Motion Recovery from Radon Transformed Image Using Neural Networks

Editors: Under Review for MIDL 2020

Abstract

In this paper, we address motion correction in image reconstruction. Patient motion, including breathing, is a persistent problem in medical imaging. Cardiac motion is particularly enigmatic and often ignored, except in high-speed imaging modalities like MRI. Motions may create artifacts to the extent that the image may have to be discarded. Since the beginning of medical imaging, motion correction remained an important subcategory of research. Motion corrections may be applied during or after tomographic image reconstruction. In this work, we considered motion as a Gaussian blur at the image level. Discrete radon transform is applied to the blurred images to create corresponding noisy sinograms that mimic real imaging scenarios. Our deep learning-based tool recovered accurately (1) the blurring functions with an artificial convolutional neural network (CNN) directly from the sinograms, and (2) successfully reconstructed (inverse radon transformed) the noise-free images utilizing an adaptation of the convolutional encoder-decoder network (CED) from the literature. Our work shows that neural networks are not only capable of eliminating systematic noise in reconstruction but can also recover the noise model.

Keywords: Motion correction, motion modeling, image reconstruction, deep learning

1. Introduction

Motion is a source of major systematic noise in medical imaging. The thoracic motion caused by respiration results in motion in organs in the torso is somewhere between 5-25 mm (Dawood et al., 2006). Random patient motion is also a nuisance factor, particularly during a long scan. Cardiac motion is often beyond the temporal resolution limit of most medical imaging modalities and appears as a blur because of its lower spatial amplitude with respect to the size of the heart. These motions affect the usability of medical images by degrading the generated image quality (Faranesh et al., 2013).

Detecting noise generated from motion and applying correction at the time of image reconstruction has been attempted in the literature for a long time (Huang and Yu, 1992). Possibly, the most successful motion correction software was designed by the Artificial Intelligence in Medicine group from Cedars Sinai. Their tool, named *MoCo* (Matsumoto et al., 2001), embedded with tomographic image reconstruction algorithms, is routinely deployed by vendors. A similar software tool called *STASYS* (Bai et al., 2009) was subsequently reported in the literature to have better performance addressing cardiac motion blur before reconstruction, as our current work targets too. The *SinoCor* (Eiland et al., 2012), a patient motion correction tool for SPECT (with rotating camera head) before reconstruction and at the sinogram level, was developed by a different group. This included the motion blur borrowing from the Richardson-Lucy algorithm invented in the context of astronomy (Hanisch et al., 1996; Richardson, 1972; Lucy, 1974) and computational photography, to estimate the

convolution function causing motion-blur, while denoising the image simultaneously. The technique is very similar to our objective in this project. They used iterative statistical fixed-point algorithm, whereas we train deep neural networks for the same purpose.

King et al. (King et al., 2012) applied dimension reduction to approximately model respiratory motion, and Smith et al. (Smith et al., 2019) used a sparse approach in sampling to recover the motion model. These works evince the difficulty in motion capture even in modalities with high temporal resolution leading to sparse sampling.

In this work, we address motions, like that of cardiac motion, as symmetric Gaussian blur and try to recover that directly from the motion affected blurred sinogram data. For simplicity of fast experimentation, we used 2D synthetic data that may be easily extended to realistic clinical or pre-clinical 3D data. Rather, we used different shapes (motivated by that of the heart) and locations of the target object to prove the generalizability of our results. For the recovery of ground truth motion function and the image, we trained a convolutional neural network (CNN). Training phase needs large time but once trained, inferencing takes very little time to produce target results. We not only recovered the ground truth motion model (Gaussian function) accurately but also reconstructed the ground truth image from the noisy sinogram. Two different CNN models are trained for our twin objectives: motion model recovery and image reconstruction. Advantages from our method are: once trained, our model is agnostic to any modality, subject to the assumption that timing resolution of the modality is much lower compared to the motion frequency, which causes the motion to appear as a blur. Another advantage is that we recover an abstract mathematical model of the motion, the implication of which may be felt in diagnosis based on heart motion (Puyol-Anton et al., 2017). As mentioned before, this is similar to the Richardson-Lucy algorithm for motion recovery (Hanisch et al., 1996), but with CNN that has much faster inferencing time than the former iterative algorithm.

