IOP Publishing Phys. Med. Biol. 65 (2020) 08NT03 https://doi.org/10.1088/1361-6560/ab7d13

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

NOTE

Single patient convolutional neural networks for real-time MR reconstruction: coherent low-resolution versus incoherent undersampling

Bryson Dietz[1,4], **Jihyun Yun**[2], **Eugene Yip**[2], **Zsolt Gabos**[1], **B Gino Fallone**[1,2,3] **and Keith Wachowicz**[1,2]

[1] Department of Oncology, University of Alberta, Edmonton, Canada
[2] Department of Medical Physics, Cross Cancer Institute, Edmonton, Canada
[3] Department of Physics, University of Alberta, Edmonton, Canada
[4] Author to whom any correspondence should be addressed.

E-mail: bdietz@ualberta.ca

Keywords: MRI, reconstruction, neural networks

RECEIVED
23 September 2019

REVISED
30 January 2020

ACCEPTED FOR PUBLICATION
5 March 2020

PUBLISHED
23 April 2020

Abstract

Accelerated MRI involves undersampling k-space, creating unwanted artifacts when reconstructing the data. While the strategy of incoherent k-space acquisition is proven for techniques such as compressed sensing, it may not be optimal for all techniques. This study compares the use of coherent low-resolution (coherent-LR) and incoherent undersampling phase-encoding for real-time 3D CNN image reconstruction. Data were acquired with our 3 T Philips Achieva system. A retrospective analysis was performed on six non-small cell lung cancer patients who received dynamic acquisitions consisting of 650 free breathing images using a bSSFP sequence. We retrospectively undersampled the data by 5x and 10x acceleration using the two phase-encoding schemes. A quantitative analysis was conducted evaluating the tumor segmentations from the CNN reconstructed data using the Dice coefficient (DC) and centroid displacement. The reconstruction noise was evaluated using the structural similarity index (SSIM). Furthermore, we qualitatively investigated the CNN reconstruction using prospectively undersampled data, where the fully sampled training data set is acquired separately from the accelerated undersampled data. The patient averaged DC, centroid displacement, and SSIM for the tumor segmentation at 5x and 10x was superior using coherent low-resolution undersampling. Furthermore, the patient-specific CNN can be trained in under 6 h and the reconstruction time was 54 ms per image. Both the incoherent and coherent-LR prospective CNN reconstructions yielded qualitatively acceptable images; however, the coherent-LR reconstruction appeared superior to the incoherent reconstruction. We have demonstrated that coherent-LR undersampling for real-time CNN image reconstruction performs quantitatively better for the retrospective case of lung tumor segmentation, and qualitatively better for the prospective case. The tumor segmentation mean DC increased for all six patients at 5x acceleration and the temporal (dynamic) variance of the segmentation was reduced. The reconstruction speed achieved for our current implementation was 54 ms, providing an acceptable frame rate for real-time on-the-fly MR imaging.

1. Introduction

The use of convolutional neural networks (CNNs) is becoming an increasingly popular trend in image processing and reconstruction (Kensuke *et al* 2017, Valverde *et al* 2017, Qin *et al* 2017, Schlemper *et al* 2018, Zeng *et al* 2018, Dietz *et al* 2019, Eppenhof *et al* 2019). As linac-MR (LMR) systems are becoming increasingly popular with many sites across the globe, there will be an increase in demand for real-time MR imaging techniques for adaptive radiotherapy purposes. There are several LMR systems available ranging in field strengths from 0.35 T to 1.5 T (Fallone *et al* 2009, Raaymakers *et al* 2009, Fallone 2014, Keall *et al* 2014, Mutic *et al* 2014, Acharya *et al* 2016, Fischer-Valuck *et al* 2017, Liney *et al* 2018). Real-time on-the-fly

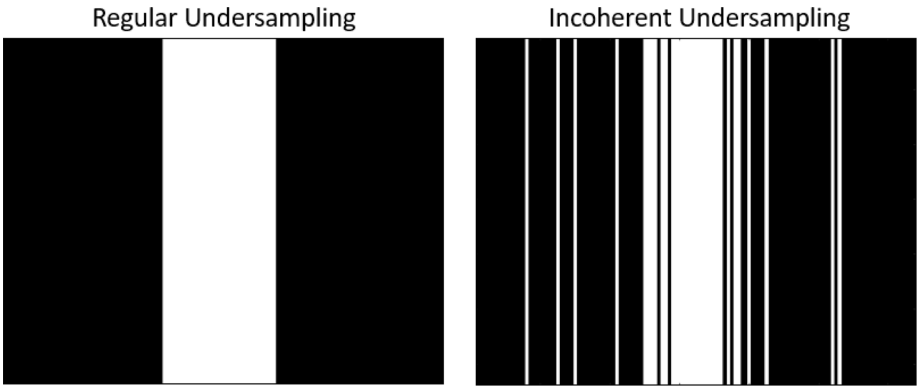

Figure 1. Coherent-LR verses incoherent phase encoding schemes. At 4x acceleration, each undersampling scheme uses 32 phase encodes (assuming 128 × 128 image). The coherent-LR undersampling only samples the central band of k-space, whereas the incoherent undersampling is focused on the central band, but also utilizes a pseudo-random sampling of the outer k-space regions.

imaging using MRI is challenging due to the slow acquisition of k-space. One way to speed up the acquisition is to skip lines of k-space, also known as phase-encode lines. However, this results in a violation of the Nyquist criterion, resulting in artefacts occurring throughout the image. Parallel imaging is a common technique that is used to correct for fold-over aliasing using a coherently undersampled acquisition (Griswold *et al* 1999, Pruessmann *et al* 1999). Parallel imaging requires intricate coil construction to achieve high acceleration factors and can still suffer from suboptimal results (Glockner *et al* 2005).

Another common technique for reconstructing undersampled data is compressed sensing (CS) (Lustig *et al* 2007); however, CS reconstruction can be time restrictive in a real-time application. There have been CS reconstruction techniques that address reconstruction speed; however, higher acceleration factors still suffer from artefacts (Yip *et al* 2014, 2017). The CS technique utilizes incoherently acquired k-space, with a greater density of the k-space acquisition near the central (low spatial resolution) region and pseudo-random sampling of the outer (high spatial resolution) k-space. Reconstruction of the incoherent acquisition results in a noise-like artefact occurring throughout the image that must then be suppressed by the reconstruction algorithm. The implementation of incoherent undersampling is not without challenges: its use can present confounding eddy current artefacts. The increased eddy current contributions are caused by the rapid gradient switching during the inconsistent phase-encoding jumps required by the pseudo-random sampling. These eddy currents can result in suboptimal reconstruction when using techniques that rely on previously acquired (fully sampled) data, such as CNN reconstruction techniques. The suboptimal reconstruction is due to the fully sampled acquisition used for training having a differing eddy current contribution from the incoherently acquired (undersampled) data. Techniques can be used to mitigate eddy currents during imaging (Bieri *et al* 2005); however, they may not reduce all the eddy current artefacts and can be restrictive in terms of sequence playout.

One way to avoid issues associated with incoherent encoding, is to acquire only the central portion of k-space and exclude the outer regions collected during the incoherent acquisition. This alternative k-space acquisition technique results in a low-resolution reconstruction in image space. While this does not give sufficient data to generate a fully-resolved reconstruction on its own, when paired with a CNN trained on a previously acquired fully-sampled data set the un-sampled data can be inferred during reconstruction, whether in the high spatial resolution region alone or scattered throughout k-space as in the incoherent-sampled case. An example of the two phase encoding (PE) schemes are shown in figure 1. This coherent low-resolution (hereafter known as coherent-LR) undersampling is ideal, as it is simple to implement on any MR system and does not lead to eddy current artefacts, due to both the fully sampled and coherent-LR undersampled data having identical sampling intervals. The previously acquired fully sampled data, on which the CNN is trained, will more closely resemble the data acquired via the coherent-LR undersampled acquisition. It is the purpose of the trained CNN to generate this high resolution data based on fully sampled data, from which relationships between low resolution structure and high resolution detail can be established It is the authors' hypothesis that this will be a more straightforward task for the CNN to fulfil and may result in a more robust algorithm. Our study investigates and compares the use of incoherent and coherent-LR undersampled k-space acquisitions for real-time on-the-fly 3D CNN reconstruction, both retrospectively and prospectively.

