Received: 3 July 2019 | Revised: 24 October 2019 | Accepted: 24 October 2019

DOI: 10.1002/nbm.4225

SPECIAL ISSUE RESEARCH ARTICLE

NMR IN BIOMEDICINE WILEY

Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation

Kamlesh Pawar[1,2] | Zhaolin Chen[1] | N. Jon Shah[1,3] | Gary F. Egan[1,2]

[1]Monash Biomedical Imaging, Monash University, Melbourne, Australia

[2]School of Psychological Sciences, Monash University, Melbourne, Australia

[3]Research Centre Jülich, Institute of Medicine, Jülich, Germany

Correspondence
Kamlesh Pawar, Monash Biomedical Imaging, Monash University, 770 Blackburn Road, Notting Hill, VIC-3168. Australia.
Email: kamlesh.pawar@monash.edu

The suppression of motion artefacts from MR images is a challenging task. The purpose of this paper was to develop a standalone novel technique to suppress motion artefacts in MR images using a data-driven deep learning approach. A simulation framework was developed to generate motion-corrupted images from motion-free images using randomly generated motion profiles. An Inception-ResNet deep learning network architecture was used as the encoder and was augmented with a stack of convolution and upsampling layers to form an encoder-decoder network. The network was trained on simulated motion-corrupted images to identify and suppress those artefacts attributable to motion. The network was validated on unseen simulated datasets and real-world experimental motion-corrupted in vivo brain datasets. The trained network was able to suppress the motion artefacts in the reconstructed images, and the mean structural similarity (SSIM) increased from 0.9058 to 0.9338. The network was also able to suppress the motion artefacts from the real-world experimental dataset, and the mean SSIM increased from 0.8671 to 0.9145. The motion correction of the experimental datasets demonstrated the effectiveness of the motion simulation generation process. The proposed method successfully removed motion artefacts and outperformed an iterative entropy minimization method in terms of the SSIM index and normalized root mean squared error, which were 5–10% better for the proposed method. In conclusion, a novel, data-driven motion correction technique has been developed that can suppress motion artefacts from motion-corrupted MR images. The proposed technique is a standalone, post-processing method that does not interfere with data acquisition or reconstruction parameters, thus making it suitable for routine clinical practice.

KEYWORDS
deep learning, MR image reconstruction, MR motion correction

1 | INTRODUCTION

MRI[1] is an established medical imaging modality used to produce detailed images of different tissue types. However, despite its advantages, image acquisition times can be long, and any patient motion during scanning causes artefacts in the resultant MR images.[2] Since the advent of MRI, the removal of motion artefacts has continued to be an unresolved problem for most clinical applications. In particular, motion artefacts present

Abbreviations: CNN, convolutional neural network; EPIK, echo planar imaging with keyhole; MoCoNet, motion correction network; PROPELLER, periodically rotated overlapping parallel line with enhanced reconstruction; SNR, signal to noise ratio; SSIM, structural similarity..

significant difficulties when scanning paediatric, elderly, claustrophobic and stroke patients. Apart from patient discomfort, motion artefacts also result in financial losses[3] for imaging sites. A number of methods can be used to correct for motion artefacts, the most commonly used ones being rigid motion correction techniques. These techniques can be broadly classified into three categories: navigator sequence-based approaches, external motion detection devices and data-driven approaches.

The navigator-based approaches involve inserting *k*-space, or image navigators, into the imaging sequences. The navigators are inserted within the sequence at different time points and are compared with a reference to estimate motion parameters. The parameters can subsequently be used to perform either prospective or retrospective motion correction. The first navigator sequence to estimate in-plane linear motion was proposed by Ehman and Felmlee[4] and involves acquiring the zeroth phase encode in both the *x* and *y* directions. The phase difference between navigators at different time points provides a measure of motion in the *x* and *y* directions. However, this technique only compensates for translation motion. An alternative technique, known as PROPELLER (periodically rotated overlapping parallel line with enhanced reconstruction),[5] is a self-navigation technique in which a number of straight lines called blades are acquired at different angles to cover the whole *k*-space. The acquisition of blades at various angles results in the acquisition of the centre of *k*-space multiple times, and the points that are sampled multiple times are used as navigators to extract rotation and translation information. A limitation of the PROPELLER technique is that it can only provide in-plane motion parameters. This limitation is, however, not specific to PROPELLER, and is a limitation found in all 2D imaging sequences, as the through-plane motion disrupts signals from different slices.

Another technique that enables compensation for linear motion, based on the acquisition of data on a Cartesian grid, is EPIK (echo planar imaging with keyhole).[6] EPIK is an interleaved technique that acquires the central line of *k*-space in each shot and therefore enables motion correction along a single axis. Application of EPIK to diffusion imaging with motion correction has been demonstrated in Reference 7. Multiple implementations of the use of *k*-space navigators to estimate motion parameters in all six degrees of freedom have been developed,[8-13] along with image navigators[14-17] for 3D imaging sequences. Despite the advantages afforded by EPIK, its implementation requires all navigator sequences to be carefully designed, resulting in an inevitable increase in acquisition time and a reduction in signal to noise ratio (SNR). Constraints are also imposed on protocol parameters, including $T_R/T_E/T_I$. Recently, a novel approach for motion correction using a spherical navigator based on a Lissajous trajectory has been demonstrated in a proof of concept study. Here, an isotropic translational accuracy of 0.2 mm and a rotation estimate of 0.5 degrees was demonstrated. The Lissajous navigator has only modest gradient demands, covers the poles of the *k*-space sphere and acquires data in approximately 5 ms.[18]

External device-based approaches involve an estimation of the motion using an external device, such as an optical camera, that tracks a marker placed on the patient,[19-23] or using field probes that measure the magnetic field perturbation due to motion.[24,25] The major limitations of external device techniques are that they involve additional cost and require calibration, and the use of mounting markers on the patient may cause patient discomfort and potential patient compliance issues, making them problematic for clinical usage.

Data-driven approaches to motion correction estimate and correct motion directly from the images. Entropy focus criterion[26,27] techniques belong to the category of data-driven approaches that have been used for retrospective motion correction. These approaches correct motion artefacts by minimizing image entropy and do not require additional navigators. However, the entropy focus criteron method relies on computationally demanding iterative algorithms.

Recently, convolutional neural networks (CNNs) have been used in MRI for image reconstruction from undersampled data[28-34] and for MR image analysis.[35] The basic operation in CNNs is the convolution operation. Convolution operations are widely used in image processing applications, such as image segmentation, edge detection, image de-noising, Gaussian blurring and median filtering, to name a few. For instance, in order to remove noise from images, a simple averaging kernel for convolution can be used, which is equivalent to low pass filtering. In these applications, the convolution kernels are hand-engineered using prior knowledge of the imaging process, noise statistics and the nature of the artefacts. However, algorithms for such applications, including image de-noising and segmentation, are being rapidly replaced by deep learning[31] approaches.

The use of CNNs to detect motion[36] or to correct motion artefacts[37-43] has been explored in previous studies.[36-43] The concept of using a CNN for end-to-end motion correction was first presented in our early work, published as a conference abstract.[42] Similar work, including Reference 43, applied a CNN to individual patches of 2D spin echo images to correct for simulated motion artefacts. In Reference 37 a CNN was used as a pre-processing block to improve the motion correction physics model, and in Reference 38 it was used as a regularizer for iterative model-based motion correction. A 3D patch-based approach using a condition-generative adversarial network was proposed in Reference 40. A feasibility study of retrospective motion correction using CNNs for rigid and non-rigid motion correction in 2D imaging was presented in Reference 41. Most of the methods[37,38,41,43] applied the CNN approach to motion correction in 2D imaging; some methods[38,43] applied CNNs on small patches of 2D images; while other methods[37,38] used a CNN to improve the motion-based model. Motion simulation for 3D images was also performed in Reference 40 and the authors tested their network in both simulated and experimental datasets. However, the 3D images in Reference 40 were from a FLASH sequence at 7 T with an extremely narrow field of view in both the sagittal and coronal views, and the motion simulation was limited to only 2–3 motion events during the scan, which may not be a realistic estimate.

The ringing, blurring and ghosting artefacts caused by motion can be suppressed with convolution operations. However, a set of convolution kernels that can be used to remove motion artefacts may not be analytically determined due to the random nature of the motion that causes the

artefacts. The objective of this study was to simulate motion-corrupted images and subsequently use them to train a deep learning network that could learn appropriate convolution kernels to identify and correct for artefacts arising as a result of motion. We leverage the power of data-driven deep learning methods to determine sets of convolution kernels that can efficiently remove artefacts arising due to motion. We focus on the widely used 3D T_1 MPRAGE images, where motion artefacts are often observed at 3 T MRI. Extensive motion simulation covering the whole head in the three orthogonal views was performed. This paper describes the following.