In the next section, we provide how we generate training data and validate our results. Actual methods of training and validation are described in section 3 and the results are provided in section 4. The last section provides concluding discussion and points to the future line of work in the project.

2. Data Generation

We first created four basic 2D shape-masks in 64x64 pixel window (fig 1), and then augmented them by changing the positions (translation) and orientations (rotation) of these images producing 2560 different images as ground truths. We used 8 different 2D 5×5 Gaussian filters to blur these images. Finally, the discrete Radon transform was applied to those blurred images to produce blurred sinograms. Subsequently, Poisson noise is added to the sinograms. Fig 2 shows this process. Totally, there are $2560 \times 8 = 20480$ blurred sinograms created. We randomly chose 20000 of them to be the training set, and the rest of them to be the validation set.

In order to simulate the shape of the heart, we created another data set which contains different annular elliptical rings. These images are generated by two different ellipse functions: the outer ellipse and the inner ellipse. By changing the center, the width, and the height, we got 640 different images (fig 3). Then we did the same augmentation operations as before for this dataset that is included for training and validation along with the previous

Figure 1: Four basic shapes.

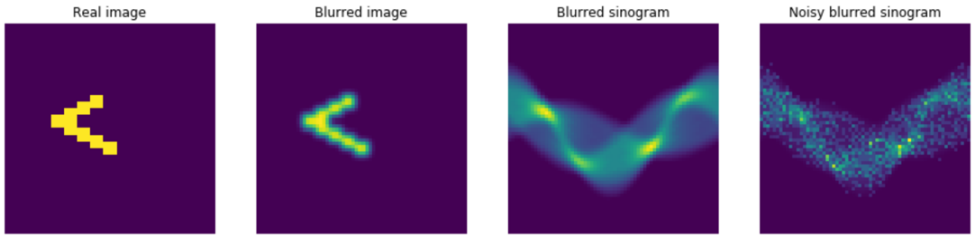

Figure 2: Steps to create a noisy blurred sinogram. (From left to right) Based on the real image, we used Gaussian filter to blur it to get a blurred image. Then, Radon transform this blurred image to create the blurred sinogram. Finally, by adding Poisson noise we get the noisy blurred sinogram that is used as input.

abstract shapes. The size of this new data set is 25600. We randomly chose 25000 of them to be the training set, and the rest of them to be the testing set.

Figure 3: Annular elliptical rings. They are generated by two ellipse functions. The outer ellipse is fixed. By changing the center, the width and the height of the inner ellipse, we get different images to simulate hearts.

3. Methods

Say, f is the Gaussian filter, g is the image, s is the sinogram, R is the Radon transform, and $*$ is the notation for the convolution operation, then the equation we used to generate the sinogram s is:

$$s = R(f * g) \tag{1}$$

Our problem is: given an s, can we recover f and g by training a deep learning model. We developed and trained two independent neural networks to: 1) learn to recover the filter f from a motion-blurred sinogram; and 2) learn to reconstruct the noise free image g from the motion-blurred sinogram.

3.1. Neural Network Model

3.1.1. Motion function recovery

A convolutional neural network (CNN) was used to recover the motion function. It was trained to extract the filter from the blurred sinogram. It contains 3 convolutional layers and two dense layers (fig 4). The number of the epoch is 20, batch size 50. The optimization function is Adam (Kingma and Ba, 2014), and the loss function is the mean squared error.