A common paradigm for the use of (convolutional) neural networks involves training a network using a large cohort of patient data in order to recover information from a single patient. Our paradigm differs, in

that we focus on a personalized healthcare approach, such that a CNN is specifically trained for each patient. This requires having a previously acquired fully sampled dynamic data set for training the CNN, which is then applied to reconstruct the accelerated data acquisition.

In this study we present the use of 3D CNNs to retrospectively reconstruct undersampled dynamic data from six non-small cell lung cancer patients on a patient-by-patient basis. A quantitative comparison of coherent-LR and incoherent acquisitions for the use of CNN MR image reconstruction was conducted. We focus on the ability to track and segment a tumorous region located in the lung. Lung was specifically chosen, as breathing motion is a challenging component to account for using conventional linac systems. We further validate and compare the two encoding schemes in the realistic prospective imaging scenario, where the fully sampled training data is acquired separately from the accelerated undersampled data. By reducing the images acquisition and reconstruction time, a higher frame rate with low latency can be achieved, which is of importance for adaptive radiotherapy The aim of this study is to develop a patient-specific 3D CNN reconstruction technique for use in real-time on-the-fly scenario for LMR adaptive radiotherapy purposes.

2. Materials and methods

2.1. Convolutional neural network

The original CNN code used for this study was written by Schlemper *et al* in Python using Theano and Lasagne packages (Schlemper *et al* 2018), that has been adapted for the use in real-time dynamic image reconstruction using 3D CNNs (Dietz *et al* 2019). A dynamic CNN network was proposed by Schlemper that aimed to reconstruct the entire temporal domain in spatial patches; however, the application to rapid real-time reconstruction was not addressed. Our adaptation allows for real-time on-the-fly reconstruction that incorporates both spatial and temporal data. We further adapted the CNN code to allow the use of coherently undersampled data.

The CNN image reconstruction (x_{cnn}) can be expressed as

$$x_{cnn} = f_{cnn}\left(\boldsymbol{x_u}|\vec{\theta}, \lambda, \boldsymbol{\Omega}\right),$$

where the term $\boldsymbol{x_u}$ is the zero-filled undersampled image and $f_{cnn}\left(\boldsymbol{x_u}|\vec{\theta}, \lambda, \boldsymbol{\Omega}\right)$ is the CNN reconstructed image. The function f_{cnn} attempts to reconstruct the undersampled image $\boldsymbol{x_u}$ and depends on the trained CNN parameters $\vec{\theta}$, a data fidelity weighting term λ, and the known k-space locations $\boldsymbol{\Omega}$. The data fidelity term λ is calculated such that $\lambda = q/\sigma$, where q is a hyper-parameter and σ^2 is the noise power. The λ term is initialized to 0.025 as it was empirically shown that this value performs well (Caballero *et al* 2014). Given the fully sampled input training data **D**, the CNN is trained to reconstruct the data by minimizing the following cost function,

$$Z(\theta) = \sum_{(\boldsymbol{x}_u, \boldsymbol{x}_{gnd}) \in D} \|\boldsymbol{x}_{gnd} - \boldsymbol{x}_{cnn}\|_2^2.$$

The terms x_{gnd} and $x_{cnn} = f_{cnn}\left(\boldsymbol{x_u}|\vec{\theta}, \lambda, \boldsymbol{\Omega}\right)$ are the ground truth image and the CNN reconstructed images (from the undersampled ground truth data), respectively. Once trained, the parameters can then be rapidly applied directly to the undersampled data (either prospectively or retrospectively). The Adam optimizer was implemented for the CNN training (Kingma *et al* 2014).

The CNN is comprised of a deep network of alternating convolutions layers and data consistency layers. The convolutional layer convolves filters with the undersampled data in order to inhibit the aliasing artefacts or low-resolution blur corresponding to the raw undersampled data. The convolutional layer applies N convolutions sequentially, which are each followed by the rectifier linear units (ReLU) nonlinearity activation function. The final N^{th} convolution output is summed with the initial input data, via the process known as residual connection (He *et al* 2016). Following the convolutional layer the data is passed to the consistency layer, which enforces data fidelity by ensuring that the acquired portion of k-space is always incorporated into the 3D CNN. The data consistency step takes the image x_{cnn} (output from the convolutional layer), which is Fourier transformed (FT) into k-space. If no k-space was initially acquired, it is set equal the value from the FT of x_{cnn}; however, if k-space was initially acquired then the value is equal to a weighted sum of the original k-space and the FT of x_{cnn}. The process of convolutions followed by data consistency continues for M data consistency layers.

2.2. CNN hyperparameter optimization

Finding the optimal hyperparameters for our CNN was done systematically by setting the number of convolutional layers, data consistency layers, and number of convolution filters. The CNN was trained for a reduced number of iterations (25), and the reconstructed data was quantitatively evaluated using metrics

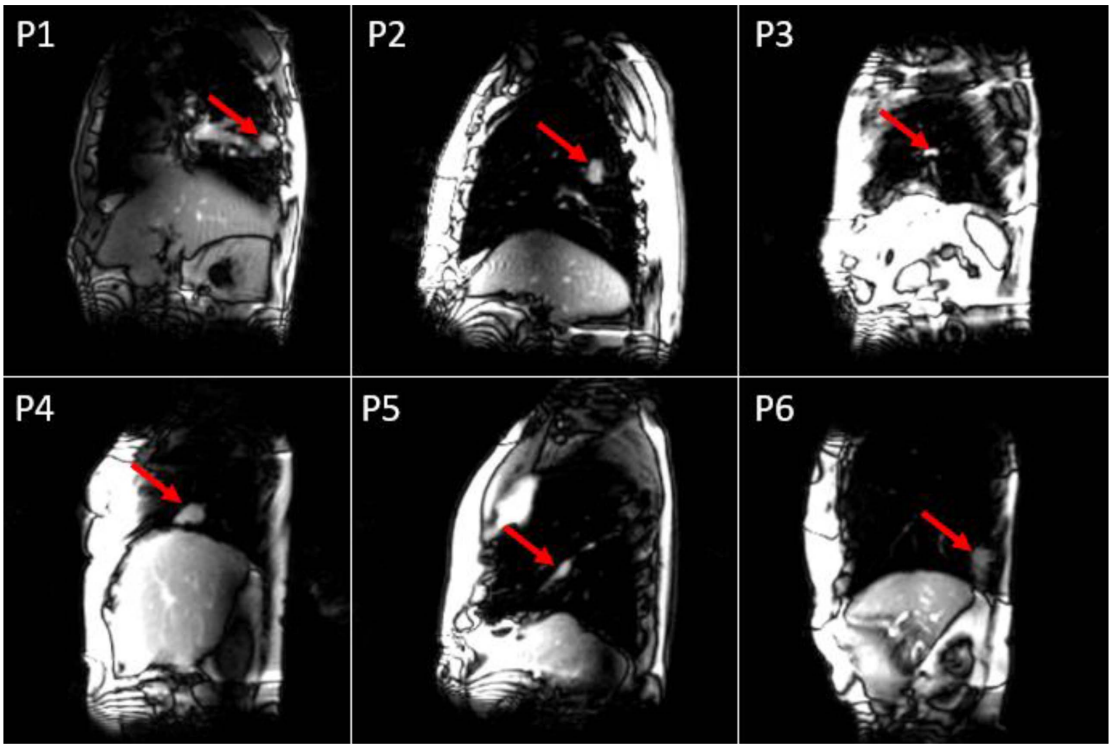

Figure 2. The six non-small cell lung cancer patients. These data sets consisted of 650 dynamic frames which were retrospectively undersampled to quantitatively evaluate and compare the incoherent and coherent-LR undersampled phase encoding schemes using our 3D CNN reconstruction technique.