1. The application of multi-resolution Inception-ResNet as an encoder that consists of different convolutional kernel sizes at each resolution scale to capture both local and global motion artefacts.
2. A method to design and simulate realistic motion curves to capture both abrupt motion and gradual motion of the real-world scenarios.
3. Training and validation of deep learning models with a large dataset containing 1572 subjects for training and 384 subjects for validation. The training and validation sets consisted of approximately 235 K and 57 K 2D slices respectively.

Here, we have substantially improved our previous method[42] and have used a multi-resolution Inception-ResNet[44] network architecture for the suppression of motion artefacts. We also demonstrate motion suppression for both simulation and experimental datasets, which were subsequently compared with an iterative entropy minimization technique. The proposed method, referred to as MoCoNet (motion correction network), is a data-driven approach based on a CNN that learns to identify and suppress motion artefacts from images retrospectively.

2 | METHODS

2.1 | Model of motion artefact in MRI

Motion during MR data acquisition manifests itself as artefacts in the reconstructed images. The nature of the artefacts found in structural scans depends on the magnitude of the motion and the time point at which it occurred. In this work, we have neglected the motion during read-out. Consider a movement (ie position change) that occurs at time point t during data acquisition (ie during the nth phase encoding step). The motion-corrupted k-space can be considered as a linear combination of two partial sets of k-spaces (Figure 1C and 1L) before and after motion occurrence. The corresponding motion-corrupted image (Figure 1M) can then be modelled by a linear combination of motion-free images (before and after change of position, Figure 1D, G) convolved with a sampling kernel (Figure 1E, H). Mathematically, the motion-corrupted image can be represented as

$$I_m^k = \sum_{t=0}^{t=T} I_t^k \circledast h_t = \sum_{t=0}^{t=T} (C_k M_t I_0) \circledast h_t \tag{1}$$

where I_m^k is the motion-corrupted image from the kth channel, C_k is the coil sensitivity map, T is the number of times the subject moved, I_t^k is the object (after motion) at time point t, I_0 is the motion-free image, M_t is the motion matrix at time t, $\circledast$ is the convolution operation and h_t is the convolution kernel determined by the time point t. The individual images from multiple channels can be combined to generate a single motion corrupted image I_m as

$$I_m = \left| \sum_{k=1}^{k=N_{ch}} C_k^* I_m^k \right| = \left| \sum_{k=1}^{k=N_{ch}} C_k^* * C_k \sum_{t=0}^{t=T} (M_t I_0) \circledast h_t \right| \tag{2}$$

where C_k^* is a conjugate of coil sensitivity and N_{ch} is the number of channels. Inspired by the similarity between the convolutional motion model presented in Equation (2) and a CNN, we investigated the use of a deep learning CNN to solve the motion correction problem. In order to estimate the motion-free image I_0 in Equation (2), the training process was applied to estimate a deep learning model that can identify and suppress the motion artefacts.

2.2 | Deep learning network architecture

The distribution of motion artefacts within the image is both local and global. The distribution of local artefacts varied within the different brain regions, with the cortical region of the brain having comparatively more artefacts than the mid-brain. The motion also resulted in ringing artefacts in the whole brain including outside the brain region in the background, which we consider as global artefacts. Since the neural network needs to learn both the global and local artefacts at multiple resolution scales, we designed a multiple resolution approach where the architecture in the

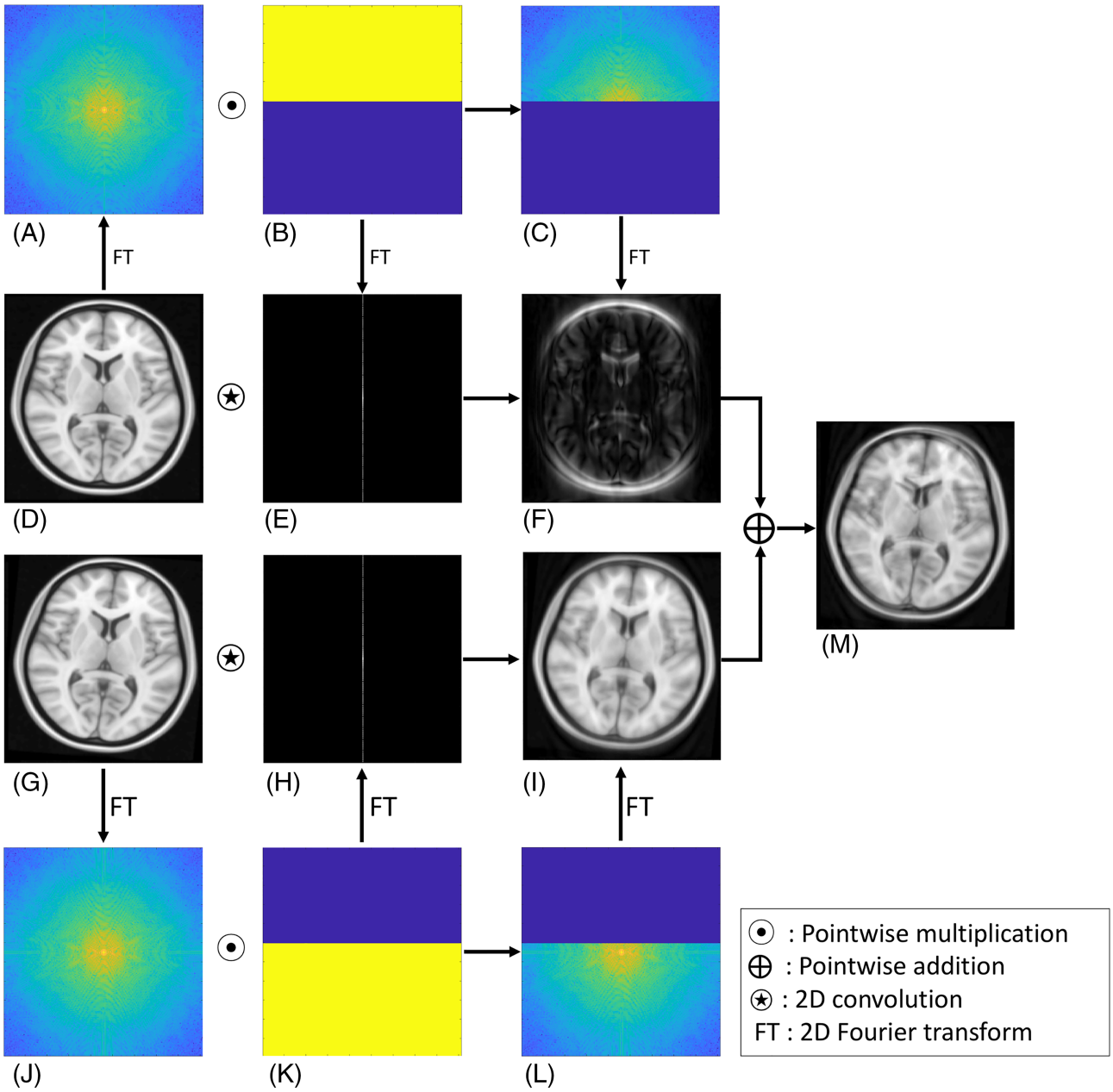

FIGURE 1 Convolution motion model, showing that the motion-corrupted image can be modelled as a linear combination of convolution of rigidly transformed, fully sampled images (D, G) and convolution kernels (E, H). A, *k*-space of fully sampled image; B, *k*-space sampling mask (yellow represents 1 and blue represents 0); C, partial *k*-space of the object; D, object of interest; E, convolution kernel determined by the sampling mask (B); F, image formed by partial *k*-space; G, object of interest rotated by $\Delta\theta$ and translated by Δx and Δy; H, I, J-L, the same as E, F, A-C respectively for the rotated and translated object (G); M, the motion-corrupted image

encoder consisted of parallel paths with different convolution kernel sizes. The different kernel sizes result in a different effective field of view at each resolution scale.

Specifically, we used Inception-ResNet V2 architecture for the encoder network, which consisted of parallel paths with different convolutional kernel sizes. Since the depth of the complete encoder-decoder architecture was large, we also used a residual connection,[45] as this only need to learns a residual function between the skip connected layers, resulting in faster training and convergence.[44] Residual blocks are the concatenation connection from the last layer of the encoder to the first layer of the decoder at the same scale. The net effect of the residual connection is that the network between the residual connections only learns the residual function instead of the complete function. In the context of MR image reconstruction, the network block between the residual connections only has to learn the difference between an artefact image and a clean image instead of learning to reconstruct a completely clean image.

