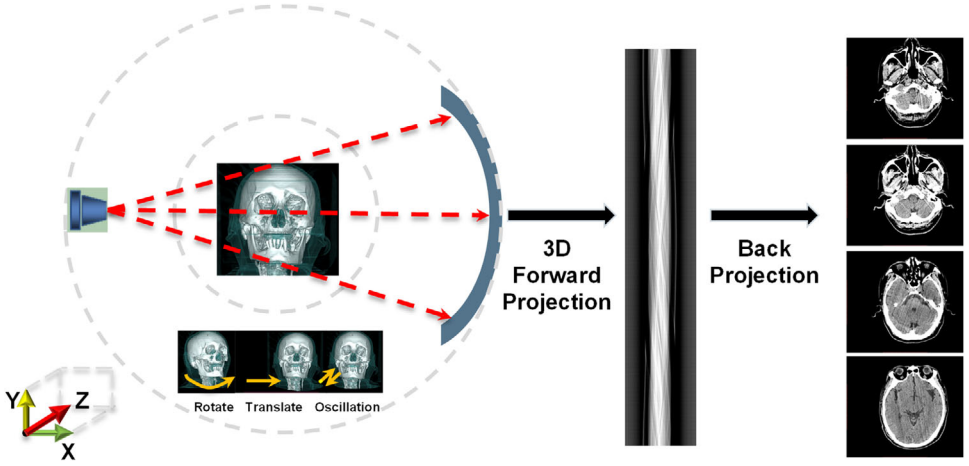

FIGURE 2 The schematics of motion simulation

The image volume after 3D motion simulation $I_v(x, y, z)$ can be represented by the motion-free image volume $I_0(t, s, l)$ in the coordinate transformation system, where the two coordinate systems are related by the following sets of equations:

$$\begin{cases} x = t \cdot \cos(\Delta\theta y(v)) \cdot \cos(\Delta\theta z(v)) + s \cdot \sin(\Delta\theta y(v)) \\ \quad \cdot \cos(\Delta\theta z(v)) + l \cdot \sin(\Delta\theta z(v)) \\ y = -t \cdot \sin(\Delta\theta y(v)) + s \cdot \cos(\Delta\theta y(v)) \\ z = -t \cdot \cos(\Delta\theta y(v)) \cdot \sin(\Delta\theta z(v)) - s \cdot \sin(\Delta\theta y(v)) \\ \quad \cdot \sin(\Delta\theta z(v)) + l \cdot \cos(\Delta\theta z(v)) \end{cases} \tag{1}$$

$$\begin{cases} x = t + \Delta x(v) \\ y = s + \Delta y(v) \\ z = l + \Delta z(v) \end{cases} \tag{2}$$

where $v \in [1, 2, \ldots, N_V]$ is the view index, N_v is the number of views. Equation (1) represents a coordinate transformation of rotation. $\Delta\theta y(v)$ and $\Delta\theta z(v)$ are the rotation angles formed with the x-axis in XY plane and XZ plane respectively. Equation (2) represents a coordinate transformation of translation and oscillation. $\Delta x(v)$, $\Delta y(v)$ and $\Delta z(v)$ are the translation distances in the x, y and z directions, respectively.

The projection data p_v under view angle θ_v, $v \in \{1, 2, \ldots, N_V\}$ can be modeled by Equation (3), where P is the 3D forward projection operator

$$p_v = P(I_v, \theta_v) \tag{3}$$

Denoting a full circle of simulated projections as $p = \{p_v, \text{ for all } v\}$, the simulated 3D motion image volume is calculated as:

$$I_m = R(p) \tag{4}$$

TABLE 1 The parameters of motion simulation

Motion type	Amplitude	Frequency	Dimension
Rotation	$0 \sim 15^o$	–	3D
Translation	$0 \sim 5$ mm	–	3D
Oscillation	$0 \sim 5$ mm	<3 Hz	1D

where R denotes the 3D filtered back-projection operator.