Table 1. Tumor statistics for each of the six patients. The area contains the mean and standard deviation over the 650 fully sampled dynamic images. It is evident from the motion extents (superior-inferior and anterior-posterior) that the tumors ranged in size and motion.

Patient	Area (cm^2)	Sup-Inf Extent (cm)	Ant-Post Extent (cm)
1	2.57 (0.09)	1.97	0.40
2	3.76 (0.25)	2.63	0.75
3	1.36 (0.16)	1.57	0.66
4	4.96 (0.31)	0.94	0.39
5	2.09 (0.19)	1.07	0.72
6	4.99 (0.30)	0.87	0.17

described below. The trained CNN was then used to reconstruct the data set, which was evaluated using various imaging metrics (described in depth below). The systematic hyperparameter search was conducted for each retrospective patient. We evaluated the convolutional layers from two to six, data consistency layers from one to five, and the number of filters was set to either 16, 32, or 64. This resulted in a total of 75 sets of hyperparameters for each patient.

2.3. Tumor segmentation via autocontouring

An automatic contouring program previously developed by our group was used for the automatic segmentation of the lung tumors (Yun *et al* 2012, 2013, 2016). Using contours drawn on several phases of the breathing cycle as input data (a total of 30 images were used), the program trains a neural network to find the region of interest (tumor) and segment it. This contouring program was used for all segmentations performed in this study, which was trained using contours drawn by a radiation oncologist.

2.4. Retrospective data analysis

The acquisition of data from six non-small cell lung cancer patients was subject to approval by our institutional board of ethics. Lung cancer was chosen as it often involves large tumor motion from breathing, making it an ideal challenging case for real-time imaging and segmentation for LMR tumor tracking applications. Figure 2 contains images displaying the field of view (FOV) with the tumor present. Table 1 contains tumor specific information for each patient.

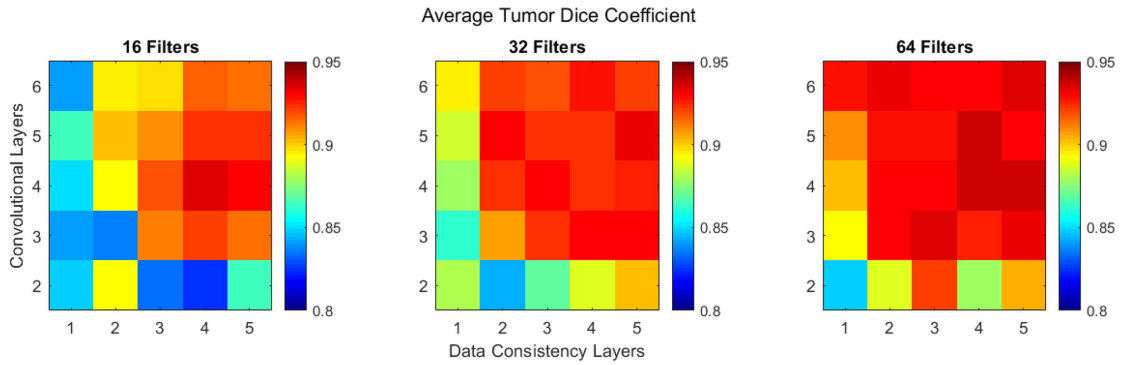

Figure 3. Dice coefficient results of the hyperparameter search. Based on the results we chose: 4 convolutional layers, 4 data consistency layers, and 16 filters. The reconstruction time increases as the number of filters, convolutional layers, or data consistency layers increase.

The retrospective data sets comprised of 650 dynamic frames acquired with a Cartesian bSSFP sequence using a multi-channel chest array. The data were acquired using our 3 T Philips Achieva system (Philips Medical Systems, Cleveland, OH, USA). The MR imaging parameters consisted of a 40 $\times$ 40 cm^2 FOV, a slice thickness of 2 cm, 128 $\times$ 128 voxels, a flip angle of 20^o, and TR/TE = 2.2/1.1 ms. The dynamic data were acquired at a rate of 4 frames per second. The data was split up such that the first 450 dynamic frames were used to train the CNN and the remaining 200 dynamic frames were used to quantitatively evaluate the CNN. In order to robustly train the CNN, a 5-fold cross validation was used on the training data portion (450 dynamic frames). Thus, each CNN k-fold iteration used 90 dynamic frames for test data and trained on the remaining 360 dynamic frames. The training data was mini-batched into groups of 5 images and shuffled to further increase the generalization of the model.

2.5. Prospective data analysis

The prospective data was acquired from a volunteer. The acquisition imaging parameters were identical to the retrospective case. However, here we have two separate acquisitions; a fully sampled acquisition (exact same parameters as the retrospective case) and an undersampled acquisition at 4x acceleration. The training of the CNN was conducted similarly to the retrospective case; except the training data comprised of 650 frames instead of the 450 frames as in the retrospective training (since the undersampled data is acquired separately). The undersampled acquisition contained the same parameters, except that only the central 32 lines of k-space were acquired, resulting in data of the size 32 $\times$ 128 pixels per dynamic frame. The remaining 96 unacquired lines of k-space were zero-filled (48 on either side of the acquired 32 lines), ultimately resulting in a matrix size of 128 $\times$ 128 voxels.

2.6. Quantitative analysis

The metrics we have chosen are specific for the analysis of tumor segmentations, being that our focus lies on the ability to segment a tumor for the purpose of real-time adaptive radiotherapy. The image quality, while an important aspect for image reconstruction, is secondary to the tumor segmentation.

2.6.1. Dice coefficient

The Dice coefficient is an ideal way to compare how well the tumor can be resolved in the CNN reconstructed data, by comparing the Boolean segmentation with the fully sampled segmentation (Dice 1945). This metric compares how well the two segmentations overlap. A DC is equal to unity if the two segmentations overlap entirely, meaning there is no depreciable difference in the image reconstruction using the CNN technique.

The Dice coefficient can be measured as,

$$DC = 2 \cdot \frac{ROI_{FS} \cap ROI_{US}}{ROI_{FS} + ROI_{US}}$$

where ROI_{FS} is the contour region of interest (ROI) for the fully sampled data and ROI_{US} is the contour for the undersampled CNN reconstructed data.