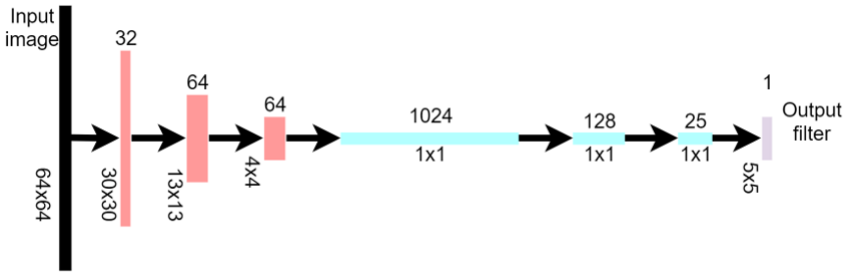

Figure 4: Convolutional neural network. It's used to extract the filter from the sinogram.

3.1.2. Image reconstruction from noisy blurred sinogram

For this experiment, we used the *convolutional encoder-decoder architecture* (CED, fig 5 (Häggström et al., 2019)). The number of the epoch is 100. Batch size is 50, optimization function is Adam, and the loss function is MSE.

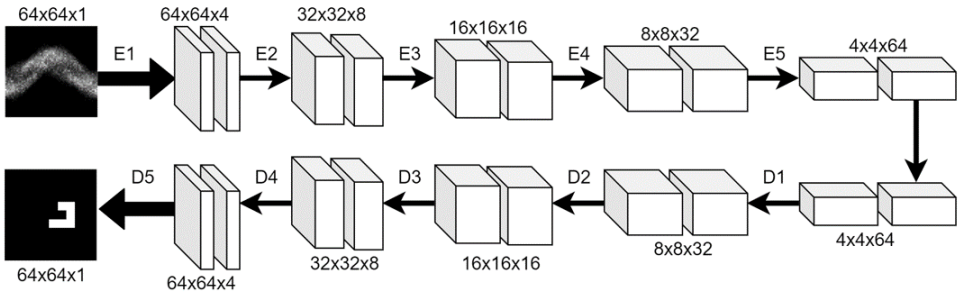

Figure 5: CED to reconstruct the image from a noisy blurred sinogram.

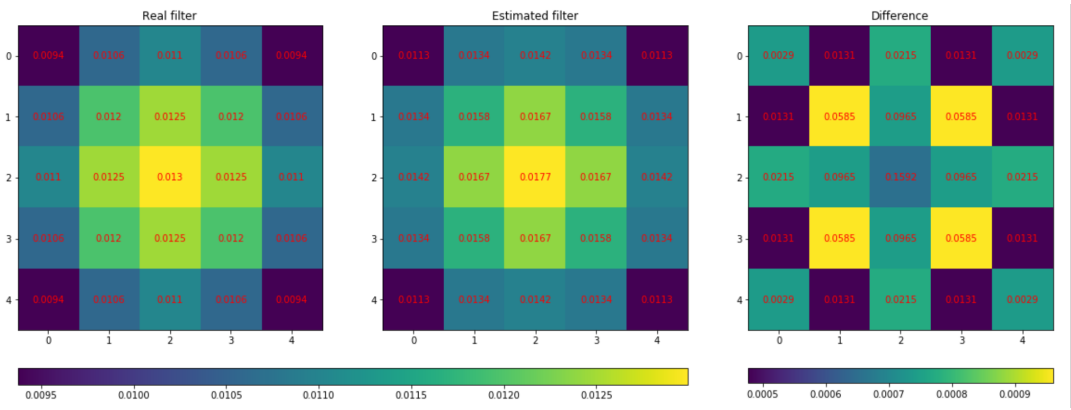

Figure 6: A sample result. Compare the real filter and the estimated filter. The third one shows the absolute difference between the real filter and the estimated filter.