The motion simulation algorithm tries to simulate head motion similar to that seen clinically. The rotation and translation simulation algorithms simulate head rigid motion during a CT scan. Such motion can start at any view angle and extend to a certain number of views. Therefore, some linear motion during certain views is defined. The start view and end view are randomly chosen. The rotation angle is restricted to be less than 15^o in any direction to most realistically match motions seen in clinical situations. Similarly, the translation starts at a random view and extends for a random number of views. The translation distances are restricted to be no more than 5 mm. The oscillation simulation algorithm simulates the couch oscillations during movement, which may be caused by the mechanical resonance or unstable interface between head holder and couch. The oscillation is simulated with an amplitude of less than 5 mm and a frequency of less than 3 Hz. Considering the mechanical structure of the couch, oscillations are more likely to move up and down, so we just simulated oscillations in that direction.

The typical parameters used in the simulation algorithm are shown in Table 1.

2.2 | Motion correction

Given the paired motion and motion-free image volume, a 3D neural network can be constructed to correct the

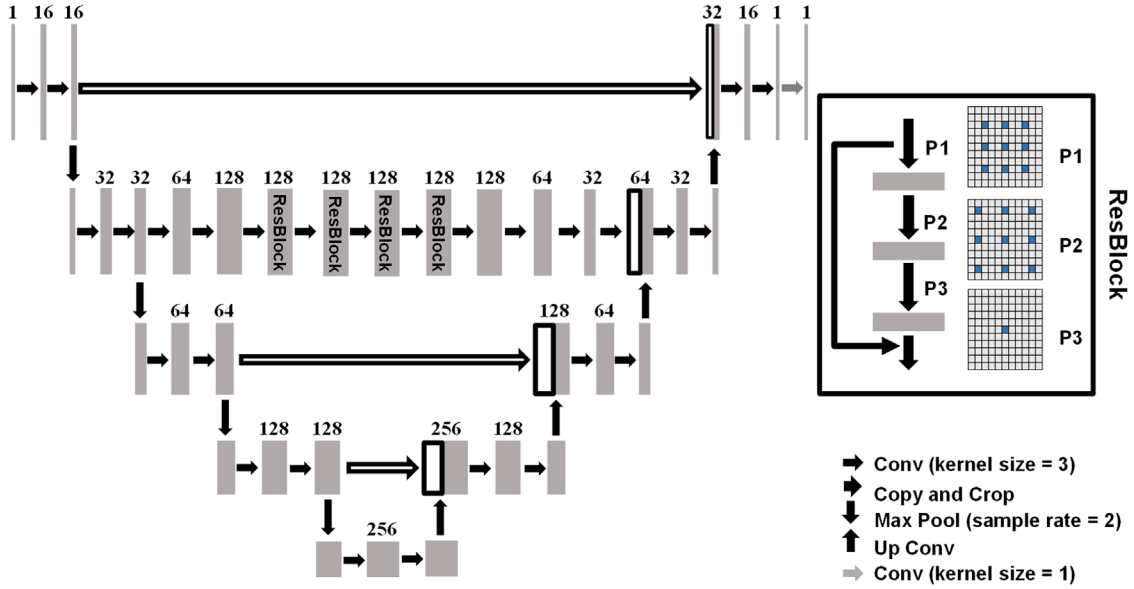

FIGURE 3 The structure of the proposed network

motion-induced artifacts. Considering that the motion-induced artifacts mainly originate from the skull and extend far beyond their origin, a neural network with a large reception field is required. Also, the spatial resolution should not be degraded during the processing by the network.

2.2.1 | Neural network structure

The network for correcting motion-induced artifacts is a modification of the U-Net.[15] We introduced a ResNet[16] in the middle level of the U-Net. The multiple pooling layers in the U-Net guarantee a large reception field, which contributes to extracting widely extended artifacts. However, the cascaded pooling operation could lose spatial resolution which cannot be fully recovered even with deconvolution operations. Therefore, an additional ResNet branch is applied in the middle level of the U-Net to retain the resolution. The original ResNet is a cascaded convolution operation with some shortcut connections to ease the gradient back-propagation. The absence of a down-sampling operation retains the resolution while sacrificing the large reception field. In order to balance the resolution and reception field, we applied a dilation structure in the ResNet branch.