2.6.2. Centroid displacement

The centroid displacement offers another metric to evaluate how well the CNN reconstructed segmentation matches the fully sampled segmentation. The centroid displacement is simply the difference in the center of

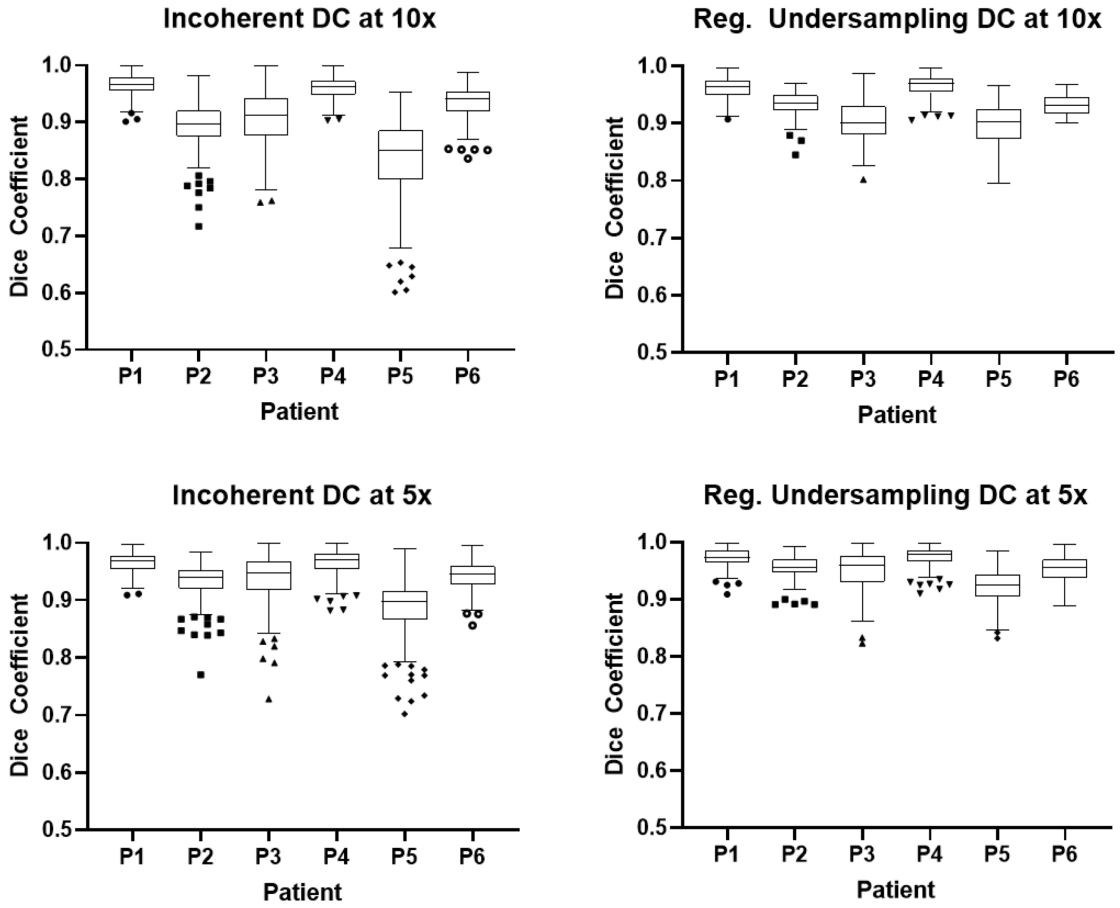

Figure 4. Tukey box plots containing the DC results from the 200 CNN reconstructed images using the incoherent and coherent-LR undersampled phase encoding schemes at 5x and 10x acceleration. It is evident that that the coherent-LR undersampling scheme for both 5x and 10x acceleration provide a larger mean DC, and less variation than the incoherent phase encoding.

masses between the two segmentations. The smaller the value, the closer the two overall segmentations overlap. This metric evaluated in concert with the DC gives a more complete picture of how well the CNN reconstruction is performing, for the purpose of segmentability.

2.6.3. Structural similarity index

The structural similarity index (SSIM) attempts to evaluate the perceived quality through the comparison of two data sets (Wang *et al* 2004). The SSIM requires the computation of three distinct terms: the luminance, contrast, and structural terms. Combining these terms, we get an expression

$$SSIM = \frac{(2\mu_{FS}\mu_{US} + C_1)(2\sigma_{FS,US} + C_2)}{(\mu_{FS}^2 + \mu_{US}^2 + C_1)(\sigma_{FS}^2 + \sigma_{US}^2 + C_2)}.$$

The terms μ_{FS}, μ_{US}, σ_{FS}, and σ_{US} are the means and standard deviation of the fully sampled and undersampled data respectively. The term $\sigma_{FS,\,US}$ is the cross-correlation of the fully and undersampled images. The constants C_1 and C_2 are based on the dynamic range (DR) of the two image sets; these values are typically set to $C_1 = (0.01 \cdot DR)^2$ and $C_2 = (0.03 \cdot DR)^2$.

3. Results

3.1. CNN hyperparameter search

An exploration of the optimal CNN hyperparameters was conducted. Figure 3 contains a visual representation of the searched hyperparameter space. Based on the results of the hyperparameter search we chose to use four convolutional layers, four data consistency layers, and 16 filters. This proved to be robust on all retrospective patients and had a short reconstruction time of 54 ms per dynamic frame. For the prospectively acquired volunteer data, we chose to use four convolutional layers and four data consistency

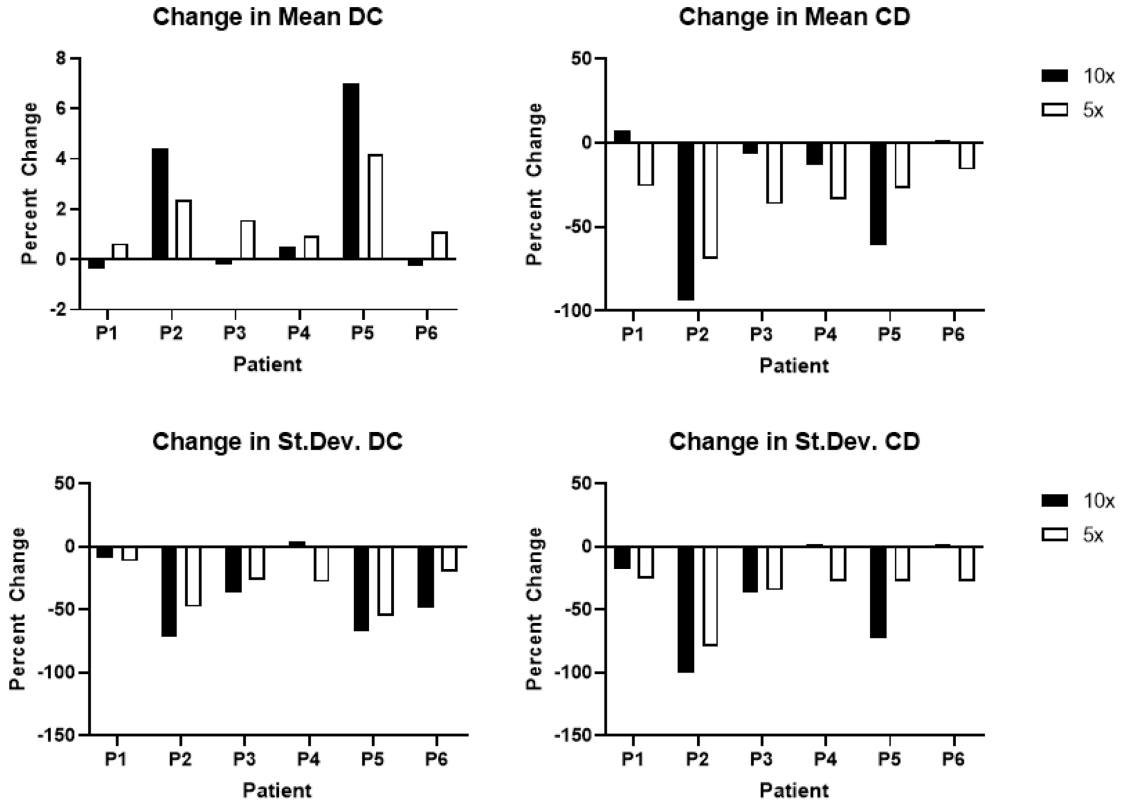

Figure 5. The percent change in the mean and standard deviation for the Dice coefficient (DC) and centroid displacement (from incoherent to coherent-LR). Nearly all patients had an increased DC when using of the coherent-LR phase-encoding scheme, with the exception of P1, P3, and P6 at 10x acceleration. The DC standard deviation decreased or remained unchanged for all patients except P4 at 10x acceleration. The mean CD decreased for all patients except for P1 and P6 at 10x acceleration. The CD standard deviation decreased for all patients, except for P4 and P6 which only very slightly increased.

layers; however, we chose 64 filters. While this resulted in a slightly increased reconstruction time, it provided a qualitatively better result (data not shown) than using 16 or 32 filters.