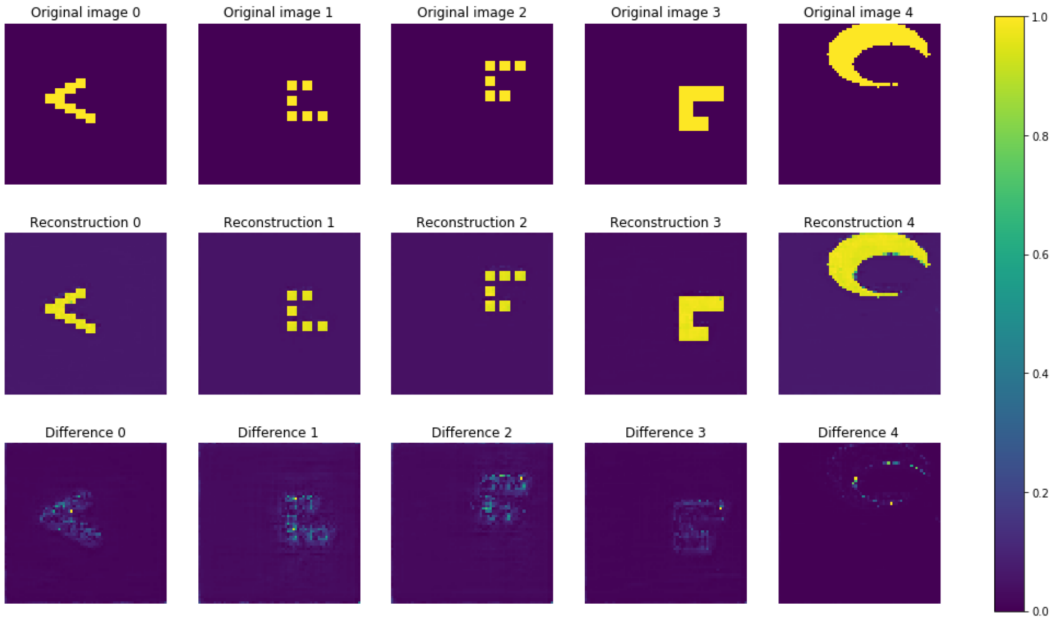

Figure 7: Image reconstruction with adapted CED. Annular elliptical ring on the last row simulates shapes of hearts. First row shows the ground truths. Second row shows the reconstructions. Third row is the difference between the ground truth images and the reconstruction images.

4. Results

4.1. Extract the filter from motion-blurred sinogram

We used mean squared error to measure the model loss. After 100 epochs, the training loss is 1.4249e-6 (RMS = 1.1937e-3), and the validation loss is 1.6612e-5 (RMS = 4.0758e-3). Each epoch of training takes approximately 3 sec on our machine (CPU is Intel(R) Core(TM) i7-9700K @ 3.60GHz. GPU is Nvidia Titan Xp. RAM is 32 GB). Fig 6 shows one sample result of the estimated filter (from the model output) compared to the ground truth filter.

4.2. Reconstruction from noisy motion-blurred sinogram

We trained the network on a GPU node on a cluster with the dataset which contains the ellipse rings. The configuration of this GPU node is: 2 x 10 core Intel Xeon @ 2.30GHz, 131GB of RAM, 4 x Nvidia Tesla K40m. After 100 epochs, the training loss is 4.0035e-04 (RMS = 0.02001), and the testing loss is 1.928e-4 (RMS = 0.01825). Training time is 24302.638 seconds, and the average reconstruction time per image is 2.710e-3 seconds.

5. Conclutions

According to our results, we believed the neural networks have a great potential to recover motion model and reconstruct images from motion-blurred sinograms. It can improve the reconstruction process in nuclear imaging such as PET, SPECT, and CT, for example, for noise reduction. It also provides a fast way to reconstruct the image in place of iterative reconstruction methods, though it takes much more time to train the network.

In order to avoid motion generated noise, respiratory and cardiac gating methodologies have been developed. They are quite useful for static imaging protocols with radiotracers and contrast agents, where tracer agent-concentrations are allowed to stabilize in body and imaging takes place after some wait time. However, in dynamic imaging protocols, which provide better quantitative and diagnostic information (El Fakhri et al., 2009), and where imaging starts immediately after injection to study the pharmacokinetics of the imaging agent, gating techniques are not as effective as in static imaging. Our proposed technique will be very useful in dynamic imaging as the trained neural network will incorporate the motion model. In the near future we want to develop such motion corrected dynamic image reconstruction deep learning models.