The detailed structure of the proposed neural network is illustrated in Figure 3. Each gray box represents a learned feature map, and the white box represents the copied feature map. The channel number of features is at the top of the box. Arrows with different orientations and styles represent different operations and are defined in the legend at the bottom right of the figure. The neural network consists of a contraction path (left) and an expansion path (right). The contraction path executes two repetitive unpadded convolutions with a kernel size of 3, each followed by a batch normalization (BN), an activation function of leaky rectified linear unit (LeakyReLU) and a maximum pooling structure with a sample rate of 2 for down-sampling. The number of features is doubled after each down-sampling step. In the expansion path, an up-sampling operation is performed to double feature size which follows a convolution with a kernel size of 1 to reduce the feature number by half. After that, the features are concatenated with the corresponding copied features and two repetitive unpadded convolutions with a kernel size of 3. Each use BN and LeakyReLU. It is necessary to crop feature maps because of lost boundary pixels in each convolution. In the last layer, a convolution with a kernel size of 1 is used to get the final output image.

The second concatenation connection is replaced by a ResNet branch. The ResNet branch consists of five convolution layers and four cascaded ResBlocks. The number of features expands before each ResBlock, and contracts respectively after the ResBlock. Shortcut connections are attached between the input and output of sequential convolutional layers in ResBlock. Adding these shortcut connections solves the gradient vanishing problem, helping the network to learn more quickly.[15] Each ResBlock contains three sequenced convolution layers with the same feature number and dilation rates of 2, 3 and 5.[17] The dilation rate configuration facilitates the large reception field and avoids a gridding effect. BN is performed right after each convolution and before the LeakyReLU activation function.

To improve the network learning process, it is important to reduce the solution space by compressing the mapping range.[18] We observed that the residual artifact images between the motion and motion-free images

have a significant range reduction in CT values. As a result, the residual images are introduced to be the gold standard. The network is trained to map the artifact images from the motion-induced images, and finally, the motion-induced images are subtracted from the learned results to get motion-corrected images.

2.2.2 | Network training

Paired motion and motion-free images are the training dataset. Fifty motion-free volumetric images were acquired from the United Imaging Healthcare (UIH) clinical database. All images were reconstructed with a 512×512 image size and 230 mm reconstruction diameter. All the images were identified to be artifact free and satisfy clinical diagnostic requirements.

By the aforementioned motion simulation algorithm, five sets of motion image volumes were simulated for each set of the motion-free image volume. Each image volume contained about 60 images, so we had about $50 \times 60 \times 5 = 15\ 000$ paired images in total. We randomly selected 12 000 image pairs as the training dataset, and the remaining 3000 image pairs were used to validate the trained network.

We applied mini-batch adaptive moment estimation (Adam)[19] to minimize the loss function (Equation (5)).

$$\text{Loss} = (L(In, w) - Tar)^2 + \mu * |L(In, w) - Tar| \quad (5)$$

In and *Tar* represent the input volume of the multi-scale network and the gold standard. $L(:)$ represents the processing of multi-scale networks, where w is the learnable parameters of the network. The objective function in Equation (5) was minimized with respect to the network parameters w, which correspond to the weights of the convolution filters. For 3D network, the calculation is done in each image volume. The first term on the right side of the equation is L_2 loss function, and the last term on the right side of the equation is L_1 loss function. μ controls the weighting of L_1 loss and L_2 loss, which is set it to be 1 mandatorily. Each loss function has distinct benefits: L_2 loss speeds up the learning in early epochs, while L_1 loss generates better results in later epochs. By using mini-batch Adam, the parameters got more frequent updates and the training efficiency was greatly improved. Here, we set the mini-batch size to be 16 and iteratively trained the model for 1000 epochs to get the optimal parameters.

3 | RESULTS

3.1 | The simulated motion artifacts

The simulated motion artifacts with rotation, translation, oscillation and a combination of the three patterns are shown in Figure 4. Streak artifacts extended into brain tissue obscuring the brain structure. With a certain deflection angle, the streak artifacts showed up with a broadened width correlated with the deflection angle as shown in the rotation simulation. With a movement of fixed direction, the streak artifacts appear with different angles correlated with the skull structure as shown in the translation simulation. For movement of cycling oscillation, the artifacts show up with a star-like shape originating from the skull structure with a spike as shown in oscillation simulation. Similar artifacts originated in the inner ear. Using different parameters set during simulation, the streak artifacts appear with different widths and strengths. The mixture of different motion patterns gives more complicated artifacts having both streak and star-like artifacts.