3.2. Retrospective CNN analysis

3.2.1. Dice coefficient

The DC was calculated for both the incoherent and coherent-LR undersampled phase encoding schemes at 5x and 10x acceleration for the six retrospective lung tumor. Figure 4 contains Tukey box and whisker plots of the DC, performed on the 200 dynamic frames that were not included in the training of the network.

The percent changes (from incoherent to coherent-LR) for the mean and standard deviations of the DC and CD are shown in figure 5. It is evident from figures 4 and 5 that the DC from the coherent-LR undersampled PE scheme performed better than the incoherent PE scheme with less variability and resulted in a higher mean value. Patients 2 and 5 had the greatest increase in mean DC (2.3% and 4.2% at 5x acceleration; 4.4% and 7.0% at 10x acceleration) and reduction in variance of the DC values (−48% and −55% at 5x acceleration; −71% and −67% at 10x acceleration); patients 1, 3, and 6 had a small decrease in mean DC (−0.4%, −0.2%, and −0.3% at 10x acceleration; respectively). However, the standard deviation for patient 1, 3, and 6 decreased by 8.6%, 36.5%, and 49.1% at 10x acceleration; respectively.

Furthermore, we have included an example of a poor segmentation for the most difficult patient (P5), as shown in figure 6 for both phase-encoding schemes. It is evident that a blurring of the tumorous region occurred, resulting in a reduced DC of 0.610 and 0.795 (at 10x acceleration) for the incoherent and coherent-LR, respectively. The reduced DC may be a potential limitation of the acceleration and segmentation itself; however, this example demonstrates how the coherent-LR improves upon the incoherent PE. This patient was particularly difficult to segment, given the long tumor shape and sporadic motion.

3.2.2. Centroid displacement

The centroid displacement (CD) was calculated for both the incoherent and coherent-LR undersampled phase encoding schemes at 5x and 10x acceleration for the six retrospective lung tumor patients for the 200 dynamic frames. Figure 5 contains the percent changes for the mean and standard deviation values for both phase-encoding schemes for each patient and acceleration factor. As shown in figure 7, the centroid

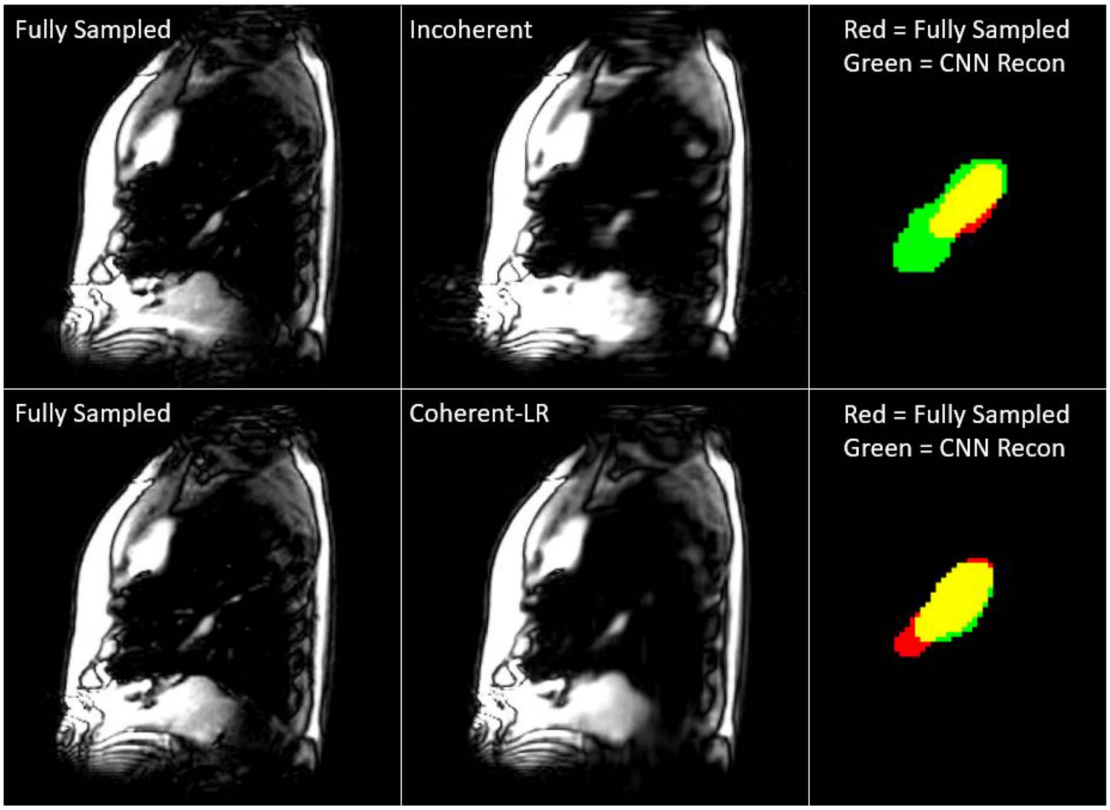

Figure 6. Example of the worst-case DC for patient 5 for the incoherent and coherent-LR phase-encoding with values of 0.610 and 0.795, respectively. It is evident for both PE schemes that some fine detail was lost; however, the incoherent had more significant blurring.

displacement is reduced for the coherent-LR undersampled case for both acceleration factors, with the exception of patients 1 and 6 at 10x acceleration (which can be seen in further detail via figure 5). The variation of the centroid displacement values was reduced, similarly as the DC values, suggesting the coherent-LR undersampled phase encoding scheme is superior to the incoherent phase encoding scheme for our purpose. Furthermore, the CD had a greater improvement using the coherent-LR phase-encoding scheme than the DC for both acceleration factors.

3.2.3. Structural similarity index

The SSIM was averaged over the 200 reconstructed frames at 5x and 10x acceleration for both phase encoding schemes. The mean and standard deviation are plotted in figure 8, which demonstrates that the coherent-LR image quality improved for all patients at both acceleration factors.

3.3. Prospective CNN analysis

Figure 9 contains spatial and temporal views of the CNN reconstructed volunteer data using the coherent-LR and incoherent undersampled phase encoding schemes. The fully sampled training data, and the prospectively acquired (zero-filled) data are also shown for comparison. While the CNN reconstruction removed much of the aliasing from the incoherently acquired data, there were still artefacts present, which manifests as a faint rippling throughout the image. The artefacts did not appear to be present in the coherent-LR reconstruction. This demonstrates that the CNN technique does in fact work with real world accelerated data that has been trained on a separate fully sampled data set (which is a different scenario than the retrospective case). This figure thus has only qualitative value, since there is no fully sampled 'ground-truth' data set to compare to.

4. Discussion

By reducing the imaging time through undersampling, a higher frame rate and reduced latency can be achieved, allowing for real-time adaptive radiotherapy treatment. The use of patient-specific trained CNNs can be interpreted as personalized healthcare for the patient. This differs from the common NN paradigm, where a generalized solution is sought to be applied to any new patient data.