An encoder-decoder architecture (Figure 2A) was trained for motion correction. The encoder network consisted of an Inception-ResNet V2 architecture,[44] and pre-trained weights trained on the ImageNet dataset[46] were used for the initialization of the encoder part of the encoder-decoder network. The inception architecture (Figure 2B) consists of parallel convolution pathways with different convolutional kernel sizes. The use of different kernel sizes results in different receptive fields, giving rise to multi-resolution feature maps. The decoder is made up of a series of convolution and upsampling layers (Figure 2C); two convolutions took place before every upsampling operation. The decoder network includes dropout layers with a dropout fraction of 0.2 after every convolution to avoid any overfitting. A stride of one was used for all the convolution layers and a stride of two for the upsampling layers in the decoder. Skip connections (concatenation) were made between the same scale feature maps from the encoder and decoder parts of the network. All the feature maps at a particular scale in the encoder were concatenated to the first layer of the decoder with the same scale. The loss function used to train the network was the mean absolute error between the output image and the target image (no motion image). The input to the network during the training was a 256 × 256 motion-corrupted image and the output was the motion-corrected image. During the inference, the input image of arbitrary size can be first zero padded to the nearest multiple of 32 and can be used as the input to the network.

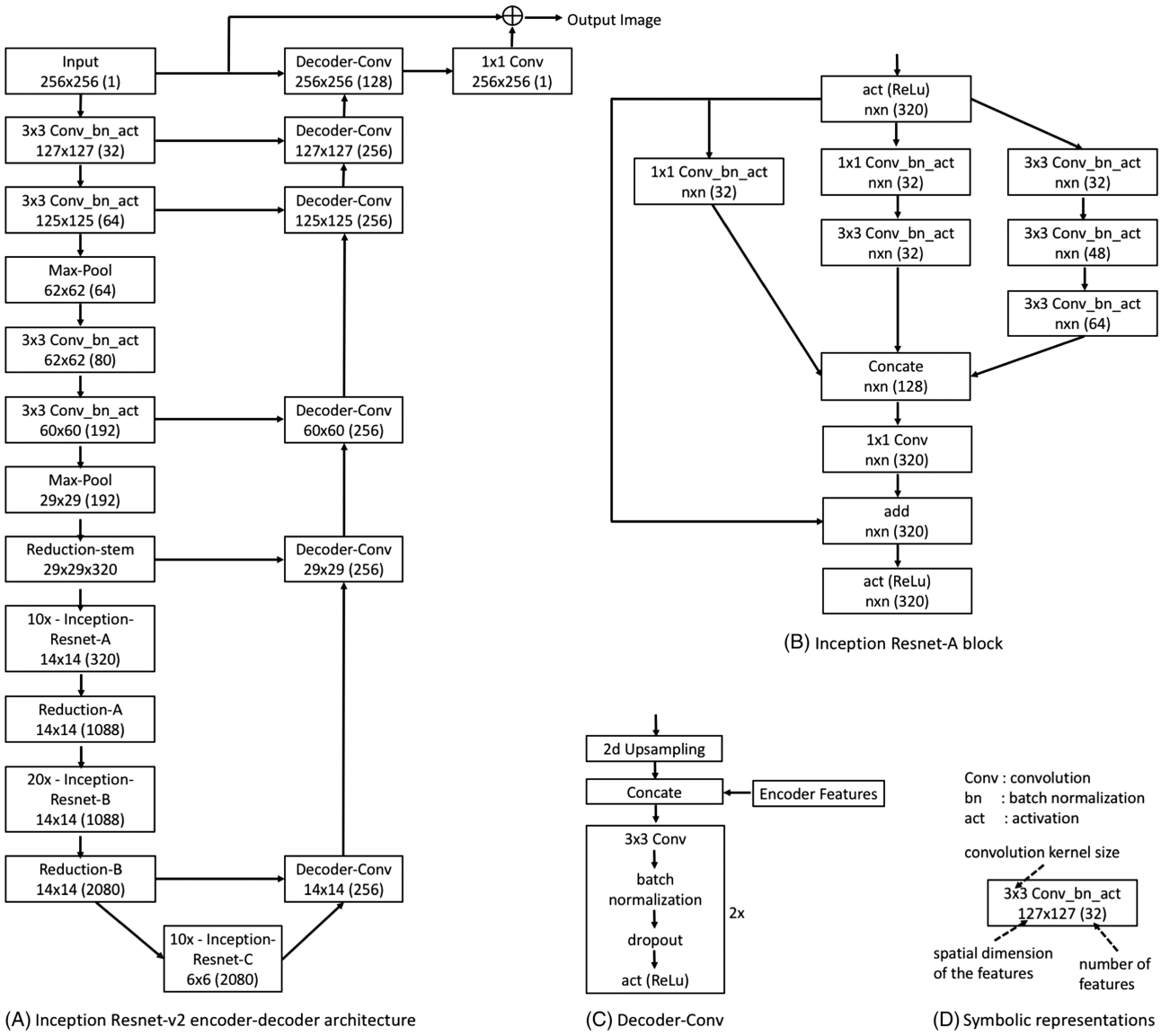

FIGURE 2 Network architecture. A, encoder-decoder CNN architecture for motion correction. The encoder is Inception-ResNet V2 architecture and the decoder is a stack of convolution and upsampling operations. B, Inception-ResNet-A block in the encoder consisting of parallel paths with different convolution kernel sizes. C, a block of convolution and upsampling in the decoder; six such blocks are used to form the decoder network. D, description of the symbolic representation used in the figure. The input to the network is a 256 × 256 motion-corrupted image and output is a clean motion-corrected image; mean absolute error was used as a loss function to train the network

2.3 | Motion simulation

Motion simulation was performed with randomly generated 3D rigid body motion between ±5 mm of translation and ± 5° of rotation on all three axes. The simulated motion range was consistent with a previous in vivo cohort study.[47] The MPRAGE images of size 256 × 256 × 256 were used for simulation purposes and the motion was simulated along the phase encode direction (anterior–posterior). Any motion within a single phase encode was neglected in the simulations. The process of motion simulation is depicted in Figure 3. The 3D image was rigidly transformed based on the six motion parameters (three translations and three rotation parameters) for each time point (phase encode). The *k*-space ($N_{pe} \times N_{fe} \times N_{se}$) corresponding to a rigidly transformed image was computed and the $1 \times N_{fe} \times N_{se}$ *k*-space corresponding to the current time point (phase encode) was extracted and appended to the motion-corrupted *k*-space. The appending of the *k*-space blocks ($1 \times N_{fe} \times N_{se}$) for all phase encodes constituted the motion-corrupted *k*-space data ($N_{pe} \times N_{fe} \times N_{se}$). The motion-corrupted, simulated *k*-space is inverse Fourier transformed to generate motion-corrupted images. The anterior–posterior (A-P) direction was considered as phase encodes, the left–right (L-R) direction was considered as slice encodes and the head-feet (H-F) direction was considered as frequency encodes for all the motion simulations.

Two different motion scenarios were simulated. (i) *Abrupt motion*, where the patient moves abruptly during the scan from time to time. We simulated up to 16 abrupt motion events during a single MPRAGE scan; the number of motion events ranging from 0 to 16 was selected using a uniform random number generator. An example of a motion curve is shown later in Figure 6. (ii) *Gradual motion*, where the patient makes gradual movements during the scan. An example of the motion curve for gradual motion is shown later in Figure 7.

In order to generate abrupt motion curves, a random number is generated for the number of motion events ranging between 1 and 16. Then, six motion parameters are randomly generated for each motion event, which results in an abrupt motion curve as shown later in Figure 7. For gradual motion, curves for abrupt motion are generated and then each motion curve is passed through a Gaussian filter, resulting in smoothly varying motion curves as shown later in Figure 6.

2.4 | Data preparation and training

3D MPRAGE datasets of 262 subjects were used to generate a training dataset, and 64 subjects were used to generate a validation dataset. Six different random motions were simulated for each subject, thus providing 1572 subjects for training and 384 subjects for validation. In order to generate motion-corrupted images, 3D *k*-space datasets of the images were distorted with a simulated 3D rigid body random motion as described in Subsection 2.3. Both the motion-corrupted image and the motion-free images were normalized to 0.5 mean and 1/6 standard deviation, so that approximately 95% of the intensity values lie between 0 and 1. After processing through the deep learning network, the output motion-corrected image was renormalized to its initial mean and standard deviation. The CNN was trained to suppress motion artefacts from individual 2D slices of the 3D MPRAGE dataset.