3.2 | Network validation

3000 paired images were used for network validation. The validation dataset contains images with translation, rotation, oscillation and combinations of the aforementioned motions. In order to prove the effectiveness of the proposed network structure, we compared learning results with that of the original U-Net and a similar multi-scale network but without dilation. Randomly chosen three sets of images are shown in Figure 5 for comparison. The artifacts include streaks with different widths and different orientations. Also, some star-like artifacts can be observed. The proposed neural network finely corrected the motion artifacts and recovered the brain tissue. The original U-Net, with a large reception field, gained the ability to learn global artifact information, but the reduced spatial resolution resulted in barely satisfactory artifact correction. When the second concatenation connection in U-Net is replaced by ResNet, substantial improvements can be expected. However, there were still some artifacts remaining in some cases. The images corrected by our proposed network showed significant improvement, validating the effectiveness and practicality.

We calculated normalized root-mean-square error (NRMSE), peak signal-to-noise ratio (PSNR), and structure similarity (SSIM) for a quantitative comparison. NRMSE reflects the normalized error of two images, which can be calculated by Equation (6). Where *Tar* means gold standard, and CNN means the image processed by CNNs:

$$\text{MRMSE} = \frac{\sum_i \sum_j \left| Tar_{ij} - CNN_{ij} \right|^2}{\sum_i \sum_j \left| Tar_{ij} \right|^2} \quad (6)$$

PSNR is often used to measure the quality of reconstructed digital signals,[20] which can be calculated by

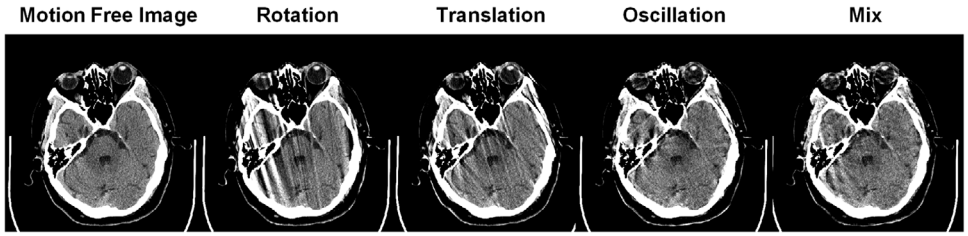

FIGURE 4 The simulated artifacts in different motion patterns

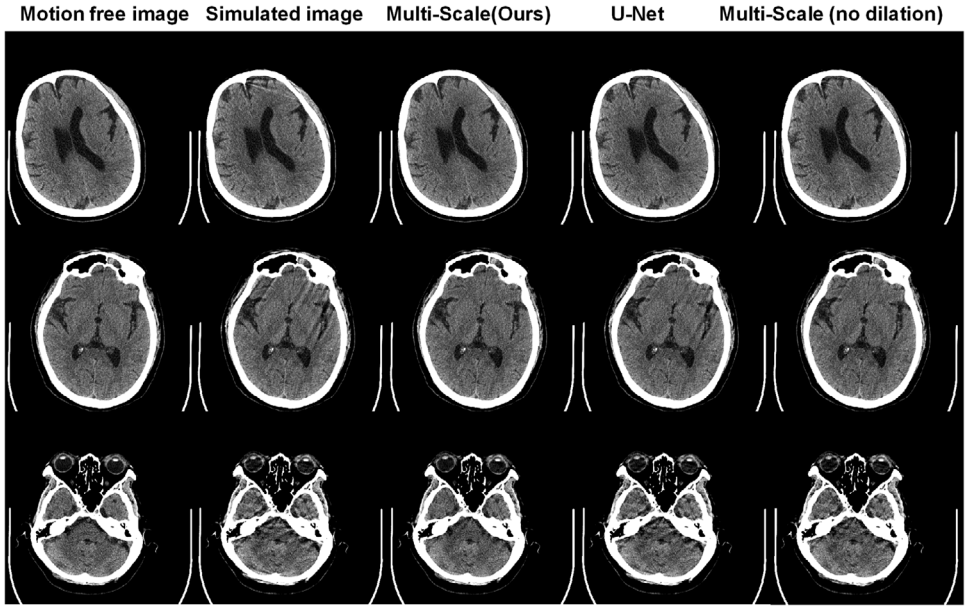

FIGURE 5 Comparison of motion correction results between different networks

TABLE 2 The quantitative comparison between different networks

	Target	Input	Multi-scale (Ours)	UNet	Multi-scale (no dilation)
NRMSE	0	0.060 ± 0.005	0.020 ± 0.005	0.050 ± 0.003	0.035 ± 0.007
PSNR	–	65.23 ± 0.008	68.96 ± 0.010	65.51 ± 0.080	67.53 ± 0.010
SSIM	1	0.940 ± 0.010	0.982 ± 0.008	0.952 ± 0.010	0.964 ± 0.007