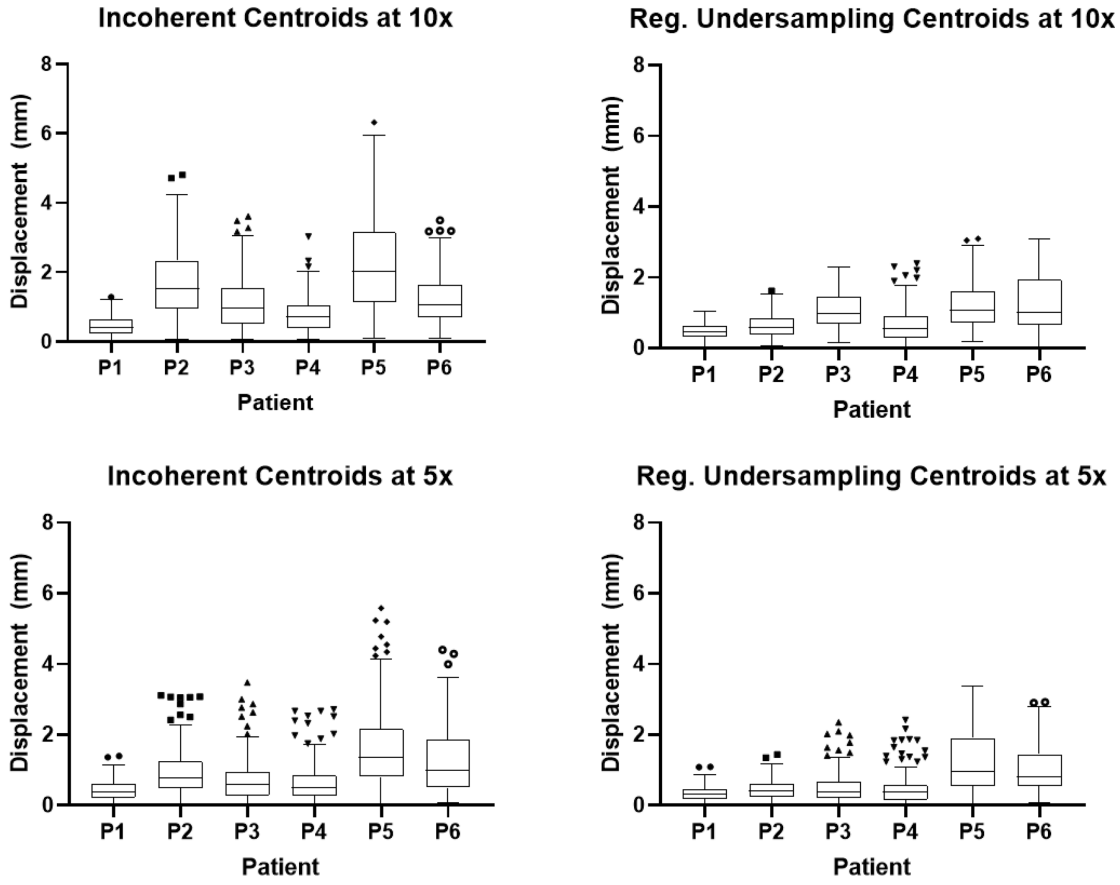

Figure 7. Tukey box plots containing the results of the centroid displacement between the CNN reconstructed data and the fully sampled data. The mean and variation of the segmented tumor was reduced when using the coherent-LR undersampling phase encoding scheme, compared to the incoherently undersampled data. It is evident that the coherent-LR (reg. undersampling) reduced the variance and outliers as comparted to the incoherent acquisition.

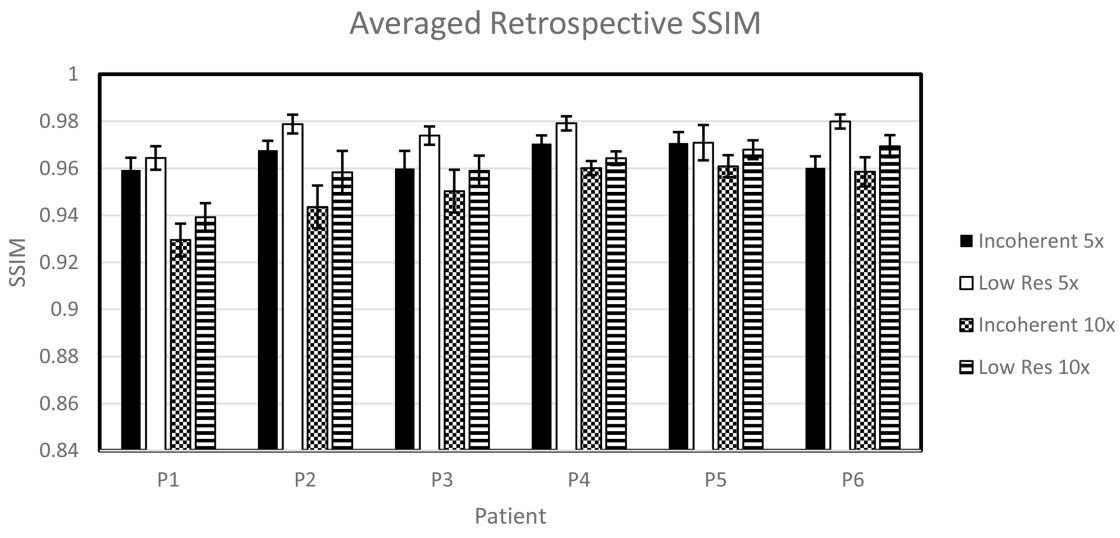

Figure 8. The SSIM showed improvement for all patients when using the coherent-LR undersampled phase encoding scheme.

For the purpose of the study, clinical interpretation of the real-time tracking is not the primary objective. Any required clinical interpretation would be achieved with other sequences in the planning or diagnostic stage of the patient's treatment. Further, delineation of the contours for training purposes would be performed on the fully-sampled images, not the CNN reconstructed data in our purposed workflow. As such, a successful match of the contour to the fully sampled case is the primary, or perhaps only requirement of these images.