The following parameters were used for training the network in the Keras deep learning library with Tensorflow backend: Adam optimizer[48] with initial learning rate = 0.0001, batch size = 4, periodic decay = 0.96 after every epoch; each epoch consists of 2500 batches, ie 2500 weights update per epoch, and a total of 100 epochs on the NVIDIA Tesla V100 GPU. The training and validation losses are shown in Figure 4.

2.5 | Experimental data acquisition

In vivo experiments were performed on a Siemens Skyra 3 T MRI scanner (Siemens Healthineers, Erlangen, Germany) with a maximum gradient strength of 45 mT/m and a maximum slew rate of 200 mT/m/ms. Informed consent was obtained from volunteers in accordance with the

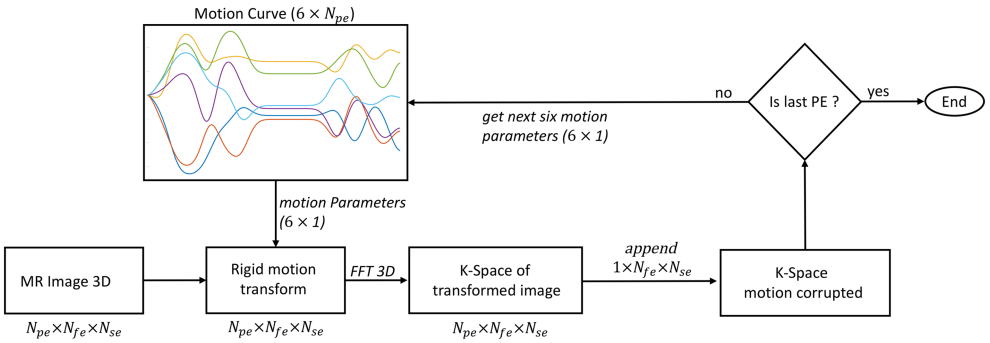

FIGURE 3 Flow chart of motion-corrupted data generation process: N_{pe}, number of phase encodes; N_{fe}, number of frequency encodes; N_{se}, number of slice encodes. It consists of rigid body transformation of a 3D MR image based on six motion parameters at a given time, followed by the extraction of $1 \times N_{fe} \times N_{se}$ *k*-space lines, which are appended for all time points, resulting in $N_{pe} \times N_{fe} \times N_{se}$ motion-corrupted *k*-space

institution's human research ethics policy. For experimental validation of the proposed motion correction technique, 3D MPRAGE datasets were acquired with $T_E/T_R/T_I$ = 3/2000/900 ms, matrix size = 256 × 256 × 192, resolution 1 mm isotropic, flip angle = 9°, acquisition time 8 min 32 s. Two separate scans were performed for each volunteer: one without motion, in which the volunteers were asked to remain still, and the second with motion, in which the volunteers were allowed to make natural movements during the scan. A total of four volunteers were scanned in this manner. The reconstructed images from the default reconstruction pipeline of the scanner were processed with the trained MoCoNet network, resulting in motion-corrected images. A structural similarity index (SSIM) was used as the quantitative metric for the evaluation of motion-corrected images.

3 | RESULTS

Both the magnitude of the motion and the *k*-space location where motion occurred determine the severity and the appearance of the artefact in the resulting image. For example, motion occurring in the outer *k*-space (ie high-frequency components) manifests as ringing in the reconstructed image (Figure 5A-D). Motion occurring near to the centre of the *k*-space (ie low-frequency components) manifests as blurring (Figure 5E-H). When the magnitude of the motion was increased while the *k*-space location where motion was occurring remained fixed (shown in Figure 5A-D), the ringing increased. In contrast, when the *k*-space location where motion was occurring was varied and the magnitude of the motion remained fixed,

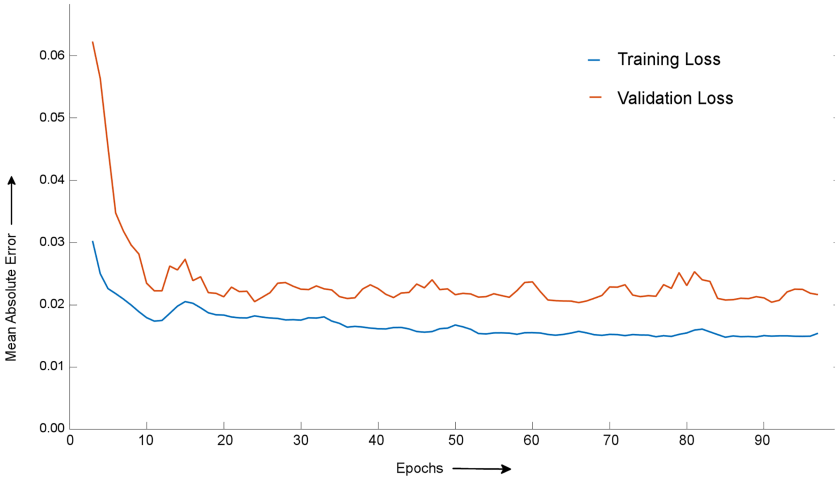

FIGURE 4 Plot of training and validation loss for the network trained with periodic learning rate decay and the Adam optimizer. Each epoch consists of 2500 iterations

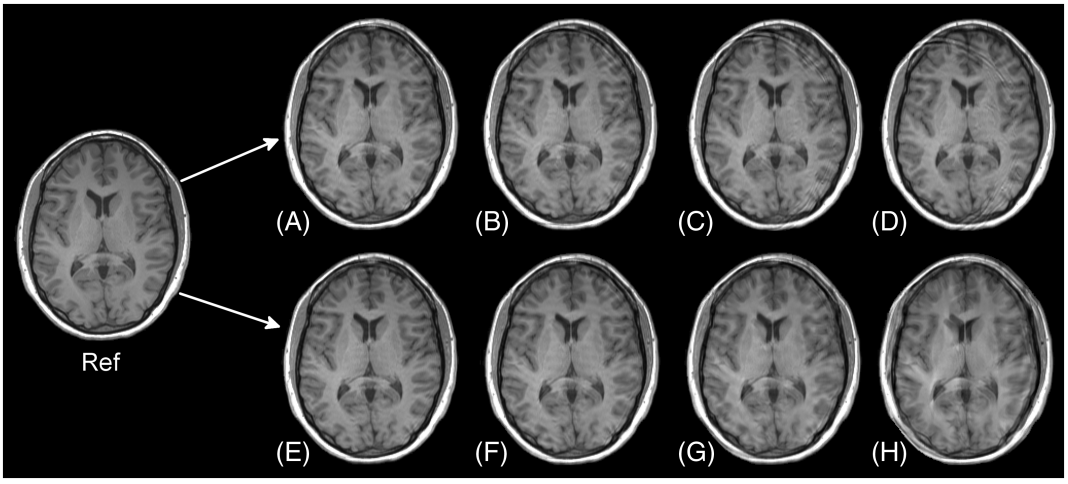

FIGURE 5 Simulation demonstrating the effect of a single motion event on reconstructed images. Ref, reference image. A-D, reconstructed images as a result of increasing the magnitude of the motion at a fixed phase encode (100th) in the acquisition process. The magnitude was increased one-, two-, four- and sixfold respectively. E-H, reconstructed images due to a fixed amount of motion at different phase encodes in the acquisition process. The phase encodes were 90th, 100th, 110th and 120th respectively