Equation (7). $2^N - 1$ is set to be 8191, which denotes the maximum CT value:

$$PSNR = 10\log_{10}\left(\frac{\left(2^N-1\right)^2}{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|Tar_{ij}-\mathrm{CNN}_{ij}\right|^2}\right) \tag{7}$$

SSIM is a metric to measure the similarity of two images, including brightness, contrast and structure. Quantitative results are summarized in Table 2. Our network has the lowest NRMSE, highest PSNR and SSIM, indicating our network gives the best approximation of the gold standard.

3.3 | Artifact correction on clinical images

Thirty sets of images were acquired from the UIH clinical dataset to do the clinical evaluation. Three randomly chosen images are shown in Figure 6. The top row lists the original images with artifacts. It can be observed that artifacts introduced by simulation closely resemble the clinically encountered artifacts. In both cases, there is basically a composition of streak and star-shaped

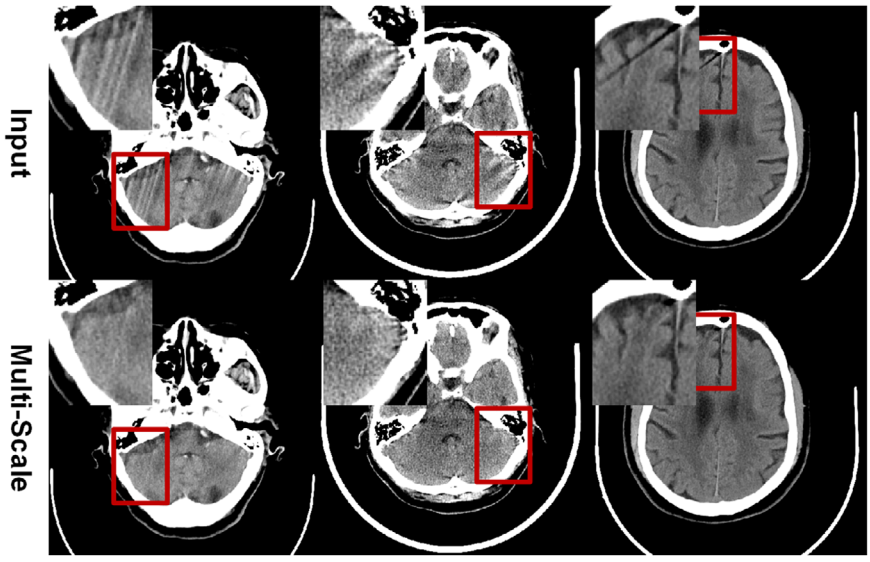

FIGURE 6 Motion correction on clinical images. Top row: motion images, bottom row: corrected images