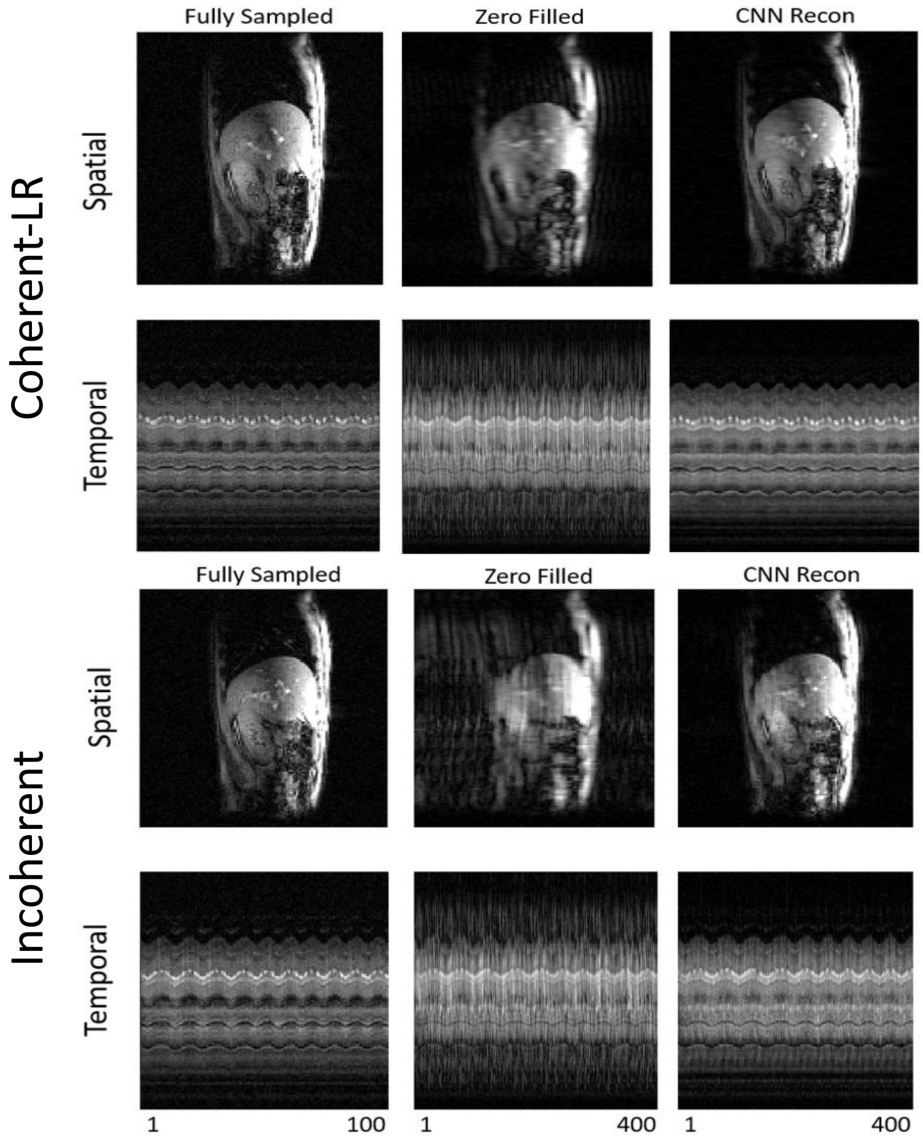

Figure 9. Prospective reconstruction results for the coherent-LR and incoherent undersampled CNN reconstructions of a volunteer data set. The first column contains the fully sampled (training) data, the second column is the prospectively acquired data prior to reconstruction (zero-filled), and the third column contains the CNN reconstructed data. For each sampling scheme, the top row contains a spatial view, while the bottom row contains a temporal view; note that there are four times more temporal frames for the prospectively acquired data due to the reduced acquisition time. It is evident that the artefacts have been removed from both the spatial and temporal views; however, the coherent-LR reconstruction produced a qualitatively better reconstruction.

As mentioned previously, the CNN training time is on the order of several hours; however, this does not pose any issue for clinical implementation, given that the pre-treatment training data would be acquired prior to the patients scheduled treatment. Patients undergoing external beam radiotherapy will generally receive a pre-treatment CT for the dosimetrist to generate patient contours and plan how the radiation will be delivered by the linac (or LMR). Institutes treating with LMR systems will also acquire pre-treatment MRI data. This previously acquired MRI data could be utilized to train a patient-specific CNN prior to the scheduled treatment, which can then be used to aid in reconstructing rapidly acquired (prospectively) undersampled data during treatment. Because the MRI data would be acquired prior to the scheduled treatment (often days prior), the training time is generally not much of concern; however, the reconstruction time must be fast enough to facilitate real-time on-the-fly imaging.

The benefits to using a coherent-LR undersampled acquisition is that it is simple to implement on an MRI scanner and has reduced artefacts as compared to the incoherent sampling scheme. We have shown that the realistic prospective imaging scenario qualitatively appears superior when using the coherent-LR undersampled acquisition (compared to the incoherent undersampling). The coherent-LR undersampling does result in a slight blurring of the reconstructed image; however, we believe that the segmentation program handles this case better than the incoherent noise-like artefact. Furthermore, not all MR systems have the software capability to implement an incoherent acquisition.

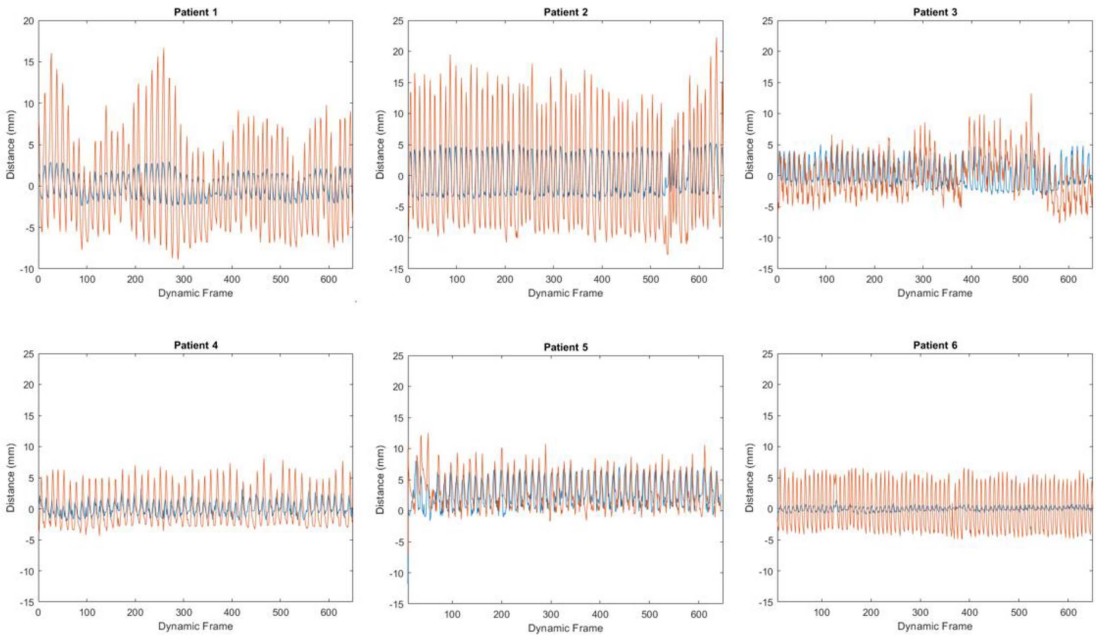

Figure 10. Centroid plots for the six patients lung tumors. The orange displays the superior-inferior motion, while the blue displays the anterior-posterior motion.

The use of low-resolution undersampling may evoke memories of keyhole imaging (van Vaals *et al* 1993), since only the central region of k-space is acquired during the real-time imaging portion of the exam. Indeed, both the incoherent and coherent-LR undersampling CNN strategies described in this work can be considered sophisticated versions of view-share, since patient-specific fully-sampled data is used to train a CNN to reconstruct a sub-sampling of that data during the real-time acquisition. However, unlike the keyhole strategy, where blocks of k-space from fully-sampled acquisitions are simply imported into newly-acquired frames, the trained CNN determines a custom k-space complement appropriate to the data acquired in the undersampled frame. This complement will be different for every frame, based only on patterns of correspondence seen in the training data.

Generally, the incoherent sampling scheme requires some correction for the increased artefacts cause by eddy currents, either pre or post reconstruction. For our incoherent acquisition we used a 'through-slice technique', where the slice preparation amplitude is increased. This creates a slight unbalancing of the slice-selection gradient in order to suppress the eddy-current contributions (Bieri *et al* 2005). When using the coherent-LR undersampling scheme, this is not a concern, given the coherent-LR sampling pattern does not require rapidly changing gradients.

The tumor specifics are stated in table 1; however, figure 9 displays the centroid motion for the six patients along the 650 fully sampled dynamic images. It can be seen patients 2, 4, 5, and 6 contain rhythmic motion for the most part; while patients 1 and 3 had tumor motion that changed during the imaging. As shown from figures 4 and 5, these patients did not correlate to a reduced DC or increased centroid displacement.