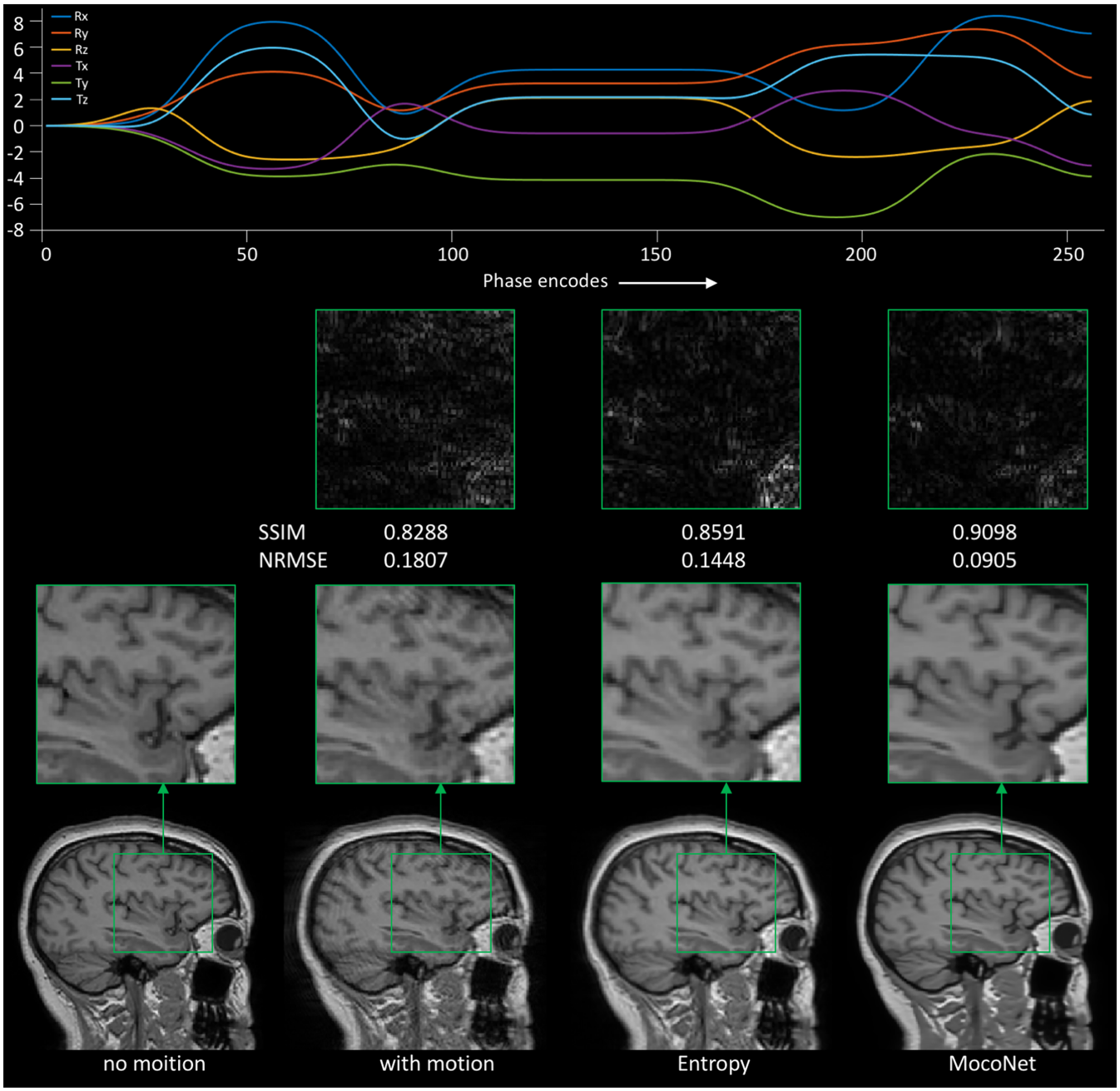

FIGURE 6 Simulated motion-correction results for gradual motion; the plot at the top shows the simulated motion parameters, where R_x, R_y, R_z are rotation in degrees, and T_x, T_y, T_z are translations in mm in the x, y, z directions, respectively. The x-axis represents time in terms of phase encodes (PE) and the time would be equal to PE T_R for the MPRAGE sequence. No motion, the motion-free clean image; with motion, the image corrupted by the simulated motion; entropy, the image corrected using the entropy minimization method; MoCoNet, the image corrected using the deep learning method; NRMSE, normalized root mean squared error

the blurring artefacts increased as the location of the motion in k-space moved towards the centre (shown in Figure 5E-H). From this simulation, it can be inferred that the location of the motion in k-space largely determines the appearance of the motion artefacts (ringing or blurring) and the magnitude of the motion largely determines the spatial distribution and severity of the motion artefacts.

Although the network was trained in the range of ±5 mm translation and ± 5° rotation, the predictions were performed on a wider range (±8 mm and ± 8°) to test the generalization capability of the network to unseen cases. The results of the motion artefact suppression are shown in Figures 6 and 7 for gradual motion and abrupt motion scenarios, respectively. In order to assess the performance of the deep learning method, the results were compared with the entropy minimization method,[27] which is also a post-processing based method. From Figures 6 and 7, we can see that the ringing artefacts were more effectively suppressed with MoCoNet, compared with the entropy minimization method, and that the SSIM indices were higher for the MoCoNet reconstructions. It is evident from the enlarged images in Figures 6 and 7 that the motion artefacts

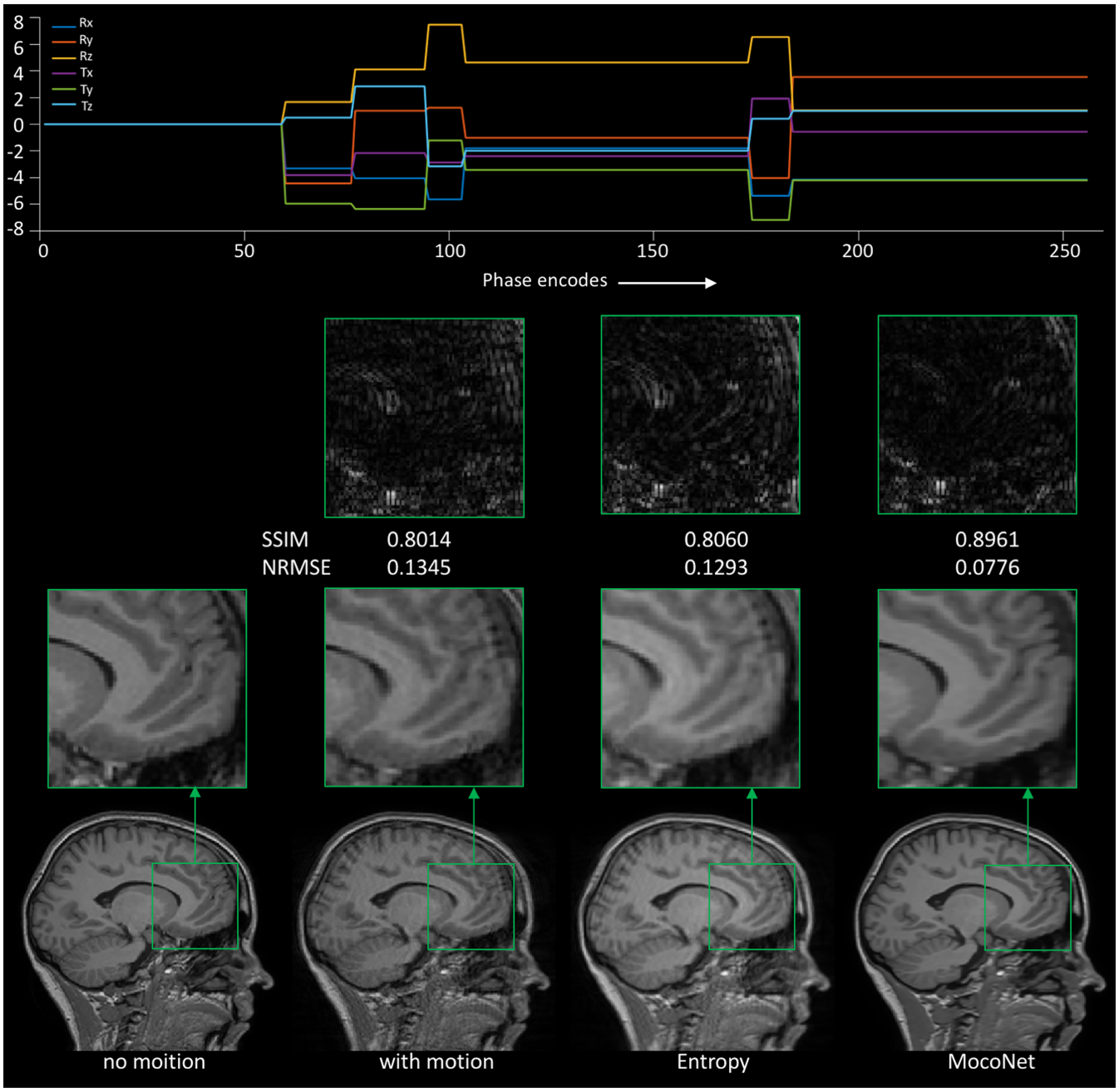

FIGURE 7 Simulated motion-correction results for abrupt motion; the plot at the top shows the simulated motion parameters, where R_x, R_y, R_z are rotation in degrees, and T_x, T_y, T_z are translations in mm in the x, y, z directions, respectively. The x-axis represents time in terms of phase encodes (PE) and the time would be equal to PE T_R for the MPRAGE sequence. No motion, the motion free clean image; with motion, the image corrupted by the simulated motion; entropy, the image corrected using the entropy minimization method; MoCoNet, the image corrected using the deep learning method; NRMSE, normalized root mean squared error

were successfully removed using MoCoNet. In contrast, the entropy minimization method resulted in residual ringing, particularly in the abrupt motion scenario. The entropy minimization method performed better for the gradual motion compared with the abrupt motion. However, MoCoNet outperformed the entropy-based method in both of the motion scenarios in terms of the SSIM index and normalized root mean squared error. Furthermore, the performance of MoCoNet was also found to be consistent in both gradual motion and abrupt motion simulations. The consistent performance of MoCoNet highlights the fact that it is robust to both gradual and abrupt motion scenarios. The mean of the SSIM index for the corrected images was improved from 0.9058 to 0.9338 and the mean of the normalized root mean squared error decreased from 0.1993 to 0.1336.