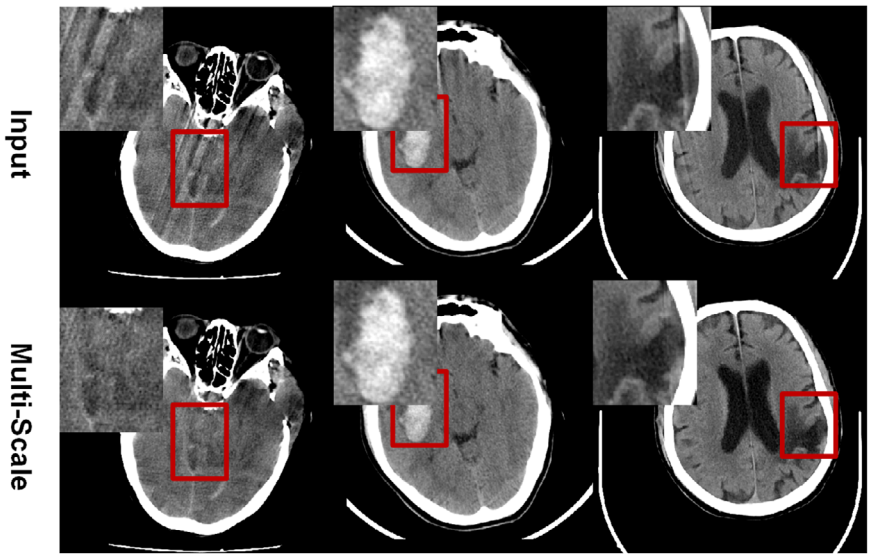

FIGURE 7 Motion correction on lesion images. Top row: motion images, bottom row: corrected images

artifacts. The bottom row lists the corrected images. The proposed method shows good performance on artifacts with different shapes and strengths. It should be pointed out that since the dose level and image thickness varied, the noise level for the three images in Figure 6 was different. However, in all cases, our proposed network eliminated the artifact, without interrupting the noise pattern.

As shown in Figure 7, we also chose another three images to show the applicability of disease presentation. These clinical cases had streak artifacts that go through the lesion region, interrupting the structure of disease and thus affecting diagnosis. The proposed method can eliminate artifacts without obscuring the lesion. The corrected images increase the diagnostic confidence compared with artifact-contaminated images.

In order to show the artifact reduction in a whole image volume, we provided a typical multiple planar reformation (MPR) image to evaluate the effectiveness. The image thickness for the MPR is 2 mm. Some ghosting boundary blurring artifacts can be observed on the bottom of the MPR. Since the head scan was done with a gantry tilt, the artifact-contaminated region has an inclination. The proposed method can eliminate the artifacts with great performance and recover the contaminated structures, as shown in Figure 8.

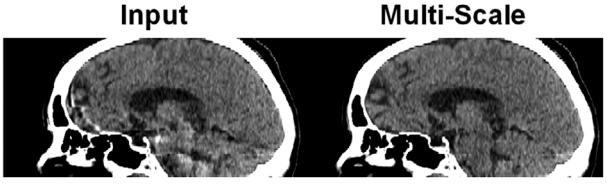

FIGURE 8 Motion correction on clinical multiple planar reformation (MPR) images

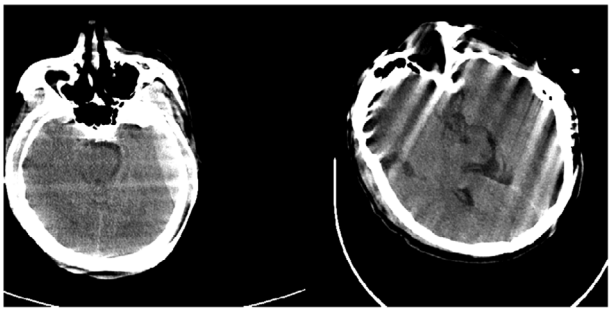

FIGURE 9 The images with doubled skull (left) and deformed skull (right)

4 | DISCUSSION

We presented a motion correction algorithm for reducing motion-induced artifacts in head CT scans. Head rigid motion, including rotation, translation, oscillation and any combination, were simulated during a full-circle data acquisition. With 3D forward projection and back projection, we synthesized images with streaks and some star-like artifacts originating from the skull. A novel 3D neural network structure, which combined the advantages of ResNet and U-Net, was proposed to map the artifact-free images from simulated images with motion artifacts. The large reception field of U-Net facilitated the learning of long-range extended artifacts, and the high resolution of ResNet avoided the loss of spatial resolution. The proposed multi-scale network outperformed results compared to other commonly used networks. Clinical results proved the validity of the proposed method. We believe that the correction method can improve the ability to successfully diagnose and avoid repeated CT scans in some clinical circumstances.

It is hard for analytical methods to solve the motion-induced artifacts. Solving equations to get artifact-free images from artifact modulated images is ill-posed. Iterative methods may approach the clean images with some penalty; however, the computation cost is huge and the penalty design calls for a rich experience. Also, it is difficult to find a penalty function that can meet different motion patterns. The deep learning method may open a new door to solve such problems. Instead of solving the inverse problem, we just do the forward calculation, and let the neural network do the inverse mapping. Taking advantage of the generalization of neural networks and diversity of simulated motion artifacts, the deep learning method can solve the problems with minimal effort.