There are several limitations to the study presented. Our patient cohort was small with only six patients; however, the tumors varied in location, size, and motion. Further, our study only evaluated data acquired at one field strength (3 T); however, the scenario of having training data acquired by a different system than the treatment Linac-MR should be investigated as the contrast may differ between field strengths. While future studies will be required to assess varying contrast etc; as a proof of concept this study provides encouraging results, and a trend toward better segmentation when using coherent undersampling with the CNN. Regarding the SNR, we have previously demonstrated that the presence of increased noise does not decrease the accuracy of the CNN reconstruction (Dietz *et al* 2019). We only investigated one specific region for tumor tracking of lung tumors; however, as outlined in table 1, the tumors varied in motion and size. Figure 10 contains plots of the tumor centroid along the 650 fully sampled dynamic frames. In the future, various other regions should be investigated such as the liver or pancreas. Furthermore, the smallest tumor present in our data is 1.36 cm; future studies are required to determine the extent of how small the tumorous region can be, while still providing an accurate reconstruction (with either phase encoding scheme).

5. Conclusion

Incoherent undersampling results in a noise-like aliasing artefact, while coherent undersampling results in a low spatial resolution blurring. We investigated the suitability of using these phase-encoding schemes for MRI reconstruction using CNNs. In particular, we investigated the use for lung tumor segmentation for adaptive radiotherapy purposes using an MR-Linac system. We believe our results support the hypothesis that the CNN with coherent undersampling performs at least as well as, and often better than the CNN with the more typical incoherent undersampling, at least with regards to this particular real-time tracking scenario. We retrospectively evaluated the technique by comparing tumor segmentations of six non-small cell lung cancer patients. The coherent-LR undersampled acquisition demonstrated larger DC values, but more importantly, there was less variation in the DC across the temporal domain (200 dynamic frames). We applied the coherent-LR undersampled acquisition in a prospectively undersampled volunteer data set. The qualitative results from the prospective case demonstrate that the artefacts were removed, with only a slight blurring occurring throughout the reconstructed images. The reconstruction speed of the 3D CNN technique was 54 ms per image.

Acknowledgment

The authors would like to acknowledge Alberta Innovates Health Solutions (AIHS) for funding used in part, to support the research presented. We would also like to acknowledge Philips for their research support.

Disclosure

The authors are affiliated with MagnetTx Oncology Solutions, of which B G F is co-founder and chair.

References

Archarya S *et al* 2016 Online magnetic resonance image guided adaptive radiation therapy: first clinical applications *Int. J. Radiat. Oncol. Biol. Phys.* **94** 394–403

Bieri O, Markl M and Scheffler K 2005 Analysis and compensation of eddy currents in balanced SSFP *Magn. Reson. Med.* **54** 129–37

Caballero J, Price A N, Rueckert D and Hajnal J V 2014 Dictionary learning and time sparsity for dynamic MR data reconstruction *IEEE Trans. Med. Imag.* **33** 979–94

Dice L R 1945 Measure of the amount of ecologic association between species *Ecology* **26** 297–302

Dietz B, Yun J, Yip E, Gabos Z, Fallone B G and Wachowicz K 2019 Single patient convolutional neural networks for real-time MR reconstruction: a proof of concept application in lung tumor segmentation for adaptive radiotherapy *Phys. Med. Biol.* **64** 195002

Eppenhof K A J *et al* 2019 Fast contour propagation for MR-guided prostate radiotherapy using convolutional neural networks *Med. Phys.* **47** 1238–48

Fallone B G 2014 The rotating biplanar linac-magnetic resonance imaging system *Semin. Radiat. Oncol.* **24** 200–2

Fallone B G, Murray B, Rathee S, Stanescu T, Steciw S, Vidakovic S, Blosser E and Tymofichuk D 2009 First MR images obtained during mega-voltage photon irradiation from a integrated Linac-MR system *Med. Phys.* **36** 2084–8

Fischer-Valuck B W *et al* 2017 Two-and-a-half-year clinical experience with the world's first magnetic resonance image guided radiation therapy system *Adv. Radiat. Oncol.* **2** 485–93

Glockner J F, Houchun H H, Stanley D W, Angelos L and King K 2005 Parallel MR imaging: a user's guide *RSNA* **25** 1279–97

Griswold M A, Jakob P M, Heidemann R M, Nittka M, Jellus V, Wang J, Kiefer B and Haase A 1999 Generalized autocalibrating partially parallel acquisitions (GRAPPA) *Magn. Reson. Med.* **47** 1202–10

He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, 27–30 June 2016)* pp 770–8

Keall P J, Barton M and Crozier S 2014 The Australian magnetic resonance imaging-linac program *Semin. Radiat. Oncol.* **24** 203–6

Kingma D P and Ba J 2014 Adam: a method of stochastic optimization arXiv: 1412.6980

Kensuke U, Junko O and Takayuki I 2017 Application of super-resolution convolutional neural network for enhancing image resolution in chest CT *J. Digit. Imaging* **31** 441–50

Liney G P, Whelan B, Oborn B, Barton M and Keall P 2018 MRI-linear accelerator radiotherapy systems *Clin. Oncol.* **30** 686–91

Lustig M, Donoho D and Pauly J M 2007 Sparse-MRI: the application of compressed sensing for rapid MR imaging *Magn. Reson. Med.* **58** 1182–95

Mutic S and Dempsey J F 2014 The ViewRay system: magnetic resonance-guided and controlled radiotherapy *Semin. Radiat. Oncol.* **24** 196–9

Pruessmann K P, Weiger M, Scheidegger M B and Boesiger P 1999 SENSE: sensitivity encoding for fast MRI *Magn. Reson. Med.* **42** 952–62

Qin C, Schlemper J, Caballero J, Price A N, Hajnal J V and Rueckert D 2017 Convolutional recurrent neural networks for dynamic MR image reconstruction *IEEE Trans. Med. Imaging* **38** 280–90

Raaymakers B W *et al* 2009 Integrating a 1.5 T MRI scanner with a 6 MV accelerator: a proof of concept *Phys. Med. Biol.* **54** N229–37

Schlemper J, Caballero J, Hajnal J V, Price A N and Rueckert D 2018 A deep cascade of convolutional neural networks for dynamic MR image reconstruction *IEEE Trans. Med. Imaging* **37** 491–503

Valverde S, Cabezas M, Roura E, Gonzalez-Villa S, Pareto D, Vilanova J C, Ramio-Torrenta L, Rovira A and Oliver A 2017 Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach *Neuroimage* **15** 159–68

Van Vaals J J *et al* 1993 'Keyhole' method for accelerating imaging of contrast agent uptake *JMRI* **3** 671–5

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12

Yip E, Yun J, Wachowicz K, Heikal A A, Gabos Z, Rathee S and Fallone B G 2014 Prior data assisted compressed sensing: a novel MR imaging strategy for real time tracking of lung tumors *Med. Phys.* **41** 1–12

Yip E, Yun J, Wachowicz K, Gabos Z, Rathee S and Fallone B G 2017 Sliding window prior data assisted compressed sensing for MRI tracking of lung tumors *Med. Phys.* **44** 84–98

Yun J, Wachowicz K, Mackenzie M, Rathee S, Robinson D and Fallone B G 2013 First demonstration of intrafractional tumor-tracked irradiation using 2D phantom MR images on a prototype Linac-MR *Med. Phys.* **40** 051718

Yun J, Wachowicz K, Rathee S, Mackenzie M, Robinson D and Fallone B G 2012 Evaluation of a lung tumor autocontouring algorithm for intrafractional tumor tracking using low-field MRI: a phantom study *Med. Phys.* **39** 1481–94

Yun J, Yip E, Gabos Z, Wachowicz K, Rathee S and Fallone B G 2016 Improved lung tumor autocontouring algorithm for intrafractional tumor tracking using 0.5 T linac-MR *Biomed. Phys. Eng. Express* **2** 067004

Zeng K *et al* 2018 Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network *Comput. Biol. Med.* **14** 133–41