Figure 8 shows the performance of the motion correction with varying levels of motion on the validation dataset. Artefacts are shown to increase from left to right. In all cases, MoCoNet was able to substantially remove the artefacts from the motion-corrupted images, resulting in

corrected images with higher SSIM and lower normalized root mean squared error. The improved delineation between white and grey matter is clearly noticeable in the enlarged motion-corrected images (Figure 8B, C) compared with the motion-corrupted images. As shown in Figure 9, using a paired t-test $p < 0.001$, which was calculated on 384 validation subjects, the SSIM statistical parameters (median, 25th, 75th and span) for MoCoNet corrected images improved compared with the uncorrected images. Likewise, a similar trend was found to be true for normalized root mean squared errors, with a paired t-test $p < 0.001$.

The proposed MoCoNet correction was also evaluated on the experimental datasets acquired as described in Subsection 2.5. The motion-corrupted DICOM (Digital Imaging and Communications in Medicine) images were processed with MoCoNet, which suppressed the motion artefacts in the images. Motion-corrected images from two volunteers for all three orthogonal slices are shown in Figures 10 and 11, along with the SSIM with respect to the no-motion image. Visual inspection of the enlarged images shows that the images corrected using the MoCoNet method have substantially reduced motion artefacts compared with the uncorrected images. The SSIM index demonstrates the improvements in image quality as a result of using MoCoNet. For Subject 1 the mean SSIM was increased from 0.9012 to 0.9396, and for Subject 2 it was increased from 0.8330 to 0.8895. Similarly, the mean normalized root mean squared error was decreased from 0.1021 to 0.0759 for Subject 1 and from 0.0976 to 0.0774 for Subject 2.

We also compared the computation times of the different correction methods. The entropy minimization is an iterative algorithm and thus the computation time is large. For the image size of 256 × 256 × 150, the execution time was 12 h 46 min (46 000 s) with a CPU for the entropy minimization method, while it was only 185 s for the MoCoNet method—a 250-fold speed-up in execution time. The implementation of MoCoNet on a GPU Nvidia-V100 further reduced the execution time to 15 s, making its use feasible in real-time clinical applications.

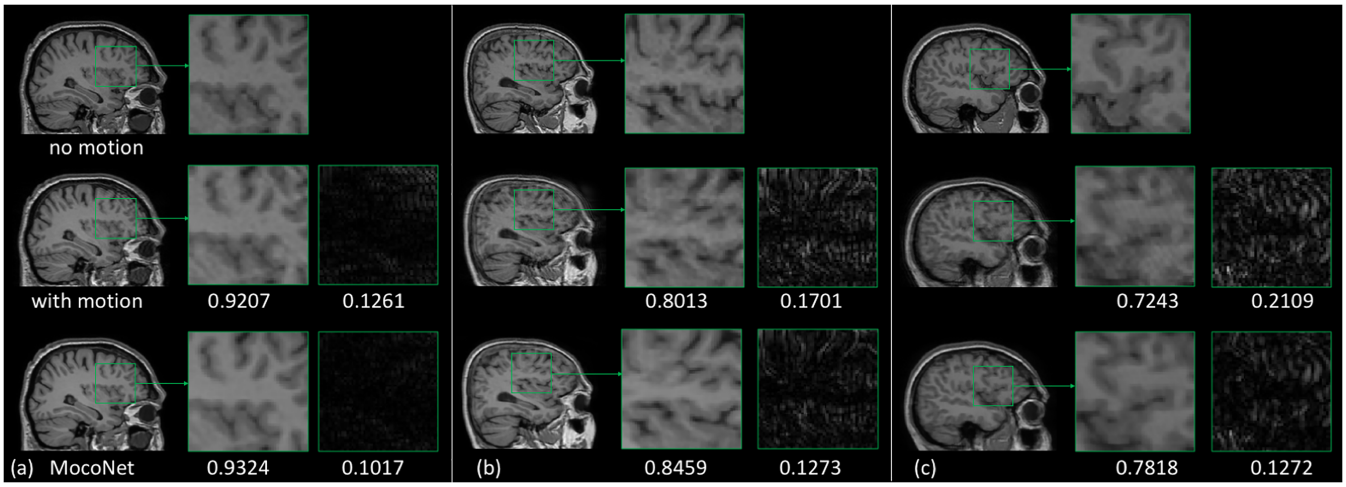

FIGURE 8 Simulated motion-correction results at different levels of motion artefacts on the validation dataset, severity increasing from left (a) to right (C): No motion, reference image without any motion artefacts; with motion, image with simulated motion artefacts; MoCoNet, images corrected with the proposed motion correction method. The numbers below the enlarged images show the SSIM index and the numbers below the difference images show the normalized root mean squared error

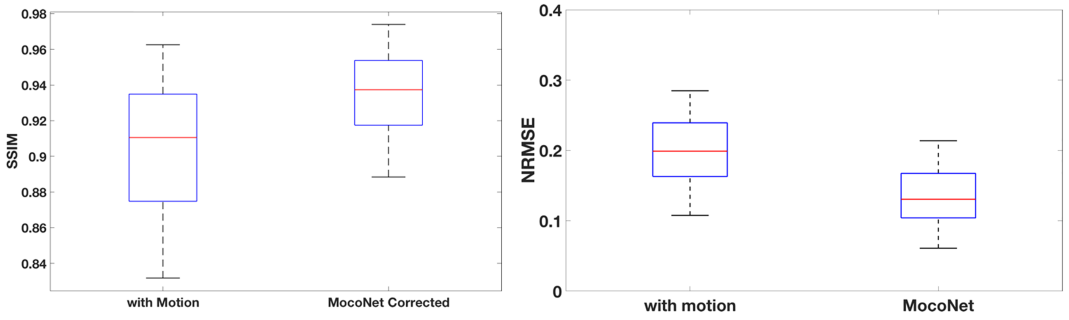

FIGURE 9 SSIM and normalized root mean squared error (NRMSE) for the motion-corrupted versus motion-corrected image: The top of the box represents the 25th percentile; the bottom of box represents the 75th percentile; the red line represents the median. The paired t-test p value was less than 0.001 for both SSIM and NRMSE