In some extreme conditions, the motion can be so severe that the FBP images showed a ghosted doubled skull or a deformed skull, as shown in Figure 9. At this time, the proposed network cannot recover such motion-induced skull deformities. Networks that can perform geometrical deformation may be needed to solve the problem.

5 | CONCLUSION

We proposed a network-based method to eliminate motion artifacts. The novel neural network worked well at artifact reduction. We believe that neural networks can be employed in more CT image processing tasks. In the future, we will also investigate applying neural networks directly to CT raw data processing.

ACKNOWLEDGMENTS

This work was supported by the state key project of "Research of Digital Theranostic Equipment" under grant 2016ZY04002047. We gratefully thank Patrick Kling for revising the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Research data is not shared due to ethical restrictions.

REFERENCES

1. Hernandez JD, Eldib ME, Hegazy M, Myung HC, Cho MH, Lee SY. A head-motion estimation algorithm for motion artifact correction in dental CT imaging. *Phys Med Biol*. 2018;63(6):065014.
2. Wagner A, Schicho K, Kainberger F, Birkfellner W, Grampp S, Ewers R. Quantification and clinical relevance of head motion during computed tomography. *Invest Radiol*. 2003;38(11):733.
3. Bontempi M, Bettuzzi M, Casali F, Pasini A, Rossi A, Ariu M. Relevance of head motion in dental cone-beam CT scanner images depending on patient positioning. *Int J Comput Assisted Radiol Surg*. 2008;3(3-4):249-255.
4. Boas FE, Fleischmann D. CT artifacts: causes and reduction techniques. *Imaging Med*. 2012;4(2):229-240.
5. Fleischmann D, Boas FE. Computed tomography—old ideas and new technology. *Eur Radiol*. 2011;21(3):510-517.
6. Pelc NJ, Glover GH. Method for reducing image artifacts due to projection measurement inconsistencies. 1986. https://www.freepatentsonline.com/4580219.html U.S, Patent 4.
7. Bruder H, Rohkohl C, Stierstorfer K, Flohr T. Compensation of skull motion and breathing motion in CT using data-based and image-based metrics, respectively. 2016:97831E. SPIE Medical Imaging. International Society for Optics and Photonics.
8. Sun T, Kim J, Fulton R, Nuyts J. An iterative projection-based motion estimation and compensation scheme for head X-ray CT. *Med Phys*. 2016;43(10).
9. Wang G. A perspective on deep imaging. *IEEE Access*. 2017;4:8914-8924.
10. Wang G, Ye JC, Man BD. Deep learning for tomographic image reconstruction. *Nat Mach Intell*. 2020;2(12):737-748.
11. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24-29.
12. Yang Q, Yan P, Wang G, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging*. 2018:1348-1357.
13. Shan H, Jia X, Yan P, Li Y, Paganetti H, Wang G. Synergizing medical imaging and radiotherapy with deep learning. *Mach Learn Sci Technol*. 2020;1(2):021001.
14. Turbell H. *Cone-beam reconstruction using filtered backprojection*. Sweden: Linkoping Studies in Science and Technology, 2001, Thesis No. 672. doi:10.1.1.134.5224

15. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-assisted Intervention, 2015. Springer; 234-241. https://doi.org/10.1007/978-3-662-54345-0_3
16. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. IEEE Conference on Computer and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016. IEEE; 2016. https://doi.org/10.1109/CVPR.2016.90
17. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *ICLR*, 2016. arXiv: 1511.07122v3
18. Fu X, Huang J, Zeng D, Huang Y, Ding X, Paisley J, Removing rain from single images via a deep detail network. IEEE Conference on Computer Vision & Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. https://doi.org/10.1109/CVPR.2017.186
19. Kingma D, Adam BaJ, : A method for stochastic optimization. International Conference on Learning Representations, 2015: 1-13. arXiv:1412.6980
20. Wang J, Zeng L, Guo Y, et al. ADMM-based deep reconstruction for limited-angle CT. *Phys Med Biol*. 2019;64(11):115011.

How to cite this article: Su B, Wen Y, Liu Y, et al. A deep learning method for eliminating head motion artifacts in computed tomography. *Med Phys.* 2022;49:411–419. https://doi.org/10.1002/mp.15354