Finally, our technique needs to be validated with real data. Availability of large real training dataset is a challenge. We believe, combination of synthetic data and limited amount of available patient data will provide quality training. Another challenge is the adaptability of a model for individualized patient. However, motion model's topological shape is expected to be very similar for all patients. If we train our network to learn different convolution functions with different shapes, then such a trained model may be robust enough to recognize an individualized motion function and inference correctly. This is another direction where our near future efforts will be directed.

References

Chuanyong Bai, Jamshid Maddahi, Joel Kindem, Richard Conwell, Michael Gurley, and Rex Old. Development and evaluation of a new fully automatic motion detection and correction technique in cardiac spect imaging. *Journal of nuclear cardiology*, 16(4):580–589, 2009.

Mohammad Dawood, Norbert Lang, Xiaoyi Jiang, and Klaus P Schafers. Lung motion correction on respiratory gated 3-d pet/ct images. *IEEE transactions on medical imaging*, 25(4):476–485, 2006.

Daniel Eiland, Debasis Mitra, Mahmoud Abdalah, Rostyslav Buchko, and Grant T Gullberg. Sinocor: Inter-frame and intra-frame motion correction tool. In *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, pages 2963–2966. IEEE, 2012.

Georges El Fakhri, Arash Kardan, Arkadiusz Sitek, Sharmila Dorbala, Nathalie Abi-Hatem, Youmna Lahoud, Alan Fischman, Martha Coughlan, Tsunehiro Yasuda, and Marcelo F Di Carli. Reproducibility and accuracy of quantitative myocardial blood flow assessment with 82rb pet: comparison with 13n-ammonia pet. *Journal of Nuclear Medicine*, 50(7): 1062–1071, 2009.

Anthony Z Faranesh, Peter Kellman, Kanishka Ratnayaka, and Robert J Lederman. Integration of cardiac and respiratory motion into mri roadmaps fused with x-ray. *Medical physics*, 40(3):032302, 2013.

Ida Häggström, C Ross Schmidtlein, Gabriele Campanella, and Thomas J Fuchs. Deeppet: A deep encoder–decoder network for directly solving the pet image reconstruction inverse problem. *Medical image analysis*, 54:253–262, 2019.

Robert J Hanisch, Richard L White, and Ronald L Gilliland. Deconvolution of hubbles space telescope images and spectra. In *Deconvolution of images and spectra (2nd ed.)*, pages 310–360. 1996.

S-C Huang and D-C Yu. Capability evaluation of a sinogram error detection and correction method in computed tomography. *IEEE transactions on nuclear science*, 39(4):1106–1110, 1992.

Andrew P King, Christian Buerger, Charalampos Tsoumpas, Paul K Marsden, and Tobias Schaeffter. Thoracic respiratory motion estimation from mri using a statistical model and a 2-d image navigator. *Medical image analysis*, 16(1):252–264, 2012.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Leon B Lucy. An iterative technique for the rectification of observed distributions. *The astronomical journal*, 79:745, 1974.

Naoya Matsumoto, Daniel S Berman, Paul B Kavanagh, James Gerlach, Sean W Hayes, Howard C Lewin, John D Friedman, and Guido Germano. Quantitative assessment of motion artifacts and validation of a new motion-correction program for myocardial perfusion spect. *Journal of Nuclear Medicine*, 42(5):687–694, 2001.

Esther Puyol-Anton, Matthew Sinclair, Bernhard Gerber, Mihaela Silvia Amzulescu, Helene Langet, Mathieu De Craene, Paul Aljabar, Paolo Piro, and Andrew P King. A multimodal spatiotemporal cardiac motion atlas from mr and ultrasound data. *Medical image analysis*, 40:96–110, 2017.

William Hadley Richardson. Bayesian-based iterative method of image restoration. *JoSA*, 62(1):55–59, 1972.

Rhodri L Smith, Paul Dasari, Clifford Lindsay, Michael King, and Kevin Wells. Dense motion propagation from sparse samples. *Physics in Medicine & Biology*, 64(20):205023, 2019.