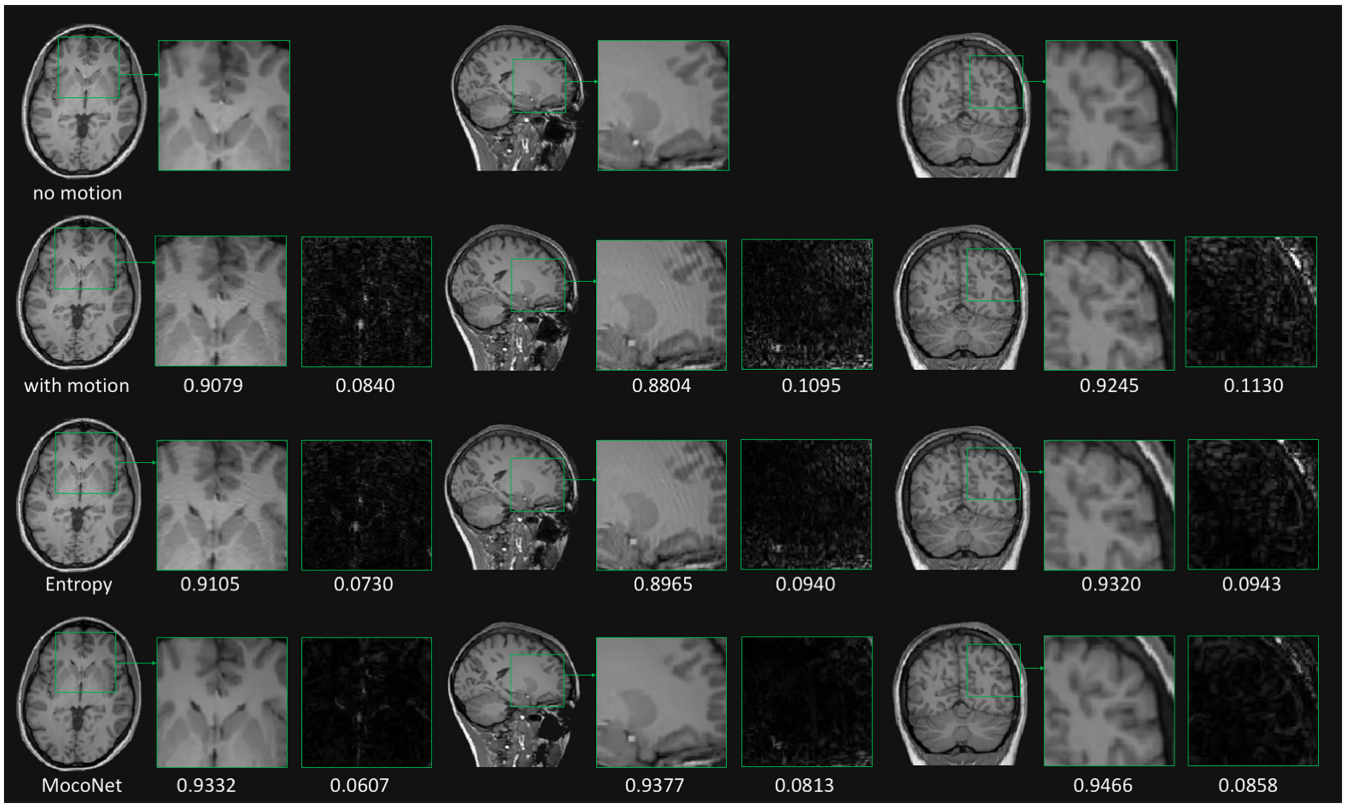

FIGURE 10 Experimental results for volunteer I, on whom two scans were performed, one with motion and one without motion: No motion, images from the scan in which the volunteer did not move; with motion, images from the scan in which the volunteer moved randomly at his/her own discretion; MoCoNet, images from the scan where the volunteer moved and the motion-corrupted images were subsequently corrected using the MoCoNet network; entropy, images corrected with the entropy minimization method. The numbers below the enlarged images show the SSIM index and the numbers below the difference images show the normalized root mean squared error

4 | DISCUSSION

We have demonstrated that the motion artefacts in MR images can be modelled as a linear combination of convolution between transformed images and convolution kernels. This convolution motion model forms the basis of using a CNN to learn a model that can correct for motion artefacts. Simulation results confirmed that motion at high frequency results in ringing, whereas motion at low frequency results in blurring and ghosting. The encoder in MoCoNet separates the motion-corrupted images into features consisting of motion artefacts. The neuron in the decoder network only allows the motion-free features to be integrated into the final motion-corrected images. Although the network was trained on simulated motion-corrupted images without considering any spin history effect, the network was able to correct for the real motion, demonstrating the robustness of the MoCoNet method.

Current state-of-the-art motion correction methods that rely on the image or *k*-space navigators interfere with the MR protocol to the detriment of contrast, scan time and SNR. Other methods based on external sensors need calibration and are complex to set up.[49] The method presented in this work is an innovative solution to the problem of motion during MRI acquisitions. First, unlike navigators, it requires neither the development nor the modification of acquisition sequences or image reconstruction methods. Second, it does not interfere with MR protocol parameters such as $T_R/T_I/T_E$, scan time or contrast. Third, it does not require any external hardware devices, thus making it relatively inexpensive to implement.

We have presented a novel method based on deep learning that is capable of removing motion artefacts from motion-corrupted MR images. The proposed method was validated on both the simulation and the experimental datasets. We have demonstrated that the network trained with simulated motion was able to generalize and was able to correct for the real motion that occurred during a cohort of subjects that were scanned. Although the motion simulation model captures a wide range of realistic motion scenarios, MR signal variations that may occur due to spin wraps, magnetic field changes and signal loss caused by motion are not modelled. This may degrade the performance of the motion suppression in a few real motion scenarios. Integrating the Bloch equation into the motion model would lead to more accurate data generation and hence a more robust motion suppression system. The proposed method can work independently without the need for any motion parameter acquisition setup. Alternatively, it can also be used as a complementary method with existing motion correction setups such as navigators and external devices.

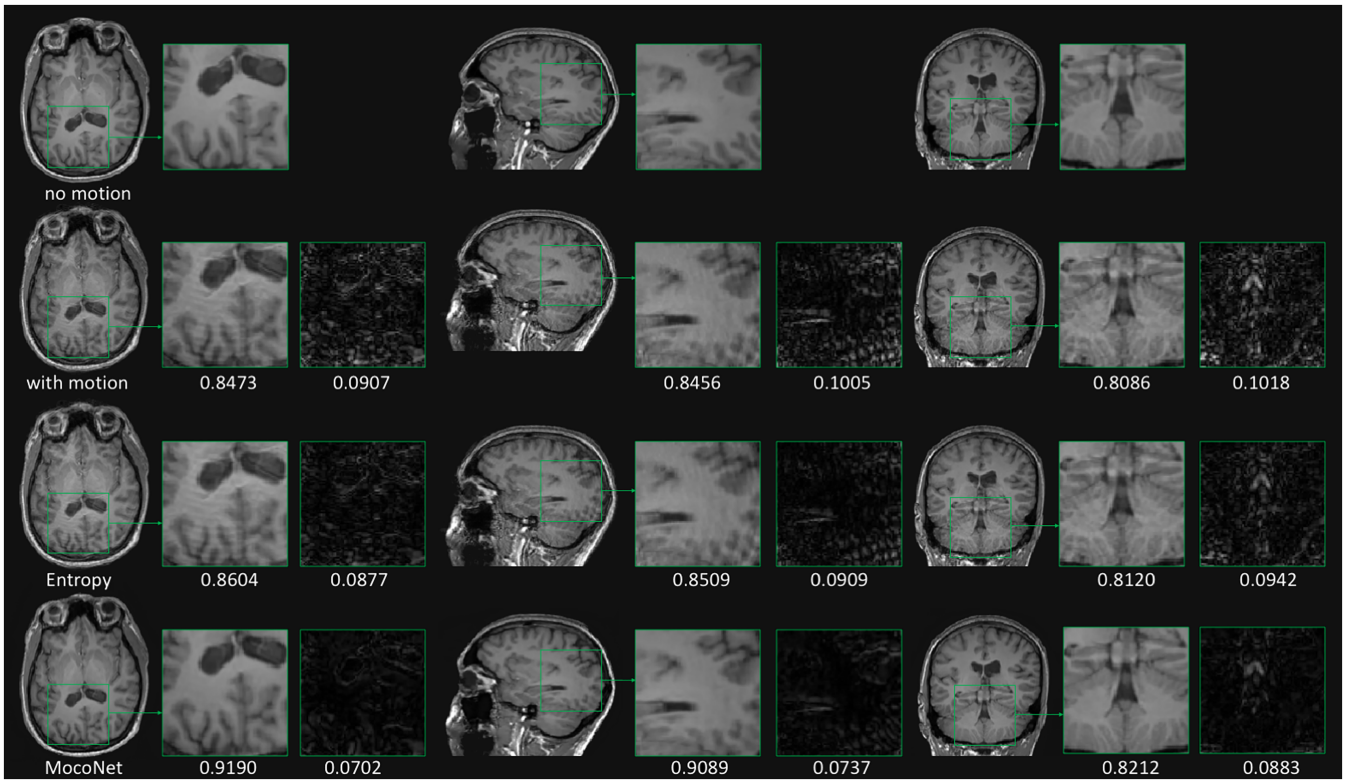

FIGURE 11 Experimental results for volunteer II, on whom two scans were performed, one with motion and one without motion: No motion, images from the scan in which the volunteer did not move; with motion, images from the scan in which the volunteer moved randomly at his/her own discretion; MoCoNet, images from the scan where the volunteer moved and the motion-corrupted images were subsequently corrected using the MoCoNet network; entropy, images corrected with the entropy minimization method. The numbers below the enlarged images show the SSIM index and the numbers below the difference images show the normalized root mean squared error

Compared with the entropy-based technique, the MoCoNet method is significantly more effective in removing motion artefacts. However, there are occasionally residual motion artefacts left, especially when motion occurred at low-frequency *k*-space locations. In order to make the motion correction more robust, a hybrid approach of combining MoCoNet with navigators could result in a better motion correction system. It is known that navigators are good at picking up large motion efficiently, hence in a hybrid motion correction system the images can be prospectively or retrospectively corrected using navigators and the residual motion in the resulting images can be corrected with MoCoNet. Another source of residual artefact could be the patient motion, which is widely different from the simulated motion case. Residual artefacts can also result from a different acquisition ordering of the phase/slice encodings that can be minimized by fine-tuning the network with a dataset corresponding to the acquisition order.

5 | CONCLUSION

An efficient data-driven technique based on a CNN has been developed for MR motion correction without the need for external sensors or internal navigators. The network is based on the Inception-ResNet architecture and is fully trained using simulated motion. Here we have demonstrated the application of the technique to 3D T_1 MPRAGE experimental datasets. MoCoNet is computationally efficient and the shorter inference time makes it practical for routine clinical practice.

ACKNOWLEDGEMENT

We would like to thank Alexander Loktyushin (Max Planck Institute for Intelligent Systems) for sharing the code for 3D motion correction using entropy focus criteria.

ORCID

Kamlesh Pawar https://orcid.org/0000-0001-6199-2312
Zhaolin Chen https://orcid.org/0000-0002-0173-6090
N. Jon Shah https://orcid.org/0000-0002-8151-6169

REFERENCES

1. Gupta MP, Shringirishi MM. Implementation of brain tumor segmentation in brain MR images using k-means clustering and fuzzy c-means algorithm. *Int J Comput Technol*. 2013;5(1):54-59.
2. Godenschweger F, Kägebein U, Stucht D, et al. Motion correction in MRI of the brain. *Phys Med Biol*. 2016;61(5):R32.
3. Andre JB, Bresnahan BW, Mossa-Basha M, et al. Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations. *J Am Coll Radiol*. 2015;12:689-695.
4. Ehman RL, Felmlee JP. Adaptive technique for high-definition MR imaging of moving structures. *Radiology*. 1989;173:255-263.
5. Pipe JG. Motion correction with PROPELLER MRI: application to head motion and free-breathing cardiac imaging. *Magn Reson Med*. 1999;42:963-969.
6. Zaitsev M, Zilles K, Shah NJ. Shared k-space echo planar imaging with keyhole. *Magn Reson Med*. 2001;45(1):109-117.
7. Nunes RG, Jezzard P, Behrens TE, Clare S. Self-navigated multishot echo-planar pulse sequence for high-resolution diffusion-weighted imaging. *Magn Reson Med*. 2005;53(6):1474-1478.
8. Fu ZW, Wang Y, Grimm RC, et al. Orbital navigator echoes for motion measurements in magnetic resonance imaging. *Magn Reson Med*. 1995;34:746-753.
9. Johnson PM, Liu J, Wade T, Tavallaei MA, Drangova M. Retrospective 3D motion correction using spherical navigator echoes. *Magn Reson Imaging*. 2016;34:1274-1282.
10. Welch EB, Manduca A, Grimm RC, Ward HA, Jack CR. Spherical navigator echoes for full 3-D rigid body motion measurement in MRI. In: Niessen WJ, Viergever MA, eds. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2001. Lecture Notes in Computer Science*. Vol.2208 Berlin, Germany: Springer; 2001:1235-1236.
11. Costa AF, Petrie DW, Yen YF, Drangova M. Using the axis of rotation of polar navigator echoes to rapidly measure 3D rigid-body motion. *Magn Reson Med*. 2005;53:150-158.
12. Petrie DW, Costa AF, Takahashi A, Yen YF, Drangova M. Optimizing spherical navigator echoes for three-dimensional rigid-body motion detection. *Magn Reson Med*. 2005;53:1080-1087.
13. Liu J, Drangova M. Rapid six-degree-of-freedom motion detection using prerotated baseline spherical navigator echoes. *Magn Reson Med*. 2011;65:506-514.
14. White N, Roddey C, Shankaranarayanan A, et al. PROMO: real-time prospective motion correction in MRI using image-based tracking. *Magn Reson Med*. 2010;63:91-105.
15. Thesen S, Heid O, Mueller E, Schad LR. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magn Reson Med*. 2000;44:457-465.
16. Gallichan D, Marques JP, Gruetter R. Retrospective correction of involuntary microscopic head movement using highly accelerated fat image navigators (3D FatNavs) at 7T. *Magn Reson Med*. 2016;75:1030-1039.
17. Tisdall MD, Hess AT, Reuter M, Meintjes EM, Fischl B, van der Kouwe AJW. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn Reson Med*. 2012;68:389-399.
18. Buschbeck RP, Yun SD, Jon SN. 3D rigid-body motion information from spherical Lissajous navigators at small k-space radii: a proof of concept. *Magn Reson Med*. 2019;82:1462-1470.
19. Schulz J, Siegert T, Reimer E, et al. An embedded optical tracking system for motion-corrected magnetic resonance imaging at 7T. *Magn Reson Mater Phys Biol Med*. 2012;25:443-453.
20. Zaitsev M, Dold C, Sakas G, Hennig J, Speck O. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *Neuroimage*. 2006;31:1038-1050.
21. Maclaren J, Armstrong BSR, Barrows RT, et al. Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain. *PLoS ONE*. 2012;7:e48088.
22. Qin L, Van Gelderen P, Derbyshire JA, et al. Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system. *Magn Reson Med*. 2009;62:924-934.
23. Ooi MB, Krueger S, Thomas WJ, Swaminathan SV, Brown TR. Prospective real-time correction for arbitrary head motion using active markers. *Magn Reson Med*. 2009;62:943-954.
24. Sengupta S, Tadanki S, Gore JC, Welch EB. Prospective real-time head motion correction using inductively coupled wireless NMR probes. *Magn Reson Med*. 2014;72:971-985.
25. Haeberlin M, Kasper L, Barmet C, et al. Real-time motion correction using gradient tones and head-mounted NMR field probes. *Magn Reson Med*. 2015;74:647-660.
26. Atkinson D, Hill DL, Stoyle PN, Summers PE, Keevil SF. Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion. *IEEE Trans Med Imaging*. 1997;16(6):903-910.
27. Loktyushin A, Nickisch H, Pohmann R, Scholkopf B. Blind retrospective motion correction of MR images. *Magn Reson Med*. 2013;70(6):1608-1618.
28. Quan TM, Nguyen-Duc T, Jeong W-K. Compressed sensing MRI reconstruction with cyclic loss in generative adversarial networks. *IEEE Trans Med Imaging*. 2018;37:1488-1497.
29. Yang G, Yu S, Dong H, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans Med Imaging*. 2018;37:1310-1321.
30. Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Processing*. 2017;26:4509-4522.
31. LeCun YA, Bengio Y, Hinton GE. Deep learning. *Nature*. 2015;521:436-444.
32. Lee D, Yoo J, Ye JC. Deep residual learning for compressed sensing MRI. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. 2017:15–18.
33. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature*. 2018;555:487-492.
34. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med*. 2018;79:3055-3071.
35. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88.

36. Sciarra A, Mattern H, Speck O. Machine learning algorithms for detection of motion artifacts: a general approach. Poster presented at: joint annual meeting ISMRM-ESMRMB; June 16-21, 2018; Paris, France
37. Bilgic B, Cauley SF, Chatnuntawech I, et al. Combining MR physics and machine learning to tackle intractable problems. *Proc Int Soc Magn Reson Med*. 2018;3374.
38. Haskell MW, Cauley SF, Bilgic B, et al. Network accelerated motion estimation and reduction (NAMER): convolutional neural network guided retrospective motion correction using a separable motion model. *Magn Reson Med*. 2019;82:1452-1461.
39. Johnson PM, Drangova M, editors. Motion correction in MRI using deep learning. *Proc Int Soc Magn Reson Med*. 2018.
40. Johnson PM, Drangova M. Conditional generative adversarial network for 3D rigid-body motion correction in MRI. *Magn Reson Med*. 2019;82(3):901-910.
41. Küstner T, Armanious K, Yang J, Yang B, Schick F, Gatidis S. Retrospective correction of motion-affected MR images using deep learning frameworks. *Magn Reson Med*. 2019;82:1527-1540.
42. Pawar K, Chen Z, Shah NJ, Egan G. Motion correction in MRI using deep convolutional neural network. *Proc Int Soc Magn Reson Med*. 2018;26:1174.
43. Sommer K, Brosch T, Wiemker R, et al. Correction of motion artifacts using a multi-resolution fully convolutional neural network. *Proc Int Soc Magn Reson Med*. 2018.
44. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the impact of residual connections on learning. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*. AAAI; 2017:4278–4284.
45. He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). IEEE; 2016:770–778.
46. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition; IEEE; 2009.
47. Chen Z, Sforazzini F, Baran J, Close T, Shah NJ, Egan GF. MR-PET head motion correction based on co-registration of multicontrast MR images. *Hum Brain Mapp*. In press.
48. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:14126980. 2014.
49. Maclaren J, Aksoy M, Ooi MB, Zahneisen B, Bammer R. Prospective motion correction using coil-mounted cameras: cross-calibration considerations. *Magn Reson Med*. 2018;79(4):1911-1921.

How to cite this article: Pawar K, Chen Z, Shah NJ, Egan GF. Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation. *NMR in Biomedicine*. 2022;35(4):e4225. https://doi.org/10.1002/nbm.4225